



**HAL**  
open science

## Stable Phase Retrieval with Mirror Descent

Jean-Jacques Godeme, Jalal Fadili, Claude Amra, Myriam Zerrad

► **To cite this version:**

Jean-Jacques Godeme, Jalal Fadili, Claude Amra, Myriam Zerrad. Stable Phase Retrieval with Mirror Descent. 2024. hal-04579446v2

**HAL Id: hal-04579446**

**<https://hal.science/hal-04579446v2>**

Preprint submitted on 20 Jun 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Stable Phase Retrieval with Mirror Descent

Jean-Jacques Godeme\*    Jalal Fadili\*    Claude Amra<sup>†</sup>    Myriam Zerrad<sup>†</sup>

**Abstract.** In this paper, we aim to reconstruct an  $n$ -dimensional real vector from  $m$  phaseless measurements corrupted by an additive noise. We extend the noiseless framework developed in [15], based on mirror descent (or Bregman gradient descent), to deal with noisy measurements and prove that the procedure is stable to (small enough) additive noise. In the deterministic case, we show that mirror descent converges to a critical point of the phase retrieval problem, and if the algorithm is well initialized and the noise is small enough, the critical point is near the true vector up to a global sign change. When the measurements are i.i.d. Gaussian and the signal-to-noise ratio is large enough, we provide global convergence guarantees that ensure that with high probability, mirror descent converges to a global minimizer near the true vector (up to a global sign change), as soon as the number of measurements  $m$  is large enough. The sample complexity bound can be improved if a spectral method is used to provide a good initial guess. We complement our theoretical study with several numerical results showing that mirror descent is both a computationally and statistically efficient scheme to solve the phase retrieval problem.

**Key words.** Phase retrieval, Noise, Stability, Inverse problems, Mirror descent.

## 1 Introduction

### 1.1 Problem statement and motivations

The phase retrieval problem arises in many applications including X-ray crystallography, diffraction imaging, and light scattering, to name just a few; see [24, 16, 19] and references therein. Phase retrieval is a very active research area and we refer to [24, 16, 19, 13, 28] for recent reviews of the current state of the art.

Our focus in this paper is phase retrieval with possibly noisy measurements. In real applications, the intensity measurements are not perfectly acquired. For instance, let us consider light scattering for precision in optics [1] which is our motivating application, where the goal is to describe the roughness of a polished surface. The latter is illuminated with a laser source, and the diffusion is measured by moving a detector. Then the power spectral density of the surface topography can be directly measured. However, during the acquisition process, different types of noise can corrupt the measurements such as photon noise, thermal noise, Johnson noise, *etc.*. Knowing the statistical model underlying the noise and the way it contaminates the measurements can prove useful to achieve robust reconstruction. The noise model can then be incorporated as the negative log-likelihood in the minimization objective. There are several noise models used in phase retrieval. One of them is the signal-dependent Poisson noise model which models the photon count noise. Another noise model is the (complex-valued) noise arising from multiple scattering, which can be modelled by the (complex) circularly-symmetric Gaussian distribution, and used to describe Rayleigh fading channels encountered in communication systems. Yet another source of noise is the thermal one or the incoherent background noise.

---

\*Normandie Univ, ENSICAEN, CNRS, GREYC, France. e-mail: jean-jacques.godeme@unicaen.fr, Jalal.Fadili@ensicaen.fr.

<sup>†</sup>Aix-Marseille Univ, CNRS, Centrale Marseille, Institut Fresnel, Marseille, France. `firstname.lastname@fresnel.fr`.

In this manuscript, and similarly to [8, 12, 9], we will work with a generic additive noise model, without any particular statistical assumption, in which the noisy phase retrieval problem reads

$$\begin{cases} \text{Recover } \bar{x} \in \mathbb{R}^n \text{ from the measurements } y \in \mathbb{R}^m \\ y[r] = |a_r^* \bar{x}|^2 + \epsilon[r], \quad r \in \llbracket m \rrbracket, \end{cases} \quad (\text{NoisyPR})$$

where  $[r]$  is the  $r$ -th entry of the corresponding vector, and  $\epsilon \in \mathbb{R}^m$  is the noise vector. Throughout the paper,  $A$  is the  $m \times n$  matrix with  $a_r^*$ 's as its rows.

Since  $\bar{x}$  is real-valued, the best one can hope for is to ensure that  $\bar{x}$  is uniquely determined from its intensities up to a global sign. Phase retrieval is in fact an ill-posed inverse problem in general, and even for  $\epsilon = 0$ , checking whether a solution to (NoisyPR) exists or not is known to be NP complete [23]. The situation is even more complicated in presence of noise. Thus, one of the major challenges is to design efficient recovery algorithms and find conditions on  $m$ ,  $(a_r)_{r \in \llbracket m \rrbracket}$  and  $\epsilon$  which guarantee stable recovery in presence of noise. This is the goal we pursue in this paper.

In this paper, we cast the noise-aware phase retrieval problem (NoisyPR) as the smooth but nonconvex minimization problem

$$\min_{x \in \mathbb{R}^n} \left\{ f(x) \stackrel{\text{def}}{=} \frac{1}{4m} \sum_{r=1}^m (y[r] - |(Ax)[r]|^2)^2 \right\}. \quad (1.1)$$

In fact, this is the same problem as in (3.1) studied in the noiseless case in [15]. There, we proposed a mirror descent algorithm based on a suitably chosen entropy. In particular, we analyzed the case where the measurements were either i.i.d standard Gaussian measurements or drawn from the Coded Diffraction Pattern (CDP) model. It is our aim in this paper to extend these results to the noisy case and prove stability guarantees for mirror descent to minimize (3.1).

## 1.2 Prior work

In phase retrieval, understanding the impact of noise is crucial because real-world measurements are invariably corrupted by it. Thus, establishing stability of phase retrieval to (small enough) noise is of paramount importance. Stability of phase retrieval to (small enough) noise has been studied by several authors with various measurement ensembles and reconstruction procedures. For convex relaxations, Candès and Li showed in [6] that a noise-aware variant of PhaseLift is stable against additive noise with a reconstruction error bound  $O\left(\|\bar{x}\|, \frac{\|\epsilon\|_1}{m\|\bar{x}\|}\right)$  as soon as  $m \gtrsim n$  (complex) Gaussian measurements are taken (see also [8] where the sample complexity was  $m \gtrsim n \log(n)$ ). This is of course only meaningful if the signal-to-noise ration is sufficiently high. This result has been extended to the case of sub-Gaussian measurements in [17]. For nonconvex formulations, Huang and Xu in [20] analyzed the performance of the Wirtinger flow and showed that any solution of this algorithm enjoys a reconstruction error upper bound  $O\left(\min\left\{\frac{\sqrt{\|\epsilon\|}}{m^{1/4}}, \frac{\|\epsilon\|}{\|\bar{x}\|\sqrt{m}}\right\}\right)$  as soon as  $m \gtrsim n$ . The amplitude, the reshaped and the Wirtinger flow algorithms are stable against additive noise as shown respectively in [32],[30] and [14]. Indeed, these authors showed that the reconstruction error scales as  $O\left(\frac{\|\epsilon\|}{\sqrt{m}}\right)$ . The convergence result is obtained under the specific assumption that  $\|\epsilon\|_\infty \lesssim \|\bar{x}\|$ . In [31], the authors study the performance of the amplitude-based model and showed the solution satisfies the following reconstruction upper bound  $O\left(\frac{\|\epsilon\|}{\sqrt{m}}\right)$  as soon as  $m \gtrsim n$ . The truncated Wirtinger flow [9], which can account even for Poisson noise, has been shown to be stable with a reconstruction error bound that scales as  $O\left(\frac{\|\epsilon\|}{\sqrt{m}\|\bar{x}\|}\right)$  under the assumption  $\|\epsilon\|_\infty \lesssim \|\bar{x}\|^2$  provided that  $m \gtrsim n$ . It was shown there that this is the best statistical guarantee any algorithm can achieve by deriving a fundamental lower bound on the minimax estimation error.

### 1.3 Contributions and relation to prior work

In this paper, we claim that mirror descent to solve (3.1) is stable against sufficiently small additive noise. This in turn provides recovery error bounds of the noisy phase retrieval problem (NoisyPR). In the deterministic case, we show that for almost all initializers, mirror descent converges to a critical point near the true vectors (up to sign ambiguity) where the objective has no direction of negative curvature. In the random case, we consider i.i.d Gaussian measurements, and in the regime where the signal-to-noise ratio is large enough (see Assumption (A.1)), we provide a complete geometric characterization of the landscape of the nonconvex objective provided that  $m \gtrsim n \log^3(n)$ . In turn, this allows us to describe the set of the critical points of  $f$  as the union of the strict saddle points and global minimizers of  $f$ . From this, we provide a global convergence to a point in  $\text{Argmin}(f)$ , which is near  $\bar{x}$  (up to sign ambiguity), as soon as the number of samples is large enough. If  $m \gtrsim n \log(n)$ , using a spectral initialization method, we provide a local convergence to a vector in the neighborhood of the target vector (up to sign ambiguity). By "near" we mean a reconstruction error that eventually scales as  $O\left(\frac{\|\epsilon\|}{\sqrt{m}\|\bar{x}\|}\right)$  which matches the minimax optimal bounded established in [9, Theorem 3]. Compared to the Wirtinger flow and variants, our algorithm, by adapting to the geometry, offers an easier and dimension-independent choice of the parameters (in fact one, the descent-step size), and has global convergence guarantees.

Our results can be easily extended to sub-Gaussian measurements with minor changes. The case where  $a_r$ 's are drawn from the CDP model is, however, far more challenging. Indeed, this model enjoys less randomness compared to the (sub-)Gaussian case and many of our arguments that require the uniformization of some bounds that are difficult to extend to the CDP model. Nevertheless, numerical experiments suggest that stable recovery still holds for our mirror descent algorithm with CDP measurements.

### 1.4 Outline of the paper

The rest of the paper is organized as follows. Section 2 gathers preliminary material necessary to our exposition. In Section 3, we recall the mirror descent algorithm with backtracking and establish its global and local convergence guarantees in the deterministic case. In Section 4, we sample complexity bounds with Gaussian measurements for our deterministic guarantees to hold with high probability. Section 5 describes the numerical experiments. The proofs of technical results are deferred to the appendix. In particular, Section C studies the landscape of the noise-aware objective with Gaussian measurements.

## 2 Preliminaries

### 2.1 Notations

We denote  $\langle \cdot, \cdot \rangle$  the scalar product and  $\|\cdot\|$  the corresponding norm.  $B(x, r)$  is the corresponding ball of radius  $r$  centered at  $x$  and  $\mathbb{S}^{n-1}$  is the corresponding unit sphere. For  $m \in \mathbb{N}^*$ , we use the shorthand notation  $\llbracket m \rrbracket = \{1, \dots, m\}$ . The  $i$ -th entry of a vector  $x$  is denoted  $x[i]$ . Given a matrix  $M$ ,  $M^\top$  is its transpose and  $M^*$  is its adjoint (transpose conjugate) when it is complex. Let  $\lambda_{\min}(M)$  and  $\lambda_{\max}(M)$  be respectively the smallest and the largest eigenvalues of  $M$ . For two real symmetric matrices  $M$  and  $N$ ,  $M \succeq N$  if  $M - N$  is positive semidefinite.

$\text{dom}(f)$  is the domain of the function  $f$ .  $f^*$  denotes the Legendre-Fenchel conjugate of  $f$ . Recall that the set of critical points of a differentiable function  $f$  is  $\text{crit}(f) = \{x \in \mathbb{R}^n : \nabla f(x) = 0\}$ .

Let us denote the set of true vectors by  $\bar{\mathcal{X}} = \{\pm \bar{x}\}$ . For any vector  $x \in \mathbb{R}^n$ , the distance to the set of true vectors is

$$\text{dist}(x, \bar{\mathcal{X}}) \stackrel{\text{def}}{=} \min(\|x - \bar{x}\|, \|x + \bar{x}\|). \quad (2.1)$$

Our limitation of the ambiguity set to  $\{\pm\bar{x}\}$  may appear restrictive since even for real vectors, the equivalence class is much larger than what we are allowing. However, our restriction will be justified in the oversampling regime.

## 2.2 Bregman toolbox

**Definition and properties** Since our focus is on phase retrieval, we will restrict our discussion here to entropy/kernel functions  $\phi$  that are strictly convex and differentiable on the whole space. As such,  $\phi$  is in fact a Legendre function; see [22, Chapter 26]. Observe that following [22, Theorem 26.5], a function is Legendre if and only if its Fenchel conjugate  $\phi^*$  is Legendre. Moreover  $\nabla\phi$  is a bijection from  $\text{int}(\text{dom}(\phi)) = \mathbb{R}^n$  to  $\text{int}(\text{dom}(\phi^*))$  with  $\nabla\phi^* = (\nabla\phi)^{-1}$ .

For any function differentiable function  $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$ , its Bregman divergence is

$$D_\phi(x, z) = \phi(x) - \phi(z) - \langle \nabla\phi(z), x - z \rangle. \quad (2.2)$$

This proximity measure is not a distance (it is not symmetric in general for instance nor it verifies the triangle inequality). If  $\phi(x) = \frac{1}{2} \|x\|^2$ , the Bregman divergence is the usual euclidean distance  $D_\phi(x, z) = \frac{1}{2} \|x - z\|^2$ .

Throughout the rest of the work, we use the following properties of the Bregman divergence.

### Proposition 2.1. (Properties of the Bregman distance)

- (i) *The Bregman divergence of  $\phi$  is nonnegative if and only if  $\phi$  is convex. If in addition,  $\phi$  is strictly convex, then its Bregman divergence vanishes if and only if its arguments are equal.*
- (ii) *Linear additivity: for any  $\alpha, \beta \in \mathbb{R}$  and any differentiable functions  $\phi_1$  and  $\phi_2$  we have*

$$D_{\alpha\phi_1 + \beta\phi_2}(x, z) = \alpha D_{\phi_1}(x, z) + \beta D_{\phi_2}(x, z), \quad (2.3)$$

for all  $x, z \in \mathbb{R}^n$ .

- (iii) *The three-point identity: For any  $x, u, z \in \mathbb{R}^n$ , we have*

$$D_\phi(x, z) - D_\phi(x, u) - D_\phi(u, z) = \langle \nabla\phi(u) - \nabla\phi(z), x - u \rangle. \quad (2.4)$$

- (iv) *Suppose that  $\phi$  is also  $C^2$  and  $\nabla^2\phi(x)$  is positive definite for any  $x$ . Then for every convex compact subset  $\Omega \subset \mathbb{R}^n$ , there exists  $0 < \theta_\Omega \leq \Theta_\Omega < +\infty$  such that for all  $x, z \in \Omega$ ,*

$$\frac{\theta_\Omega}{2} \|x - z\|^2 \leq D_\phi(x, z) \leq \frac{\Theta_\Omega}{2} \|x - z\|^2. \quad (2.5)$$

**Regularity of functions** The following definition extends the classical gradient Lipschitz continuity property to the Bregman setting. This notion is coined "relative smoothness" and is important for the analysis of optimization problems with objective functions that are differentiable but lack the popular gradient Lipschitz-smoothness. The earliest reference to this notion can be found in an economics paper [4] where it is used to address a problem in game theory involving fisher markets. Later on it was developed in [3, 5] and then in [18], although first coined relative smoothness in [18].

**Definition 2.2. ( $L$ -relative smoothness)** Let  $\phi$  and  $g$  be differentiable functions.  $g$  is called  $L$ -smooth relative to  $\phi$  if there exists  $L > 0$  such that  $L\phi - g$  is convex on  $\mathbb{R}^n$ , i.e. for all  $x, z \in \mathbb{R}^n$ ,

$$D_g(x, z) \leq LD_\phi(x, z). \quad (2.6)$$

When  $\phi$  is the energy entropy, i.e.  $\phi = \frac{1}{2} \|\cdot\|^2$ , (2.6) is nothing but a manifestation of the standard descent lemma implied by Lipschitz continuity of the gradient of  $g$ .

In a similar way, we also extend the standard local strong convexity property to a relative version w.r.t to an entropy or kernel  $\phi$ .

**Definition 2.3. (Local relative strong convexity)** Let  $\mathcal{C}$  be a non-empty subset of  $\mathbb{R}^n$  and  $\phi$  and  $g$  be differentiable functions. For  $\sigma > 0$ , we say that  $g$  is  $\sigma$ -strongly convex on  $\mathcal{C}$  relative to  $\phi$  if

$$D_g(x, z) \geq \sigma D_\phi(x, z) \quad \forall (x, z) \in \mathcal{C}^2. \quad (2.7)$$

When  $\mathcal{C} = \mathbb{R}^n$ , we recover the notion of global relative strong convexity. If  $\phi$  is the energy entropy (i.e.  $\phi = \frac{1}{2} \|\cdot\|^2$ ), one recovers the standard definition of (local/global) strong convexity.

The idea of global relative strong convexity has already been used in the recent literature, see e.g. [27, Proposition 4.1] and [2, Definition 3.3]. Its local version was first proposed in [25]. Relation of global relative strong convexity to gradient dominated inequalities, which is an essential ingredient to prove global linear convergence of mirror descent, was studied in [2, Lemma 3.3].

We record the following simple lemma which will be useful to compare Bregman divergences of smooth functions through the partial Loewner ordering of their respective Hessians.

**Lemma 2.4.** *Let  $g, \phi \in C^2(\mathbb{R}^n)$ . If  $\forall u \in \mathbb{R}^n, \nabla^2 g(u) \preceq \nabla^2 \phi(u)$  for all  $u$  in the segment  $[x, z]$ , then,*

$$D_g(x, z) \leq D_\phi(x, z). \quad (2.8)$$

**Proof.** The result comes from the Taylor-MacLaurin expansion. Indeed we have  $\forall x, z \in \mathbb{R}^n$

$$\begin{aligned} D_g(x, z) &= g(x) - g(z) - \langle \nabla g(z), x - z \rangle \\ &= \int_0^1 (1 - \tau) \langle x - z, \nabla^2 g(z + \tau(x - z))(x - z) \rangle d\tau, \end{aligned}$$

and thus

$$\begin{aligned} D_\phi(x, z) - D_g(x, z) &= \\ &= \int_0^1 (1 - \tau) \langle x - z, (\nabla^2 \phi(z + \tau(x - z)) - \nabla^2 g(z + \tau(x - z)))(x - z) \rangle d\tau. \end{aligned}$$

The positive semidefiniteness assumption implies the claim.  $\square$

### 3 Deterministic Stable Recovery

#### 3.1 Mirror descent with backtracking

Observe that the objective in (NoisyPR) can be decomposed as

$$f(x) = \frac{1}{4m} \sum_{r=1}^m (|(Ax)[r]|^2 - |(A\bar{x})[r]|^2 - \epsilon[r])^2. \quad (3.1)$$

The objective  $f$  is  $C^2(\mathbb{R}^n)$  and nonconvex (in fact only weakly convex). Moreover, its gradient is not Lipschitz continuous. However, using the strongly convex entropy (see [15, Proposition 2.6] for its properties),

$$\psi(x) = \frac{1}{4} \|x\|^4 + \frac{1}{2} \|x\|^2, \quad (3.2)$$

$f$  turns out to be smooth relative to  $\psi$ .

**Lemma 3.1.** *Let  $f$  and  $\psi$  defined in (1.1)-(3.2).  $f$  is  $L$ -smooth relative to  $\psi$  on  $\mathbb{R}^n$  for any  $L \geq \frac{1}{m} \sum_{r=1}^m \|a_r\|^2 (3 \|a_r\|^2 + \|\epsilon\|_\infty)$ .*

See Section A.1 for the proof. This estimate of the modulus of relative smoothness  $L$  is crude and depends also on the noise. This estimate will be largely improved for Gaussian measurements.

This relative smoothness property is the key motivation behind considering the framework of mirror descent or Bregman gradient descent. The mirror descent scheme with backtracking is detailed in Algorithm 1.

---

**Algorithm 1:** Mirror Descent for Phase Retrieval

---

**Parameters:**  $0 < L_0 \leq L$  (see Lemma 3.1),  $\kappa \in ]0, 1[$ ,  $\xi \leq 1$ .

**Initialization:**  $x_0 \in \mathbb{R}^n$ ;

**for**  $k = 0, 1, \dots$  **do**

**repeat**

$\gamma_k = \frac{1-\kappa}{L_k}$ ;

$x_{k+1} = F(x_k) = \nabla\psi^*(\nabla\psi(x_k) - \gamma_k\nabla f(x_k))$ ;

$L_k \leftarrow L_k/\xi$ ;

**until**  $D_f(x_{k+1}, x_k) \leq \xi L_k D_\psi(x_{k+1}, x_k)$ ;

$L_{k+1} \leftarrow \xi L_k$ ;

**Output:**  $x_{k+1}$ .

---

The pair  $(f, \psi)$  defined in (1.1)-(3.2) satisfies [5, Assumptions A, B, C, D] and thus the mapping  $F$  in Algorithm 1 is well-defined and single-valued on  $\mathbb{R}^n$ . Moreover, the mirror step  $\nabla\psi^*(z)$  can be computed easily as  $\nabla\psi^*(z) = t^*z$ , where  $t^*$  is the unique real positive root of the third-order polynomial  $t^3 \|z\|^2 + t - 1 = 0$ ; see [15, Proposition 2.8].

### 3.2 Deterministic recovery guarantees by mirror descent

Before the deterministic result, we start by recalling the notion of strict saddles.

**Definition 3.2 (Strict saddle point).** A point  $x_* \in \text{crit}(f)$  is a strict saddle point of  $f$  if  $\lambda_{\min}(\nabla^2 f(x_*)) < 0$ . The set of strict saddle points of  $f$  is denoted  $\text{strisad}(f)$ .

We now claim that mirror descent is stable against additive noise, as demonstrated in the following theorem.

**Theorem 3.3.** Consider the noisy phase retrieval problem cast as (1.1). Let  $(x_k)_{k \in \mathbb{N}}$  be a bounded sequence generated by Algorithm 1. Then,

(i) the sequence  $(x_k)_{k \in \mathbb{N}}$  has a finite length, converges to a point in  $\text{crit}(f)$  and the values  $(f(x_k))_{k \in \mathbb{N}}$  are nonincreasing.

Take  $L_k = L, \forall k \geq 0$ . Then,

(ii) for Lebesgue almost all initializers  $x_0$ , the sequence  $(x_k)_{k \in \mathbb{N}}$  converges to a critical point which cannot be a strict saddle, i.e.  $x_k \rightarrow \tilde{x} \in \text{crit}(f) \setminus \text{strisad}(f)$ .

(iii) Assume that  $\text{Argmin}(f) \neq \emptyset$ . Let  $\rho, \sigma > 0$  such that  $\rho > \frac{\sqrt{2}\|\epsilon\|}{\sqrt{m\sigma}}$  and define the radius  $r \leq$

$\sqrt{\frac{\rho^2 - \frac{2\|\epsilon\|^2}{m\sigma}}{\max(\Theta(\rho), 1)}}$ . If the initial point  $x_0 \in B(\bar{\mathcal{X}}, r)$  and  $f$  is  $\sigma$ -strongly convex relative to  $\psi$  on  $B(\bar{\mathcal{X}}, \rho)$  then  $x_k \in B(\bar{\mathcal{X}}, \rho), \forall k \in \mathbb{N}$ , and

$$\text{dist}^2(x_k, \bar{\mathcal{X}}) \leq (1 - \gamma\sigma)^{k-1} \rho^2 + 2 \frac{\|\epsilon\|^2}{m\sigma}, \quad (3.3)$$

See Section A.2 for the proof.

Some remarks are in order.



**Remark 3.4.**

- Clearly, claim (i) suggests that even in the presence of the noise, any bounded sequence of Algorithm 1 will converge to a critical point of  $f$  with decreasing values. Let us observe that the sequence generated by our algorithm is bounded if for instance  $f$  is coercive, in which case  $\text{Argmin}(f)$  is also a non-empty compact set. This happens to be true when  $A$  is injective, *i.e.* in the oversampling regime as we will show in the random case.
- Claim (ii) shows that when the initial guess  $x_0$  is chosen according to a distribution that has a density *w.r.t* the Lebesgue measure with constant step-size, then the sequence generated by mirror descent converges to a critical point where  $f$  has no direction of negative curvature.
- Concerning our local results in claim-(iii), if mirror descent is well initialized *i.e.*, in a ball of sufficiently small radius  $r < \rho$  around the true vectors  $\bar{\mathcal{X}}$ , and if  $f$  is strongly convex relative to  $\psi$  on the larger ball  $B(\bar{\mathcal{X}}, \rho)$ , then all the iterates  $(x_k)_{k \in \mathbb{N}}$  will remain in  $B(\bar{\mathcal{X}}, \rho)$ . Moreover, the sequence  $(x_k)_{k \in \mathbb{N}}$  will converge to a critical point  $\tilde{x}$  obeying

$$\text{dist}(\tilde{x}, \bar{\mathcal{X}}) \leq \frac{\sqrt{2} \|\epsilon\|}{\sqrt{m\sigma}}.$$

## 4 Stable Recovery from Gaussian Measurements

The deterministic stable recovery results of Theorem 3.3(ii)-(iii) require for instance a local relative strong convexity condition around  $\pm \bar{x}$  and possibly a good enough initial guess. A natural question to ask is when these conditions hold true. It turns out that this is indeed the case in the oversampling regime with i.i.d Gaussian measurements, and if the noise is small enough. This section is devoted to rigorously show these statements.

We consider that the sensing vectors  $(a_r)_{r \in [m]}$  are drawn i.i.d from a real zero-mean standard Gaussian distribution. We also work under the following assumption on the noise  $\epsilon$ .

**Assumption (A.1).** Denote  $\tilde{\epsilon} = \frac{1}{m} \sum_{r=1}^m \epsilon[r]$ . Given  $\lambda \in ]0, 1[$ , we suppose that

$$0 \leq \frac{\tilde{\epsilon}}{\min(\|\bar{x}\|^2, 1)} < \lambda \text{ and } \|\epsilon\|_\infty \leq c_s \min(\|\bar{x}\|^2, 1),$$

$$\text{for some constant } c_s \in \left] 0, \frac{(1-\lambda) \|\bar{x}\| \sqrt{\lambda \min(\|\bar{x}\|^2, 1) - \tilde{\epsilon}}}{2\sqrt{6} \min(\|\bar{x}\|^2, 1)} \right]. \quad (4.1)$$

To get better understanding of this assumption, we observe that it implies that

$$\frac{\|\epsilon\|}{\sqrt{m} \|\bar{x}\|^2} \leq \frac{\|\epsilon\|_\infty}{\|\bar{x}\|^2} < \frac{(1-\lambda)\sqrt{\lambda}}{2\sqrt{6}} < 1.$$

On the other hand, for the observation model (NoisyPR) with i.i.d real Gaussian sensing vectors, the signal-to-noise ratio (SNR) is captured by

$$\text{SNR} \stackrel{\text{def}}{=} \frac{\sum_{r=1}^m |a_r^\top \bar{x}|^4}{\|\epsilon\|^2} \approx \frac{3m \|\bar{x}\|^4}{\|\epsilon\|^2}.$$

In other words, Assumption (A.1) amounts to imposing that the SNR is large enough, *i.e.*

$$\sqrt{\text{SNR}} \gtrsim \frac{6\sqrt{2}}{(1-\lambda)\sqrt{\lambda}}.$$



Let us also observe that Assumption (A. 1) imposes that the empirical mean  $\tilde{\epsilon}$  is non-negative. This is a practical assumption that is fulfilled in many applications and will be helpful to describe the landscape of the noise-aware objective for Gaussian measurements. However, it was not used to have the deterministic guarantees.

Some of our stable recovery guarantees will be local provided that Algorithm 1 is initialized with a good guess. For this, we will use a spectral initialization method; see for instance [7, 9, 21, 32, 30, 29]. The procedure consists of taking  $x_0$  as the leading eigenvector of a specific matrix as described in Algorithm 2.

---

**Algorithm 2: Spectral Initialization.**

---

**Input:**  $y[r], r = 1, \dots, m$

**Output:**  $x_0$

Set  $\lambda^2 = n \frac{\sum_r y[r]}{\sum_r \|a_r\|^2} = n \left( \frac{\sum_r \langle a_r, \bar{x} \rangle^2}{\sum_r \|a_r\|^2} + \frac{\sum_r \epsilon_r}{\sum_r \|a_r\|^2} \right);$

Take  $x_0$  the top eigenvector of  $Y = \frac{1}{m} \sum_{r=1}^m y[r] a_r a_r^\top$  normalized to  $\|x_0\| = \lambda$ .

---

To lighten notations and clarify our proof, we consider the following events on whose intersection our deterministic convergence result will hold with high probability. Let fix  $\varrho \in ]0, 1[$  and  $\lambda \in \left] \frac{1}{9\sqrt{2}}, 1 \right[$ .

- The event

$$\mathcal{E}_{\text{strictsad}} = \{\text{crit}(f) = \text{Argmin}(f) \cup \text{strisad}(f)\} \quad (4.2)$$

means that the set of critical points of the function  $f$  reduces to the set global minimizers of  $f$  and the set of strict saddle points.

- The event

$$\mathcal{E}_{\text{conH}} = \left\{ \forall x \in \mathbb{R}^n, \quad \|\nabla^2 f(x) - \mathbb{E}(\nabla^2 f(x))\| \leq \varrho \left( \|x\|^2 + \|\bar{x}\|^2 / 3 + \|\epsilon\|_\infty \right) \right\} \quad (4.3)$$

captures the deviation of the Hessian of  $f$  around its expectation.

- The event

$$\mathcal{E}_{\text{inj}} = \left\{ \forall x \in \mathbb{R}^n, \quad (1 - \varrho) \|x\|^2 \leq \frac{1}{m} \|Ax\|^2 \right\} \quad (4.4)$$

corresponds to the injectivity of the measurement operator  $A$ .

- Let us denote by  $\mathcal{E}_{\text{smad}}$  the event on which the function  $f$  is  $L$ -smooth relative to  $\psi$  with  $L = 3 + \tilde{\epsilon} + \varrho \max(\|\bar{x}\|^2 / 3 + 1, 1)$ .
- Let us define  $\rho = \frac{(1-\lambda)\|\bar{x}\|}{\sqrt{3}} > 0$  and  $\mathcal{E}_{\text{scvx}}$  is the event on which  $f$  is  $\sigma$ -strongly convex relative to  $\psi$  locally on  $B(\bar{\mathcal{X}}, \rho)$ , with  $\sigma = \lambda \min(\|\bar{x}\|^2, 1) - \tilde{\epsilon} - \varrho \max(\|\bar{x}\|^2 / 3 + \|\epsilon\|_\infty, 1)$ .
- We end up denoting

$$\mathcal{E}_{\text{conv}} = \mathcal{E}_{\text{strictsad}} \cap \mathcal{E}_{\text{conH}} \cap \mathcal{E}_{\text{inj}} \cap \mathcal{E}_{\text{smad}} \cap \mathcal{E}_{\text{scvx}}. \quad (4.5)$$

Our main result for Gaussian measurements is the following.

**Theorem 4.1.** Fix  $\lambda \in \left] \frac{1}{9\sqrt{2}}, 1 \right[$  and  $\varrho \in ]0, \frac{\lambda \min(\|\bar{x}\|^2, 1) - \tilde{\epsilon}}{2 \max(\|\bar{x}\|^2 / 3 + \|\epsilon\|_\infty, 1)} [$ . Let  $(x_k)_{k \in \mathbb{N}}$  be the sequence generated by Algorithm 1. Under Assumption (A. 1), let us define for any  $\kappa \in ]0, 1[$

$$\nu = \frac{(1 - \kappa) \left( \lambda \min(\|\bar{x}\|^2, 1) - \tilde{\epsilon} - \varrho \max(\|\bar{x}\|^2 / 3 + \|\epsilon\|_\infty, 1) \right)}{3 + \tilde{\epsilon} + \varrho \max(\|\bar{x}\|^2 / 3 + \|\epsilon\|_\infty, 1)} \in [0, 1[$$

and

$$\varsigma = \frac{2\sqrt{2} \|\epsilon\|}{\sqrt{m \left( c_s \min(\|\bar{x}\|^2, 1) - \tilde{\epsilon} \right)}}.$$

- (i) If the number of measurements  $m$  is large enough i.e.,  $m \geq C(\varrho)n \log^3(n)$ , then for almost all initializers  $x_0$  of Algorithm 1 with the step-size  $\gamma \equiv \frac{1-\kappa}{3+\tilde{\epsilon}+\varrho \max(\|\bar{x}\|^2/3+\|\epsilon\|_\infty, 1)}$ , we have

$$x_k \rightarrow x^* \in \text{Argmin}(f) \cap B(\bar{\mathcal{X}}, \varsigma)$$

and  $\exists K > 0$  such that for all  $k \geq K$ , we have

$$\text{dist}^2(x_k, \bar{\mathcal{X}}) \leq \frac{\|\bar{x}\|^2}{3}(1-\nu)^{k-K} + \varsigma^2. \quad (4.6)$$

Both above events hold with a probability at least  $1 - 2e^{-\frac{m(\sqrt{1+\varrho}-1)^2}{8}} - e^{-\Omega(m)} - 5e^{-\zeta n} - 4/n^2 - c/m$ , where  $c, \zeta$  are fixed numerical constants.

- (ii) Suppose that  $\varrho$  obeys (B.16). If  $m$  is such that  $m \geq C(\varrho, \|\epsilon\|_\infty)n \log(n)$ , and Algorithm 1 is initialized with the spectral method in Algorithm 2, then (4.6) holds for all  $k \geq K = 0$  with probability at least  $1 - 2e^{-\frac{m(\sqrt{1+\varrho}-1)^2}{8}} - e^{-\Omega(m)} - 5e^{-\zeta n} - 4/n^2$ , where  $\zeta$  is a fixed numerical constant.

The choice of parameters can be made easier to grasp when  $\|\bar{x}\| = 1$  as assumed in many works.

**Corollary 4.2.** Suppose that  $\|\bar{x}\| = 1$  and the noise is small enough. Fix  $\lambda \in \left] \frac{1}{9\sqrt{2}}, 1 \right[$  and  $\varrho \in \left] 0, \frac{\lambda-\tilde{\epsilon}}{2} \right[$ . Let  $(x_k)_{k \in \mathbb{N}}$  be the sequence generated by Algorithm 1 with the step-size  $\gamma \equiv \frac{1-\kappa}{3+\tilde{\epsilon}+\varrho}$ , where  $\kappa \in ]0, 1[$ . Then the statements of Theorem 4.1 hold true with

$$\nu = \frac{(1-\kappa)(\lambda-\tilde{\epsilon})}{3+\tilde{\epsilon}+\varrho} \quad \text{and} \quad \varsigma = \frac{2\sqrt{2}\|\epsilon\|}{\sqrt{m(c_s-\tilde{\epsilon})}}.$$

**Remark 4.3.**

- When the number of measurements is sufficiently large as in claim (i), the SNR is large enough and the initial point  $x_0$  is chosen randomly from a measure that has a density w.r.t Lebesgue measure, then mirror descent converges, eventually linearly, to an element of  $\text{Argmin}(f)$  which is within a factor of the noise level from  $\bar{\mathcal{X}}$ . The local convergence rate is dimension-independent. To the best of our knowledge, this is the first kind of results for the noisy phase retrieval problem.
- When the number of measurements is in the less demanding regime of the second claim, then mirror descent with spectral initialization again converges to a noise region around  $\bar{\mathcal{X}}$ .
- We recover the rate of [15, Theorem 3.2] in the noiseless case.
- In the normalized setting of Corollary 4.2, the convergence rate behaves as

$$(1-\nu) \leq \frac{2}{3} + O((1-\lambda) + \kappa + \tilde{\epsilon} + \varrho).$$

- It is important to observe that in the noisy case, the true vectors  $\pm \bar{x}$  are not even critical points of  $f$ . Nonetheless, Lemma A.2 will show that  $\pm \bar{x}$  are actually  $\frac{\|\epsilon\|^2}{m}$ -minimizers.

**Proof.**

- (i) We prove this claim by combining Theorem 3.3 and the characterization of the structure of  $\text{crit}(f)$  that we provide in Theorem C.3. For the moment, let us assume that the event  $\mathcal{E}_{\text{conv}}$  holds true.
- By construction,  $\mathcal{E}_{\text{conv}} \subset \mathcal{E}_{\text{inj}}$  which means that the operator  $A$  is injective showing the coercivity of the objective  $f$  which implies that the sequence  $(x_k)_{k \in \mathbb{N}}$  generated by Algorithm 1 is bounded.
  - From the event  $\mathcal{E}_{\text{smad}}$ , we deduce that the function  $f$  is  $L$ -smooth relative to  $\psi$  with  $L = 3 + \tilde{\epsilon} + \varrho \max(\|\bar{x}\|^2/3 + \|\epsilon\|_\infty, 1)$ . Since the initializer is chosen at random with a fixed stepsize Theorem 3.3-(i)(ii) guarantees that  $(x_k)_{k \in \mathbb{N}}$  converges to  $x^* \in \text{crit}(f) \setminus \text{strisad}(f)$  and  $(f(x_k))_{k \in \mathbb{N}}$  also converges to  $f(x^*)$ .

- The event  $\mathcal{E}_{\text{scvx}}$  shows that the function  $f$  is  $\sigma$ -strongly convex relative to  $\psi$  on  $B(\bar{\mathcal{X}}, \rho)$  with  $\sigma = \lambda \min(\|\bar{x}\|^2, 1) - \tilde{\epsilon} - \varrho \max(\|\bar{x}\|^2/3 + \|\epsilon\|_\infty, 1)$ . Given that Assumption (A. 1) holds, Corollary B.6 implies that  $\rho^2 - \frac{4\|\epsilon\|^2}{m\sigma} > 0$  where we recall that  $\rho = \frac{(1-\lambda)\|\bar{x}\|}{\sqrt{3}}$ . Let us denote  $r^2 = \frac{\rho^2 - \frac{4\|\epsilon\|^2}{m\sigma}}{\max(\Theta(\rho), 1)}$ .
- Moreover,  $\mathcal{E}_{\text{strictsad}}$  holds true, and thus  $\text{crit}(f) \setminus \text{strisad}(f) = \text{Argmin}(f)$ , which means that for almost all initalizers  $x_0$ ,

$$x_k \rightarrow x^* \in \text{Argmin}(f) \text{ and } f(x_k) - \min f \rightarrow 0. \quad (4.7)$$

We now claim that  $\exists K > 0$ , large enough, such that  $\forall k \geq K$ ,  $\text{dist}(x_k, \bar{\mathcal{X}}) \leq r$  which will allow to invoke Theorem 3.3(iii). By Lemma B.3, we have for any  $k \in \mathbb{N}$

$$\text{dist}(x_k, \bar{\mathcal{X}}) \leq \|x_k - x^*\| + 8 \frac{\|\epsilon\|}{\sqrt{m} \|\bar{x}\|}$$

with probability at least  $1 - e^{-\Omega(m)}$ . Since  $x_k \rightarrow x^*$ , there exists  $K$  large enough such that  $\forall k \geq K$

$$\text{dist}(x_k, \bar{\mathcal{X}}) \leq 9 \frac{\|\epsilon\|}{\sqrt{m} \|\bar{x}\|}$$

with the same probability. To conclude, it is then sufficient to show that

$$9 \frac{\|\epsilon\|}{\sqrt{m} \|\bar{x}\|} \leq r.$$

This is true for sufficiently high SNR, *i.e.* under our Assumption (A. 1). Therefore we deduce from Theorem 3.3-(iii) that the sequence  $(x_k)_{k \geq K} \in B(\bar{\mathcal{X}}, \rho)$  and for  $k \geq K$ ,

$$\begin{aligned} \text{dist}^2(x_k, \bar{\mathcal{X}}) &\leq \left(1 - \frac{(1-\kappa)\sigma}{L}\right)^{k-1} \rho^2 + \frac{4\|\epsilon\|^2}{m\sigma}, \\ &\leq (1-\nu)^{k-K} \rho^2 + 4 \frac{\|\epsilon\|^2}{m\sigma}, \\ &\leq (1-\nu)^{k-K} \rho^2 + \zeta^2, \end{aligned}$$

where we have used (B.12) *i.e.*,  $\sigma \geq \frac{c_s \min(\|\bar{x}\|^2, 1) - \tilde{\epsilon}}{2}$  which implies that  $\text{dist}(x_k, \bar{\mathcal{X}}) \leq \zeta$ .

Let us now compute the probability that  $\mathcal{E}_{\text{conv}}$  occurs. The events  $\mathcal{E}_{\text{smad}}$ ,  $\mathcal{E}_{\text{scvx}}$  are contained in  $\mathcal{E}_{\text{conH}}$ , see respectively Lemma B.4 and Lemma B.5. From Lemma B.2,  $\mathcal{E}_{\text{conH}}$  occurs with a probability  $1 - 5e^{-\zeta n} - 4/n^2 - 2e^{-\frac{m(\sqrt{1+\varrho}-1)^2}{8}}$  as soon as  $m \geq C(\varrho)n \log(n)$ . Besides, close observation of the Hessian concentration (noisy part) highlights that it implies the injectivity of the measurements thus  $\mathcal{E}_{\text{inj}}$  is also contained in  $\mathcal{E}_{\text{conH}}$ . Thanks to Theorem C.3, the event  $\mathcal{E}_{\text{strictsad}}$  holds with a probability  $1 - c/m$  as soon as  $m \geq C(\varrho)n \log^3(n)$ . Finally, we conclude with a union bound that  $\mathcal{E}_{\text{conv}}$  holds with a probability at least  $1 - 2e^{-\frac{m(\sqrt{1+\varrho}-1)^2}{8}} - 5e^{-\zeta n} - 4/n^2 - c/m$  ( $\zeta, c$  are a fixed numerical constant) for  $m \geq C(\varrho)n \log^3(n)$ , which complete the proof.

- (ii) By [15, Lemma B.4], the operator  $A$  is injective entailing by the coercivity of  $f$  that the sequence  $(x_k)_{k \in \mathbb{N}}$  is bounded. From Corollary B.6, when the signal-to-noise coefficient  $c_s$  satisfies (4.1),  $r$  is well-defined. Lemma B.8 shows that when  $\varrho$  obeys (B.16), the initial point  $x_0$  given by Algorithm 2 is in the right  $f$ -attentive topology at the distance at most  $r = \sqrt{\rho^2 - \frac{4\|\epsilon\|^2}{m\sigma}}$ . Thanks to Lemma B.5,  $\rho$  is the radius of the ball  $B(\bar{\mathcal{X}}, \rho)$  where we have  $\sigma$ -strong convexity relative to  $\psi$  with  $\sigma = \lambda \min(\|\bar{x}\|^2, 1) - \tilde{\epsilon} - \varrho \max(\|\bar{x}\|^2/3 + \|\epsilon\|_\infty, 1)$ . The last point to check before

applying Theorem 3.3 comes from Lemma B.4 which shows that  $f$  is  $L$ -smooth relative to  $\psi$  with  $L = 3 + \tilde{\epsilon} + \varrho \max(\|\bar{x}\|^2/3 + \|\epsilon\|_\infty, 1)$ . We deduce from Theorem 3.3 that

$$\begin{aligned} \text{dist}^2(x_k, \bar{\mathcal{X}}) &\leq \left(1 - \frac{(1 - \kappa)\sigma}{L}\right)^{k-1} \rho^2 + \frac{4\|\epsilon\|^2}{m\sigma}, \\ &\leq (1 - \nu)^{k-K} \rho^2 + 4\frac{\|\epsilon\|^2}{m\sigma} \\ &\leq (1 - \nu)^{k-K} \rho^2 + \zeta^2. \end{aligned}$$

Let us observe that this statement is true only on the intersection of the above events, we call it  $\mathcal{E}'_{\text{conv}}$ . Let now compute the probability that  $\mathcal{E}'_{\text{conv}}$  occurs. We conclude with an appropriate union bound similar to the previous one, taking into account the fact that the spectral initialization event is contained in  $\mathcal{E}_{\text{conH}}$ . Finally, the statement holds with a probability at least  $1 - 2e^{-\frac{m(\sqrt{1+\varrho}-1)^2}{8}} - 5e^{-\zeta n} - 4/n^2$  ( $\zeta$  is a fixed numerical constant) for  $m \geq C(\varrho)n \log(n)$ , which completes the proof.  $\square$

## 5 Numerical experiments

In this section, we discuss some experiments to illustrate and validate numerically the efficiency of our phase recovery algorithm. In each instance, we measured the relative error between the reconstructed vector  $\tilde{x}$  and the true signal one  $\bar{x}$  as

$$\frac{\text{dist}(\tilde{x}, \bar{\mathcal{X}})}{\|\bar{x}\|}. \quad (5.1)$$

In the experiments, we set  $\|\bar{x}\| = 1$  and  $\tilde{x}$  was the output of Algorithm 1 at iteration  $K$  large enough.

### 5.1 Experiments with Gaussian sensing vectors

#### 5.1.1 Reconstruction of 1D signals

The aim is to reconstruct a randomly generated one-dimensional signal of length  $n = 128$  from  $m$  noisy observations where the sensing vectors were drawn i.i.d from the standard Gaussian ensemble. The noise vector  $\epsilon$  is chosen uniform such that  $\tilde{\epsilon} = 10^{-5}$ .

In Figure 1, the blue line shows the evolution of the objective (left) and the relative error (right) using Algorithm 1 with  $m = 128 \times \log^2(128)$ , where the initial point was drawn from the uniform distribution. The red line is the result of reconstruction from  $m = 5 \times 128 \times \log(128)$  measurements where Algorithm 1 was initialized using the spectral method in Algorithm 2 (the top eigenvalue was computed with the power iteration method using 200 iterations). In both cases, we run Algorithm 1 with a constant step-size  $\gamma = \frac{0.99}{3+10^{-5}}$  (see Theorem 4.1). As predicted by the latter, the curves in blue (*i.e.* with random initialization) have two regimes: a sublinear regime and then a local linear regime. Moreover, with spectral initialization (red curves),  $f$  and the relative error converge linearly at the same rate as in the local regime of the blue curves, hence confirming our theoretical findings. As anticipated also, both  $f$  and the relative error eventually stabilize at a plateau whose level is governed by the noise level.

#### 5.1.2 Comparison with the Wirtinger flow

We compared our mirror descent algorithm (with and without spectral initialization) with the Wirtinger flow [7]. However, the Polyak subgradient method proposed in [11], that we included in our comparison in [15], is only applicable to the noiseless case as it needs the value of  $\min f$  which is no longer known in

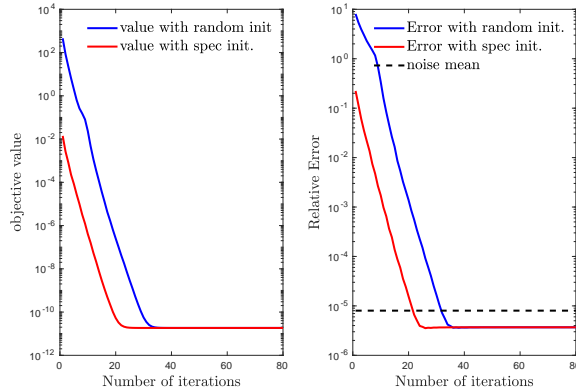


Figure 1: Reconstruction of signal from Gaussian measurements. The noise mean is  $\tilde{\epsilon}$ .

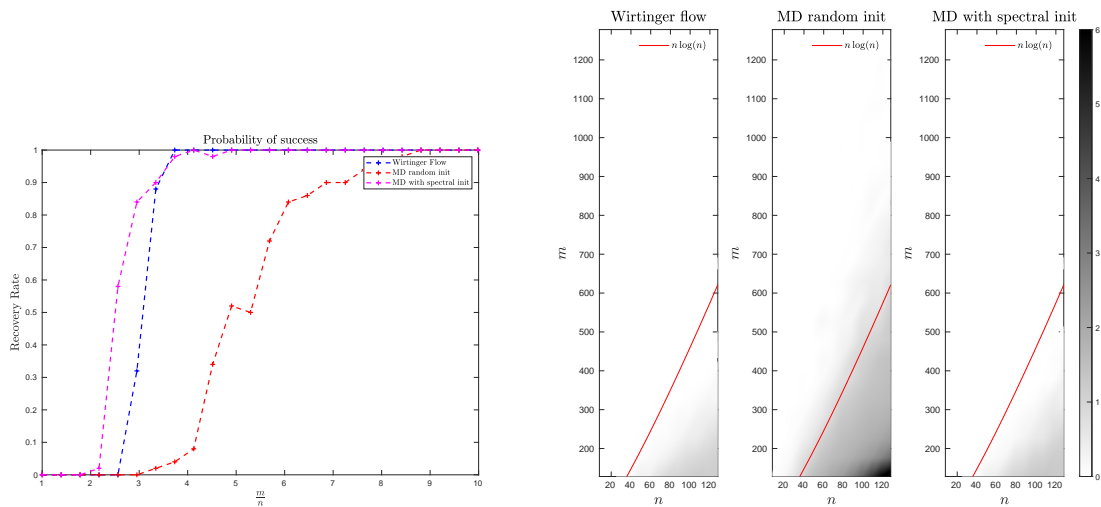


Figure 2: Phase diagrams for Gaussian measurements.

presence of noise. We used the spectral method in Algorithm 2 for the Wirtinger flow, and we compared with mirror descent with and without spectral initialization.

For this, we report the results of an experiment designed to estimate the phase retrieval probability of success of each algorithm as we vary  $n$  and  $m$ . The results are depicted in Figure 2. For each pair  $(n, m)$ , we generated 100 instances and solved them with each algorithm. Each diagram shows the empirical success probability (among the 100 instances) of the corresponding algorithm. An algorithm is declared as successful if the relative error (5.1) is less than  $\frac{2\|\epsilon\|}{\sqrt{m\sigma}} \approx 10^{-5}$ . The grayscale of each point in the diagrams reflects the observed probability of success, from 0% (black) to 100% (white). The solid curve marks the prediction of the phase transition edge. On the left panel of Figure 2, we also plot a profile of the phase diagram extracted at  $n = 128$ .

One observes a phase transition phenomenon that is in agreement with the predicted sample complexity bound shown as a solid line. Moreover, mirror descent performs better than the Wirtinger flow with both use spectral initialization. Mirror descent with uniform random initialization has a weaker recovery performance with a transition to success occurring at a higher threshold compared to the version of mirror descent with spectral initialization. This is in agreement with our theoretical findings as more measurements are needed in this case to ensure stable recovery.

## 5.2 Experiments with the CDP model

We now turn to the case of structured measurements from the CDP model. This model uses  $P$  coded diffraction patterns/masks followed by a Fourier transform. The observation model is given by

$$y = \left( \left| \sum_{\ell=0}^{n-1} \bar{x}_\ell d_p[\ell] e^{-i \frac{2\pi j \ell}{n}} \right|^2 + \epsilon_{j,p} \right)_{j,p}, \quad (5.2)$$

where  $j \in \{0, \dots, n-1\}$  and  $p \in \{0, \dots, P-1\}$ ,  $\epsilon$  is the noise. The total number of measurements is thus  $m = nP$  (i.e. the oversampling factor is  $P$ ).  $(d_p)_{p \in [P]}$  are i.i.d copies of a random variable  $d$ , and in our experiment,  $d$  takes values in  $\{-1, 0, 1\}$  with probability  $\{1/4, 1/2, 1/4\}$ . Here we performed a similar experience to the Gaussian case described in Section 5.1.1, where we chose the number of masks  $P = 7 \times \log^3(128)$  and a constant step-size  $\gamma = \frac{0.99}{2+\tilde{\epsilon}}$ , with  $\tilde{\epsilon} = 10^{-5}$ . Despite the lack of theoretical guarantees for the CDP model in the noisy case, that we conjecture are true, one can observe in Figure 3 that we have very similar results to those for Gaussian measurements.

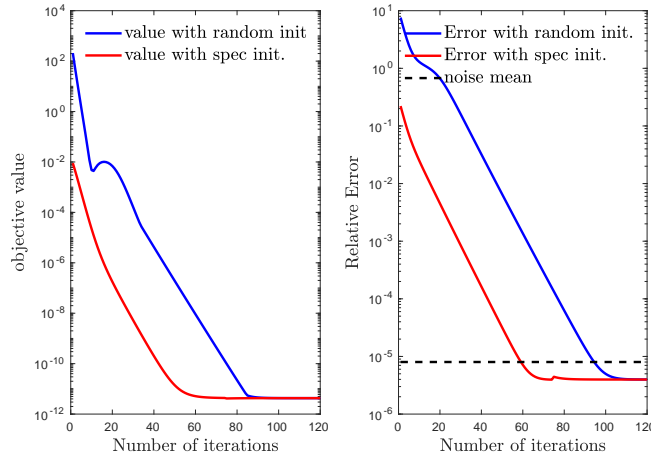


Figure 3: Reconstruction of signal from Noisy CDP. The noise mean is  $\tilde{\epsilon}$

## 5.3 Recovery of a 2D image

In this experiment, we work with the image of the beautiful Unicaen's<sup>1</sup> phoenix whose dimension is  $396 \times 396$ . Our goal is to recover the image from noisy CDP measurements with  $P = 90$  masks. The noise is chosen such that  $\tilde{\epsilon} = 10^{-5}$ . We used the spectral method to find the initial guess and run mirror descent for 1000 iterations. The results are displayed in Figure 4 showing that our algorithm converges to the desired image with a relative error of order  $10^{-2}$ .

# Appendices

## A Proofs for the Deterministic Case

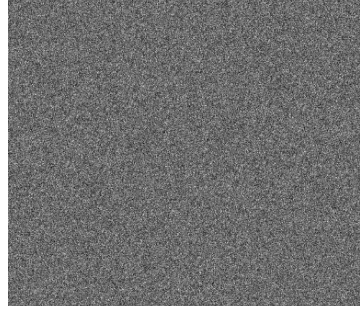
Throughout the work, we use when it is convenient the following decomposition of the objective function  $f$  in (1.1).

$$\forall x \in \mathbb{R}^n, \quad f(x) = f_{\text{NL}}(x) + f_{\text{Ny}}(x), \quad (\text{A.1})$$

<sup>1</sup>Unicaen = University of Caen



(a) Original Unicaen's phoenix



(b) The CDP measurements averaged over the  $P = 90$  masks



(c) Recovered Unicaen's phoenix

Figure 4: Reconstruction of an image from noisy CDP measurements.

where  $f_{\text{NL}}$  and  $f_{\text{Ny}}$  denote respectively the noiseless and the noisy part of  $f$  and we have

$$f_{\text{NL}}(x) = \frac{1}{4m} \sum_{r=1}^m \left( |a_r^\top x|^2 - |a_r^\top \bar{x}|^2 \right)^2, \quad f_{\text{Ny}}(x) = -\frac{1}{2m} \sum_{r=1}^m \epsilon[r] \left( |a_r^\top x|^2 - |a_r^\top \bar{x}|^2 \right) + \frac{\|\epsilon\|^2}{4m}. \quad (\text{A.2})$$

The following computations are straightforward:

$$\nabla \psi(x) = \left( \|x\|^2 + 1 \right) x, \quad \nabla^2 \psi(x) = \left( \|x\|^2 + 1 \right) \text{Id} + 2xx^\top, \quad (\text{A.3})$$

$$\nabla f(x) = \frac{1}{m} \sum_{r=1}^m \left( |a_r^\top x|^2 - |a_r^\top \bar{x}|^2 \right) a_r a_r^\top - \frac{1}{m} \sum_{r=1}^m \epsilon[r] a_r a_r^\top \quad (\text{A.4})$$

and

$$\nabla^2 f(x) = \frac{1}{m} \sum_{r=1}^m \left( 3|a_r^\top x|^2 - |a_r^\top \bar{x}|^2 \right) a_r a_r^\top - \frac{1}{m} \sum_{r=1}^m \epsilon[r] a_r a_r^\top. \quad (\text{A.5})$$

### A.1 Proof of Lemma 3.1

**Proof.** For all  $x, u \in \mathbb{R}^n$ , it easy to check that

$$\langle u, \nabla^2 \psi(x) u \rangle \geq \left( \|x\|^2 + 1 \right) \|u\|^2.$$



On the other hand, we have

$$\begin{aligned}
\langle u, \nabla^2 f(x)u \rangle &= \frac{1}{m} \sum_{r=1}^m \left( 3|a_r^\top x|^2 - |a_r^\top \bar{x}|^2 - \epsilon[r] \right) |a_r^\top u|^2 \\
&\leq \frac{1}{m} \sum_{r=1}^m \left( 3|a_r^\top x|^2 - \epsilon[r] \right) |a_r^\top u|^2 \\
&\leq \left( \frac{1}{m} \sum_{r=1}^m \left( 3\|a_r\|^2 + \|\epsilon\|_\infty \right) \|a_r\|^2 \right) (\|x\|^2 + 1) \|u\|^2.
\end{aligned}$$

Thus for any  $L \geq \frac{1}{m} \sum_{r=1}^m \|a_r\|^2 \left( 3\|a_r\|^2 + \|\epsilon\|_\infty \right)$ , we have for all  $x \in \mathbb{R}^n$

$$\nabla^2 f(x) \preceq L \nabla^2 \psi(x). \quad (\text{A.6})$$

The claim then follows from Lemma 2.4 with  $g = f$  and  $\phi = L\psi$ , and Proposition 2.1(ii).  $\square$

## A.2 Proof of Theorem 3.3

Let us start with the following intermediate results.

**Lemma A.1.** *Let the sequence  $(x_k)_{k \in \mathbb{N}}$  be generated by Algorithm 1. Then for all  $u \in \mathbb{R}^n$ ,*

$$D_\psi(u, x_{k+1}) + \gamma_k (f(x_{k+1}) - \min f) \leq D_\psi(u, x_k) - \kappa D_\psi(x_{k+1}, x_k) - \gamma_k D_f(u, x_k) + \gamma_k (f(u) - \min f). \quad (\text{A.7})$$

**Proof.** From [15, Lemma A.2], we have

$$\forall u \in \mathbb{R}^n, \quad D_\psi(u, x_{k+1}) + \gamma_k (f(x_{k+1}) - f(u)) \leq D_\psi(u, x_k) - \kappa D_\psi(x_{k+1}, x_k) - \gamma_k D_f(u, x_k).$$

Subtracting  $\min f$  from both sides yields the result.  $\square$

**Lemma A.2.** *Assume that  $\text{Argmin}(f) \neq \emptyset$ . We have*

$$0 \leq f(\pm \bar{x}) - \min f \leq \frac{\|\epsilon\|^2}{m}.$$

**Proof.** Let  $x^* \in \text{Argmin}(f)$ . We prove the claim for  $\bar{x}$ . By optimality, we have  $f(x^*) \leq f(\bar{x}) = \frac{\|\epsilon\|^2}{4m}$ , which equivalently reads

$$\sum_{r=1}^m \left( |a_r^\top x^*|^2 - |a_r^\top \bar{x}|^2 \right)^2 \leq 2 \sum_{r=1}^m \epsilon_r \left( |a_r^\top x^*|^2 - |a_r^\top \bar{x}|^2 \right).$$

Applying Young's inequality to the right-hand side then entails

$$(1 - \delta) \sum_{r=1}^m \left( |a_r^\top x^*|^2 - |a_r^\top \bar{x}|^2 \right)^2 \leq \frac{\|\epsilon\|^2}{\delta} \quad \forall \delta \in ]0, 1[. \quad (\text{A.8})$$

Consequently, using (A.8) and Young's inequality again, we get

$$\begin{aligned}
4mD_f(\pm\bar{x}, x^*) &= 4m(f(\bar{x}) - f(x^*)) \\
&= 2 \sum_{r=1}^m \epsilon[r] \left( |a_r^\top x^*|^2 - |a_r^\top \bar{x}|^2 \right) - \sum_{r=1}^m \left( |a_r^\top x^*|^2 - |a_r^\top \bar{x}|^2 \right)^2 \\
&\leq 2 \sum_{r=1}^m \epsilon[r] \left( |a_r^\top x^*|^2 - |a_r^\top \bar{x}|^2 \right) \\
&\leq \frac{\|\epsilon\|^2}{1-\delta} + (1-\delta) \sum_{r=1}^m \left( |a_r^\top x^*|^2 - |a_r^\top \bar{x}|^2 \right)^2 \leq \frac{\|\epsilon\|^2}{\delta(1-\delta)}.
\end{aligned}$$

The minimal value of the right hand side is  $4\|\epsilon\|^2$  attained for  $\delta = 1/2$ .  $\square$

**Proof.**

**(i)-(ii)** Similar to the proofs of the corresponding claims in [15, Theorem 2.11].

**(iii)** We give the proof for  $\bar{x}$  and obviously the same holds at  $-\bar{x}$ . We proceed by induction. We first have that  $x_0 \in B(\bar{x}, r) \subset B(\bar{x}, \rho)$  since  $r \leq \rho$ . Suppose now that for  $k \geq 0$ ,  $(x_i)_{0 \leq i \leq k} \subset B(\bar{x}, \rho)$ . Applying Lemma A.1 at  $\bar{x}$  and using Lemma A.2, we have

$$\begin{aligned}
D_\psi(\bar{x}, x_{k+1}) &\leq D_\psi(\bar{x}, x_k) - \kappa D_\psi(x_{k+1}, x_k) - \gamma D_f(\bar{x}, x_k) + \gamma \frac{\|\epsilon\|^2}{m} \\
&\leq D_\psi(\bar{x}, x_k) - \gamma D_f(\bar{x}, x_k) + \gamma \frac{\|\epsilon\|^2}{m} \\
&\leq (1 - \gamma\sigma) D_\psi(\bar{x}, x_k) + \gamma \frac{\|\epsilon\|^2}{m},
\end{aligned} \tag{A.9}$$

where we also used positivity of  $D_\psi$  and local  $\sigma$ -relative strong convexity of  $f$  since  $x_k \in B(\bar{x}, \rho)$ . Iterating the last inequality, we get

$$\begin{aligned}
D_\psi(\bar{x}, x_{k+1}) &\leq (1 - \gamma\sigma)^{k+1} D_\psi(\bar{x}, x_0) + \gamma \frac{\|\epsilon\|^2}{m} \sum_{i=0}^k (1 - \gamma\sigma)^i \\
&\leq (1 - \gamma\sigma)^{k+1} D_\psi(\bar{x}, x_0) + \frac{\|\epsilon\|^2}{m\sigma} \left( 1 - (1 - \gamma\sigma)^{k+1} \right) \\
&\leq D_\psi(\bar{x}, x_0) + \frac{\|\epsilon\|^2}{m\sigma}.
\end{aligned} \tag{A.10}$$

It then follows from Proposition 2.1(iv) that

$$\|x_{k+1} - \bar{x}\|^2 \leq \|x_0 - \bar{x}\|^2 \Theta(\rho) + \frac{2\|\epsilon\|^2}{m\sigma} \leq r^2 \Theta(\rho) + \frac{2\|\epsilon\|^2}{m\sigma} \leq \rho^2.$$

This shows (3.3).  $\square$

## B Proofs for Gaussian Measurements

### B.1 Expectation and deviation of the hessian

**Lemma B.1. (Expectation of the hessian)** *If the sensing vectors  $(a_r)_{r \in [m]}$  are sampled following the Gaussian model then we have for any  $x \in \mathbb{R}^n$ ,*

$$\mathbb{E}(\nabla^2 f(x)) = 3 \left( 2xx^\top + \|x\|^2 \text{Id} \right) - 2\bar{x}\bar{x}^\top - \|\bar{x}\|^2 \text{Id} - \tilde{\epsilon} \text{Id}. \tag{B.1}$$

**Proof.** The proof combines (A.5), [15, Lemma B.1] for the expectation of the first (*i.e.* noiseless part), and the last term comes the fact that the sensing vectors have zero mean and unit covariance.  $\square$

**Lemma B.2. (Concentration of the hessian)** Fix  $\varrho \in ]0, 1[$ , if the number of samples obeys  $m \geq C(\varrho)n \log n$ , for some sufficiently large constant  $C(\varrho) > 0$  then

$$\|\nabla^2 f(x) - \mathbb{E}(\nabla^2 f(x))\| \leq \varrho \left( \|x\|^2 + \frac{\|\bar{x}\|^2}{3} + \|\epsilon\|_\infty \right) \quad (\text{B.2})$$

holds simultaneously for all  $x \in \mathbb{R}^n$  with a probability at least  $1 - 5e^{-\zeta n} - \frac{4}{n^2} - 2e^{-\frac{m(\sqrt{1+\varrho}-1)^2}{8}}$ , where  $\zeta$  is a fixed numerical constant.

**Proof.** By the triangle inequality, we have

$$\begin{aligned} \|\nabla^2 f(x) - \mathbb{E}(\nabla^2 f(x))\| &\leq \left\| \frac{1}{m} \sum_{r=1}^m \left( 3|a_r^\top x|^2 a_r a_r^\top - |a_r^\top \bar{x}|^2 a_r a_r^\top \right) - \left( 6xx^\top + 3\|x\|^2 \text{Id} - 2\bar{x}\bar{x}^\top - \|\bar{x}\|^2 \text{Id} \right) \right\| \\ &\quad + \left\| \frac{1}{m} \sum_{r=1}^m \epsilon[r] a_r a_r^\top - \bar{\epsilon} \text{Id} \right\|. \end{aligned}$$

The concentration of the first term has been proved for the noiseless case (see [15, Lemma B.3]) with a probability  $1 - 5e^{-\zeta n} - \frac{4}{n^2}$ . For the noisy part, we have

$$\left\| \frac{1}{m} \sum_{r=1}^m \epsilon[r] a_r a_r^\top - \bar{\epsilon} \text{Id} \right\| \leq \frac{\|\epsilon\|_\infty}{m} \left\| \sum_{r=1}^m \left( a_r a_r^\top - \text{Id} \right) \right\| = \frac{\|\epsilon\|_\infty}{m} \|A^\top A - \text{Id}\|, \quad (\text{B.3})$$

where  $A$  is the  $m \times n$  matrix whose  $r$ -th row is the vector  $a_r^\top$ . From [15, Lemma B.4], we get that for any  $\varrho \in ]0, 1[$ ,

$$\left\| \frac{1}{m} \sum_{r=1}^m \epsilon[r] a_r a_r^\top - \bar{\epsilon} \text{Id} \right\| \leq \varrho \|\epsilon\|_\infty$$

with a probability at least  $1 - 2e^{-mt^2/2}$ , with  $\frac{\varrho}{4} = t^2 + t$  with  $m \geq \frac{16}{\varrho^2}n$ . We conclude by applying a simple union bound.  $\square$

## B.2 Optimal solution near the true vector

**Lemma B.3.** Assume that Assumption (A.1) holds and that  $m \geq cn$  where  $c$  is a positive numerical constant. Then for any  $x^* \in \text{Argmin}(f)$ ,

$$\text{dist}(x^*, \bar{\mathcal{X}}) \leq 8 \frac{\|\epsilon\|}{\sqrt{m} \|\bar{x}\|} \quad (\text{B.4})$$

holds with probability  $1 - e^{-\Omega(m)}$ .

**Proof.** Let us use the following notation:  $X^* = x^* x^{*\top}$ ,  $\bar{X} = \bar{x} \bar{x}^\top$ ,  $\bar{\epsilon} \stackrel{\text{def}}{=} \frac{\|\epsilon\|}{\sqrt{m}}$  and  $\epsilon_0 \stackrel{\text{def}}{=} 4\bar{\epsilon}$ .

By optimality of  $x^*$ , we have  $f(x^*) \leq f(\bar{x}) = \frac{\|\epsilon\|^2}{m}$ , which also implies that

$$\sum_{r=1}^m \left( |a_r^\top x^*|^2 - |a_r^\top \bar{x}|^2 \right)^2 \leq 2 \sum_{r=1}^m \epsilon_r \left( |a_r^\top x^*|^2 - |a_r^\top \bar{x}|^2 \right).$$

Applying Young's inequality to the right-hand side then entails

$$\sum_{r=1}^m (|a_r^\top x^*|^2 - |a_r^\top \bar{x}|^2)^2 \leq 4 \|\epsilon\|^2. \quad (\text{B.5})$$

Fix  $\zeta \in ]0, 1[$ . Using [9, Lemma 1], there are positive numerical constants  $C$  and  $C'$  such that if  $m \gtrsim n\zeta^{-2} \log(1/\zeta)$ , then with probability at least  $1 - C'e^{-C\zeta^2 m}$ , we have

$$\frac{1}{m} \sum_{r=1}^m (|a_r^\top x^*|^2 - |a_r^\top \bar{x}|^2)^2 \geq 0.81(1 - \zeta)^2 \|X^* - \bar{X}\|_F^2.$$

Thus, in view of (B.5), with the same probability, we have

$$\|X^* - \bar{X}\|_F \leq \frac{20}{9(1 - \zeta)} \bar{\epsilon}. \quad (\text{B.6})$$

Therefore taking  $\zeta = 0.4$  in (B.6) one has

$$\|X^* - \bar{X}\|_F \leq \epsilon_0. \quad (\text{B.7})$$

The rest of the proof is inspired by that of [8, Theorem 1.2]. Since  $\|x^*\|^2$  and  $\|\bar{x}\|^2$  are the largest eigenvalues of the rank-one symmetric matrices  $X^*$  and  $\bar{X}$ , we have from Weyl's perturbation inequality of the eigenvalues that

$$\left| \|x^*\|^2 - \|\bar{x}\|^2 \right| \leq \|X^* - \bar{X}\|_F \leq \epsilon_0.$$

Let us assume that  $\|\bar{x}\|^2 = 1$  and the general case is obtained via a simple rescaling argument. Under Assumption (A.1),  $\bar{\epsilon}$  is small enough so that  $\epsilon_0 < 1$ . We then get that  $\|x^*\|^2 \in [1 - \epsilon_0, 1 + \epsilon_0]$ . The sin- $\theta$ -Theorem [10] implies that

$$|\sin \theta| \leq \frac{\|X^* - \bar{X}\|_F}{\|x^*\|^2} \leq \frac{\epsilon_0}{1 - \epsilon_0},$$

where  $0 \leq \theta \leq \frac{\pi}{2}$  is the angle between  $x^*$  and  $\bar{x}$  which are the eigenvectors of  $X^*$  and  $\bar{X}$  associated to the eigenvalues  $\|x^*\|^2$  and 1, respectively. We can then write

$$x^* = \|x^*\| (\cos \theta \bar{x} + \sin \theta \bar{x}^\perp),$$

where  $\bar{x}^\perp$  is a unit vector orthogonal to  $\bar{x}$ . We apply the Ihâmessou-Pythagoras theorem to get

$$\|x^* - \bar{x}\|^2 = (1 - \|x^*\| \cos \theta)^2 + \|x^*\|^2 \sin^2 \theta.$$

Since  $\cos \theta = \sqrt{1 - \sin^2 \theta}$ , we have

$$1 + \epsilon_0 \geq \sqrt{1 + \epsilon_0} \geq \|x^*\| \cos \theta \geq \sqrt{1 - \epsilon_0 - \frac{\epsilon_0^2}{1 - \epsilon_0}} \geq 1 - \epsilon_0,$$

where we used that  $\epsilon_0 < 1/3$  in the last inequality. We then get that

$$(1 - \|x^*\| \cos \theta)^2 \leq \epsilon_0^2.$$

In turn

$$\|x^* - \bar{x}\|^2 \leq \epsilon_0^2 + \frac{\epsilon_0^2(1 + \epsilon_0)}{(1 - \epsilon_0)^2} \leq 4\epsilon_0^2$$

for  $\epsilon_0 < 1/3$ . We also know that

$$\|x^* - \bar{x}\| \leq 2 + \epsilon_0 \leq 7/3$$

for  $\epsilon_0 < 1/3$ . We therefore get that

$$\text{dist}(x^*, \bar{\mathcal{X}}) \leq 8 \min(\bar{\epsilon}, 1) \leq 8\bar{\epsilon},$$

where the last inequality is a consequence of Assumption (A.1). □

### B.3 Relative smoothness

**Lemma B.4.** Fix  $\varrho \in ]0, 1[$ , if the event  $\mathcal{E}_{\text{conH}}$  holds true then,

$$\forall x, z \in \mathbb{R}^n, \quad D_f(x, z) \leq \left(3 + \tilde{\epsilon} + \varrho \max\left(\|\bar{x}\|^2/3 + \|\epsilon\|_\infty, 1\right)\right) D_\psi(x, z). \quad (\text{B.8})$$

**Proof.** Let fix  $\varrho \in ]0, 1[$ , for any  $u \in \mathbb{R}^n$ , we have

$$\begin{aligned} \nabla^2 f(u) &\preceq \mathbb{E}(\nabla^2 f(u)) + \varrho(\|u\|^2 + \|\bar{x}\|^2/3 + \|\epsilon\|_\infty)\text{Id}, \\ &\preceq 3\left(2uu^\top + \|u\|^2 \text{Id}\right) - 2\bar{x}\bar{x}^\top - \|\bar{x}\|^2 \text{Id} - \tilde{\epsilon}\text{Id} + \varrho \max(\|\bar{x}\|^2/3 + \|\epsilon\|_\infty, 1)\left(\|u\|^2 + 1\right) \text{Id}, \\ &\preceq 3\left(2uu^\top + (\|u\|^2 + 1)\text{Id}\right) + \tilde{\epsilon}\text{Id} + \varrho \max(\|\bar{x}\|^2/3 + \|\epsilon\|_\infty, 1)\nabla^2 \psi(x), \\ &\preceq 3\nabla^2 \psi(u) + \tilde{\epsilon}\nabla^2 \psi(u) + \varrho \max(\|\bar{x}\|^2/3 + \|\epsilon\|_\infty, 1)\nabla^2 \psi(x), \\ &\preceq \left(3 + \tilde{\epsilon} + \varrho \max(\|\bar{x}\|^2/3 + \|\epsilon\|_\infty, 1)\right) \nabla^2 \psi(u), \end{aligned} \quad (\text{B.9})$$

We conclude by applying Lemma 2.4 in the segment  $[x, z]$ .  $\square$

### B.4 Local relative strong convexity

**Lemma B.5.** Fix  $\lambda \in \left] \frac{1}{9\sqrt{2}}, 1 \right[$  and for  $\varrho \in \left] 0, \frac{\lambda \min(\|\bar{x}\|^2, 1) - \tilde{\epsilon}}{2 \max(\|\bar{x}\|^2/3 + \|\epsilon\|_\infty, 1)} \right[$ . If the event  $\mathcal{E}_{\text{conH}}$  holds true, then for any  $x, z \in B(\bar{x}, \rho)$  or  $x, z \in B(-\bar{x}, \rho)$ , we have

$$D_f(x, z) \geq \left(\lambda \min(\|\bar{x}\|^2, 1) - \tilde{\epsilon} - \varrho \max(\|\bar{x}\|^2/3 + \|\epsilon\|_\infty, 1)\right) D_\psi(x, z), \quad (\text{B.10})$$

where  $\rho = \frac{1-\lambda}{\sqrt{3}} \|\bar{x}\|$ .

**Proof.** For any  $u \in \mathbb{R}^n$ , we have

$$\begin{aligned} \nabla^2 f(u) &\succeq \mathbb{E}(\nabla^2 f(u)) - \varrho\left(\|u\|^2 + \|\bar{x}\|^2/3 + \|\epsilon\|_\infty\right) \text{Id}, \\ &\succeq 6uu^\top + 3\|u\|^2 - 2\bar{x}\bar{x}^\top - \|\bar{x}\|^2 \text{Id} - \tilde{\epsilon}\text{Id} - \varrho \max(\|\bar{x}\|^2/3 + \|\epsilon\|_\infty, 1)(\|u\|^2 + 1). \end{aligned}$$

We then obtain, for any  $v \in \mathbb{S}^{n-1}$

$$v^\top \left( \nabla^2 f(u) + \varrho \max(\|\bar{x}\|^2/3 + \|\epsilon\|_\infty, 1) \nabla^2 \psi(u) \right) v \geq 3 \left( 2 \left( v^\top u \right)^2 + \|u\|^2 \right) - \left( 2 \left( v^\top \bar{x} \right)^2 + \|\bar{x}\|^2 \right) - \tilde{\epsilon} \text{Id}.$$

Let  $\rho > 0$  be small enough, to be made precise later. Thus for any  $u = \pm \bar{x} + \rho v$  we get

$$\begin{aligned} &v^\top \nabla^2 f(u) v + \varrho \max(\|\bar{x}\|^2/3 + \|\epsilon\|_\infty, 1) v^\top \nabla^2 \psi(u) v \\ &\geq 6 \left( v^\top \bar{x} \right)^2 + 6\rho^2 \pm 12\rho v^\top \bar{x} + 3\|\bar{x}\|^2 \pm 6\rho v^\top \bar{x} + 3\rho^2 - 2 \left( v^\top \bar{x} \right)^2 - \|\bar{x}\|^2 - \tilde{\epsilon} \\ &= 4 \left( v^\top \bar{x} \right)^2 + 9\rho^2 \pm 18\rho v^\top \bar{x} + 2\|\bar{x}\|^2 - \tilde{\epsilon}. \end{aligned} \quad (\text{B.11})$$

For the entropy  $\psi$ , we also have

$$v^\top \nabla^2 \psi(u) v = \|u\|^2 + 1 + 2 \left( v^\top u \right)^2 = 2 \left( v^\top \bar{x} \right)^2 + 3\rho^2 + \pm 6\rho v^\top \bar{x} + \|\bar{x}\|^2 + 1.$$

At this step, the proof becomes very similar to the noiseless phase retrieval (see [15, Lemma B.6]). Indeed, let us observe that we showed that for any vector  $u = \pm \bar{x} + \rho v$  with  $\rho = \frac{1-\lambda}{\sqrt{3}} \|\bar{x}\|$  we have,

$$4 \left( v^\top \bar{x} \right)^2 + 9\rho^2 \pm 18\rho v^\top \bar{x} + 2 \|\bar{x}\|^2 \geq \lambda \min(\|\bar{x}\|^2, 1) \left( 2 \left( v^\top \bar{x} \right)^2 + 3\rho^2 + \pm 6\rho v^\top \bar{x} + \|\bar{x}\|^2 + 1 \right)$$

By replacing this result in (B.11), we get

$$\begin{aligned} & v^\top \nabla^2 f(u) v + \varrho \max(\|\bar{x}\|^2/3 + \|\epsilon\|_\infty, 1) v^\top \nabla^2 \psi(u) v \\ & \geq \lambda \min(\|\bar{x}\|^2, 1) \left( 2 \left( v^\top \bar{x} \right)^2 + 3\rho^2 + \pm 6\rho v^\top \bar{x} + \|\bar{x}\|^2 + 1 \right) - \tilde{\epsilon} \\ & \geq \left( \lambda \min(\|\bar{x}\|^2, 1) - \tilde{\epsilon} \right) v^\top \nabla^2 \psi(u) v \end{aligned}$$

Finally, we have that

$$v^\top \left( \nabla^2 f(x) - \left( \lambda \min(\|\bar{x}\|^2, 1) - \tilde{\epsilon} - \varrho \max(\|\bar{x}\|^2/3 + \|\epsilon\|_\infty, 1) \right) \nabla^2 \psi(x) \right) v \geq 0$$

for all  $v \in \mathbb{S}^{n-1}$  and  $\rho \leq \frac{1-\lambda}{\sqrt{3}} \|\bar{x}\|$ . To conclude the proof, let us observe that with the prescribed bound on  $\varrho$ , we have

$$\sigma = \lambda \min(\|\bar{x}\|^2, 1) - \tilde{\epsilon} - \varrho \max(\|\bar{x}\|^2/3 + \|\epsilon\|_\infty, 1) > \frac{\lambda \min(\|\bar{x}\|^2, 1) - \tilde{\epsilon}}{2} > 0, \quad (\text{B.12})$$

where we used Assumption (A.1) on the noise. Therefore, (B.10) follows simply by invoking Lemma 2.4.  $\square$

We have the following corollary which gives a condition on the coefficient of the signal-to-noise ratio  $c_s$  ensuring that the neighborhood of strong convexity  $\rho$  is greater than the noise.

**Corollary B.6.** *For any fixed  $\lambda \in \left] \frac{1}{9\sqrt{2}}, 1 \right[$  and  $\varrho \in \left] 0, \frac{\lambda \min(\|\bar{x}\|^2, 1) - \tilde{\epsilon}}{2 \max(\|\bar{x}\|^2/3 + \|\epsilon\|_\infty, 1)} \right[$ , if Assumption (A.1) is satisfied then  $r^2 = \rho^2 - \frac{4\|\epsilon\|^2}{m\sigma} > 0$ , where  $\rho = \frac{1-\lambda}{\sqrt{3}} \|\bar{x}\|$ .*

**Proof.** To have the desired result, it suffices that  $\rho^2 - \frac{4\|\epsilon\|^2}{m\sigma} > 0$  i.e.  $\rho > 2 \frac{\|\epsilon\|_\infty}{\sqrt{\sigma}}$ . From (B.12) we have,

$$\sqrt{\sigma} > \frac{\sqrt{\lambda \min(\|\bar{x}\|^2, 1) - \tilde{\epsilon}}}{\sqrt{2}}$$

thus,

$$\frac{2\|\epsilon\|_\infty}{\sqrt{\sigma}} \leq \frac{2\sqrt{2}\|\epsilon\|_\infty}{\sqrt{\lambda \min(\|\bar{x}\|^2, 1) - \tilde{\epsilon}}} \leq \frac{2\sqrt{2}c_s \min(\|\bar{x}\|^2, 1)}{\sqrt{\lambda \min(\|\bar{x}\|^2, 1) - \tilde{\epsilon}}}.$$

Therefore it suffices to show that

$$\rho > \frac{2\sqrt{2}c_s \min(\|\bar{x}\|^2, 1)}{\sqrt{\lambda \min(\|\bar{x}\|^2, 1) - \tilde{\epsilon}}}.$$

Replacing now  $\rho$  by its expression, we get that  $c_s$  should satisfy

$$(1 - \lambda) \|\bar{x}\| \sqrt{\lambda \min(\|\bar{x}\|^2, 1) - \tilde{\epsilon}} > 2\sqrt{6}c_s \min(\|\bar{x}\|^2, 1)$$

which holds thanks to Assumption (A.1).  $\square$

**Remark B.7.** This result estimates the maximum signal-to-noise ratio for which we ensure that the neighborhood of strong convexity around the true vectors is well-defined. Let us notice that a more practical upper bound is

$$c_s < \frac{(1-\lambda)\sqrt{\lambda}}{2\sqrt{6}} \leq \frac{1}{9\sqrt{2}}. \quad (\text{B.13})$$

Indeed, it is a simple maximization of the function  $\lambda \mapsto \frac{(1-\lambda)\sqrt{\lambda}}{2\sqrt{6}}$  over  $\left] \frac{1}{9\sqrt{2}}, 1 \right[$ .

## B.5 Spectral initialization

We now show that the initial guess  $x_0$  generated by spectral initialization (Algorithm 2) belongs to a small  $f$ -attentive neighborhood of  $\bar{\mathcal{X}}$ .

**Lemma B.8.** Fix  $\varrho \in ]0, 1[$  and assume that for  $r \in [m]$ , we have  $\epsilon[r] \leq |a_r^\top \bar{x}|^2$ . If the number of samples obeys  $m \geq C(\varrho)n \log n$  for some sufficiently large constant  $C(\varrho) > 0$  and (B.2) holds true then  $x_0$  satisfies:

(i)  $\text{dist}(x_0, \bar{\mathcal{X}}) \leq \eta_1(\varrho) \|\bar{x}\|$ , where

$$\begin{aligned} \eta_1 : ]0, 1[ &\rightarrow ]0, 1[ \\ \varrho &\mapsto \left( \sqrt{2 - 2\sqrt{1 - \varrho(1 + c_s)}} + \frac{\varrho(1 + c_s)}{2} \right), \end{aligned} \quad (\text{B.14})$$

which is an increasing function.

(ii) Moreover, we have

$$f(x_0) \leq f(\bar{x}) + \left( (1 + \varrho)c_s \|\bar{x}\| + L \frac{\Theta(\eta_1(\varrho) \|\bar{x}\|)}{2} \eta_1(\varrho) \right) \eta_1(\varrho) \|\bar{x}\|^2, \quad (\text{B.15})$$

with  $L = 3 + \tilde{\epsilon} + \varrho \max(\|\bar{x}\|^2/3 + \|\epsilon\|_\infty, 1)$ .

(iii) Besides, for  $\lambda \in ]0, 1[$ , if

$$\varrho \leq \eta_1^{-1} \left( \frac{1 - \lambda}{\sqrt{3 \left( 6 \left( 1 + \frac{(1-\lambda)^2}{3} - \frac{4\|\epsilon\|^2}{m\sigma\|\bar{x}\|^2} \right) + 1 \right)} \max(\|\bar{x}\|, 1)} \right), \quad (\text{B.16})$$

then we have  $x_0 \in B \left( \bar{\mathcal{X}}, \frac{r}{\max(\sqrt{\Theta(r)}, 1)} \right)$  where  $r^2 = \frac{(1-\lambda)^2}{3} \|\bar{x}\|^2 - \frac{4\|\epsilon\|^2}{m\sigma}$ .

**Proof.** (i) Denote the matrix

$$Y = \frac{1}{m} \sum_{r=1}^m y[r] a_r a_r^\top = \frac{1}{m} \sum_{r=1}^m (|a_r^\top \bar{x}|^2 + \epsilon[r]) a_r a_r^\top.$$

We have w.h.p

$$\|Y - \mathbb{E}(Y)\| \leq \varrho(\|\bar{x}\|^2 + \|\epsilon\|_\infty) \leq \varrho(1 + c_s) \|\bar{x}\|^2.$$

Let  $\tilde{x}$  be the eigenvector associated with the largest eigenvalue  $\tilde{\lambda}$  of  $Y$  such that  $\|\tilde{x}\| = \|\bar{x}\|$  (obviously  $\tilde{\lambda}$  is nonnegative since  $Y$  is semidefinite positive). Then,

$$\begin{aligned} \varrho(1 + c_s) \|\bar{x}\|^4 &\geq \left| \tilde{x}^\top \left( Y - 2\bar{x}\bar{x}^\top - \|\bar{x}\|^2 \text{Id} - \tilde{\epsilon} \text{Id} \right) \tilde{x} \right| \\ &= \left| \tilde{\lambda} \|\bar{x}\|^2 - 2(\tilde{x}^\top \bar{x})^2 - \|\bar{x}\|^4 - \tilde{\epsilon} \|\bar{x}\|^2 \right| \end{aligned}$$



Hence

$$2(\tilde{x}^\top \bar{x})^2 \geq \tilde{\lambda} \|\bar{x}\|^2 - \|\bar{x}\|^4 - \tilde{\epsilon} \|\bar{x}\|^2 - \varrho(1 + c_s) \|\bar{x}\|^4.$$

Moreover, since  $\tilde{\lambda}$  is the largest eigenvalue of  $Y$ , applying the concentration inequality at  $\bar{x}$  *w.h.p*

$$\begin{aligned} \tilde{\lambda} \|\bar{x}\|^2 &\geq \bar{x}^\top Y \bar{x} \geq \bar{x}^\top \left( 2\bar{x}\bar{x}^\top + \|\bar{x}\|^2 \text{Id} - \tilde{\epsilon} \text{Id} \right) \bar{x} - \varrho(1 + c_s) \|\bar{x}\|^4 \\ &= 3 \|\bar{x}\|^4 + \tilde{\epsilon} \|\bar{x}\|^2 - \varrho(1 + c_s) \|\bar{x}\|^4. \end{aligned}$$

Merging the last two inequalities, we get

$$\begin{aligned} 2(\tilde{x}^\top \bar{x})^2 &\geq 3 \|\bar{x}\|^4 + \tilde{\epsilon} \|\bar{x}\|^2 - \varrho(1 + c_s) \|\bar{x}\|^4 - \|\bar{x}\|^4 - \tilde{\epsilon} \|\bar{x}\|^2 - \varrho(1 + c_s) \|\bar{x}\|^4 \\ &= 2 \|\bar{x}\|^4 - 2\varrho(1 + c_s) \|\bar{x}\|^4. \end{aligned}$$

Which implies that

$$\text{dist}(\tilde{x}, \bar{\mathcal{X}}) \leq \sqrt{2 - 2\sqrt{1 - \varrho(1 + c_s)}} \|\bar{x}\|.$$

By definition of  $x_0$  in Algorithm 2,  $x_0 = \sqrt{\frac{1}{m} \sum_r y[r]} \frac{\tilde{x}}{\|\tilde{x}\|} = \sqrt{\frac{1}{m} \sum_r (|a_r^\top \tilde{x}|^2 + \epsilon[r])} \frac{\tilde{x}}{\|\tilde{x}\|}$ , and thus *w.h.p*

$$\|x_0 - \tilde{x}\| = \left| \sqrt{\frac{m^{-1} \sum_r y[r]}{\|\tilde{x}\|^2}} - 1 \right| \|\tilde{x}\| = \left| \sqrt{\frac{m^{-1} \sum_r (|a_r^\top \tilde{x}|^2 + \epsilon[r])}{\|\tilde{x}\|^2}} - 1 \right| \|\tilde{x}\| \leq \frac{\varrho(1 + c_s) \|\tilde{x}\|}{2},$$

it comes out that,

$$\text{dist}(x_0, \bar{\mathcal{X}}) \leq \text{dist}(\tilde{x}, \bar{\mathcal{X}}) + \|x_0 - \tilde{x}\| \leq \left( \sqrt{2 - 2\sqrt{1 - \varrho(1 + c_s)}} + \frac{\varrho(1 + c_s)}{2} \right) \|\tilde{x}\|.$$

- (ii) Under our sampling complexity bound, event  $\mathcal{E}_{\text{conH}}$  defined by (4.3) holds true *w.h.p*. It then follows from Lemma B.4 applied at  $\bar{x}$  and  $x_0$ , that

$$D_f(x_0, \bar{x}) \leq LD_\psi(x_0, \bar{x}).$$

The latter implies that

$$\begin{aligned} f(x_0) &\leq f(\bar{x}) + \langle \nabla f(\bar{x}), x_0 - \bar{x} \rangle + LD_\psi(x_0, \bar{x}), \\ f(x_0) &\leq f(\bar{x}) + \|\nabla f(\bar{x})\| \|x_0 - \bar{x}\| + L \frac{\Theta(\eta_1(\varrho) \|\bar{x}\|)}{2} \|x_0 - \bar{x}\|^2, \end{aligned}$$

Since  $\nabla f(\bar{x}) = \frac{1}{m} \sum_{r=1}^m \epsilon[r] a_r a_r^\top \bar{x}$ , we obtain from (B.3) that

$$\|\nabla f(\bar{x})\| \leq \|\epsilon\|_\infty \left\| \frac{1}{m} \sum_{r=1}^m a_r a_r^\top \right\| \|\bar{x}\| \leq (1 + \varrho) \|\epsilon\|_\infty \|\bar{x}\|.$$

Combining the two last inequality yields to

$$f(x_0) \leq f(\bar{x}) + \left( (1 + \varrho) c_s \|\bar{x}\| + L \frac{\Theta(\eta_1(\varrho) \|\bar{x}\|)}{2} \eta_1(\varrho) \right) \eta_1(\varrho) \|\bar{x}\|^2.$$

- (iii) In view of (i), it is sufficient to show that  $\eta_1(\varrho) \|\bar{x}\| \leq \frac{r}{\max(\sqrt{\Theta(r)}, 1)}$ . Since from Proposition 2.1(iv), we have

$$\Theta(r) \leq 6(\|\bar{x}\|^2 + r^2) + 1 \leq \left( 6 \left( 1 + \frac{(1 - \lambda)^2}{3} - \frac{4 \|\epsilon\|^2}{m\sigma \|\bar{x}\|^2} \right) + 1 \right) \max(\|\bar{x}\|^2, 1),$$

and  $\eta_1$  is an increasing function, we conclude.  $\square$

## C Landscape of the Noise-Aware Objective with Gaussian Measurements

### C.1 Warm up: Critical points of $\mathbb{E}(f)$

We start by studying and characterizing the set of critical points of  $\mathbb{E}(f)$ . This can be seen as the asymptotic behavior of the critical points of  $f$  when the number of measurements  $m$  grows to  $+\infty$ .

**Proposition C.1.** *We have*

$$\text{crit}(\mathbb{E}(f)) = \{0\} \cup \overline{\mathcal{X}}_\epsilon \cup \left\{ x \in \mathbb{R}^n : \bar{x}^\top x = 0, \|x\|^2 = \frac{1}{3} \left( \|\bar{x}\|^2 + \tilde{\epsilon} \right) \right\},$$

where  $\overline{\mathcal{X}}_\epsilon \stackrel{\text{def}}{=} \left\{ \pm \bar{x} \sqrt{1 + \frac{\tilde{\epsilon}}{3\|\bar{x}\|^2}} \right\}$ . Those sets are respectively, the local maximizer, the set of global minimizers, and strict saddle points of  $\mathbb{E}(f)$ .

Before proving this result, we the closed form expressions of the expectation of  $f$  and its derivatives.

**Lemma C.2.** *For all  $x \in \mathbb{R}^n$ , we have:*

$$\begin{aligned} \mathbb{E}(f(x)) &= \frac{3}{4} \left( \|x\|^4 + \|\bar{x}\|^4 \right) - \frac{1}{2} \|\bar{x}\|^2 \|x\|^2 - (|\bar{x}^\top x|^2 + \frac{\|\epsilon\|^2}{4m} - \frac{\tilde{\epsilon}(\|x\|^2 - \|\bar{x}\|^2)}{2}), \\ \nabla \mathbb{E}(f(x)) &= 3 \|x\|^2 x - 2\bar{x}(\bar{x}^\top x) - \|\bar{x}\|^2 x - \tilde{\epsilon}x, \\ \nabla^2 \mathbb{E}(f(x)) &= 3 \left( 2xx^\top + \|x\|^2 \text{Id} \right) - 2\bar{x}\bar{x}^\top - \|\bar{x}\|^2 \text{Id} - \tilde{\epsilon} \text{Id}. \end{aligned}$$

**Proof.** By linearity of the expectation, we have  $\mathbb{E}(f(x)) = \mathbb{E}(f_{\text{NL}}(x)) + \mathbb{E}(f_{\text{Ny}}(x))$ . Linearity again yields

$$\begin{aligned} \mathbb{E}(f_{\text{Ny}}(x)) &= -\frac{1}{2m} \sum_{r=1}^m \epsilon[r] \left( x^\top \mathbb{E} \left( a_r a_r^\top \right) x - \bar{x}^\top \mathbb{E} \left( a_r a_r^\top \right) \bar{x} \right) + \frac{\|\epsilon\|^2}{4m}, \\ &= -\frac{1}{2m} \sum_{r=1}^m \epsilon[r] \left( \|x\|^2 - \|\bar{x}\|^2 \right) + \frac{\|\epsilon\|^2}{4m}. \end{aligned} \tag{C.1}$$

We also have

$$\mathbb{E}(f_{\text{NL}}(x)) = \frac{1}{4m} \sum_{r=1}^m \mathbb{E} \left( |a_r^\top x|^4 \right) + \frac{1}{4m} \sum_{r=1}^m \mathbb{E} \left( |a_r^\top \bar{x}|^4 \right) - \frac{1}{2m} \sum_{r=1}^m |a_r^\top x|^2 |a_r^\top \bar{x}|^2.$$

From [15, Lemma B.1], we know that

$$\forall x \in \mathbb{R}^n, \quad \mathbb{E} \left( \sum_{r=1}^m |a_r^\top x|^2 a_r a_r^\top \right) = 2xx^\top + \|x\|^2 \text{Id}.$$

Therefore we have

$$\mathbb{E} \left( \frac{1}{4m} \sum_{r=1}^m |a_r^\top x|^4 \right) = x^\top \left( \mathbb{E} \left( \frac{1}{4m} \sum_{r=1}^m |a_r^\top x|^2 a_r a_r^\top \right) \right) x = \frac{1}{4} x^\top (2xx^\top + \|x\|^2 \text{Id}) x = \frac{3}{4} \|x\|^4,$$

and

$$\mathbb{E} \left( \frac{1}{2m} \sum_{r=1}^m |a_r^\top x|^2 |a_r^\top \bar{x}|^2 \right) = x^\top \left( \bar{x}\bar{x}^\top + \frac{1}{2} \|\bar{x}\|^2 \text{Id} \right) x = |\bar{x}^\top x|^2 + \frac{1}{2} \|\bar{x}\|^2 \|x\|^2.$$

Whence we have

$$\mathbb{E}(f_{\text{NL}}(x)) = \frac{3}{4} \left( \|x\|^4 + \|\bar{x}\|^4 \right) - \frac{1}{2} \|\bar{x}\|^2 \|x\|^2 - |\bar{x}^\top x|^2. \tag{C.2}$$

The claim follows simply by summing (C.1) and (C.2).

We deduce the gradient and the hessian by straightforward derivation of  $\mathbb{E}(f)$ .  $\square$

**Proof.** (Proposition C.1)

- **The origin :** Let us observe that  $\nabla \mathbb{E}(f(0)) = 0$ , it follows that the value of the hessian at zero satisfies

$$\nabla^2 \mathbb{E}(f(0)) = -\|\bar{x}\|^2 \text{Id} - 2\bar{x}\bar{x}^\top - \tilde{\epsilon} \text{Id} \preceq -(\|\bar{x}\|^2 + \tilde{\epsilon}) \text{Id} - 2\bar{x}\bar{x}^\top \prec 0,$$

where we have used that  $\tilde{\epsilon}$  is non-negative (see Assumption (A.1)). It follows that 0 is a local maximizer of  $\mathbb{E}(f)$ .

- **Subspaces with no critical points:**

1. For any point in  $\left\{x \in \mathbb{R}^n : 0 < \|x\|^2 < \frac{\|\bar{x}\|^2 + \tilde{\epsilon}}{3}\right\}$ , we have

$$\langle x, \nabla \mathbb{E}(f(x)) \rangle = \left(3\|x\|^2 - \|\bar{x}\|^2 - \tilde{\epsilon}\right) \|x\|^2 - 2|\bar{x}^\top x|^2 \leq \left(3\|x\|^2 - \|\bar{x}\|^2 - \tilde{\epsilon}\right) \|x\|^2 < 0.$$

We deduce that,  $\langle x, \nabla \mathbb{E}(f(x)) \rangle < 0$  which implies that  $\|\nabla \mathbb{E}(f(x))\|$  is bounded away from zero on this region, and thus that there are no critical points there.

2. Consider a point in  $\left\{x \in \mathbb{R}^n : \frac{\|\bar{x}\|^2 + \tilde{\epsilon}}{3} < \|x\|^2 < \|\bar{x}\|^2 + \frac{\tilde{\epsilon}}{3}\right\}$  and recall that it is a critical point if and only if

$$\left(3\|x\|^2 - \|\bar{x}\|^2 - \tilde{\epsilon}\right) x = 2(\bar{x}\bar{x}^\top)x. \quad (\text{C.3})$$

Combining Assumption (A.1) and the fact that  $\|x\|^2 > \frac{\|\bar{x}\|^2 + \tilde{\epsilon}}{3}$ , we get that  $\left(3\|x\|^2 - \|\bar{x}\|^2 - \tilde{\epsilon}\right) / 2$  is the positive eigenvalue of the rank-one matrix  $\bar{x}\bar{x}^\top$ . This is equivalent to  $3\|x\|^2 - \|\bar{x}\|^2 - \tilde{\epsilon} = 2\|\bar{x}\|^2$ , *i.e.*,  $\|x\|^2 = \|\bar{x}\|^2 + \frac{\tilde{\epsilon}}{3}$  which contradicts the definition of this region, showing again that there are no critical points there.

3. For a point in the region  $\left\{x \in \mathbb{R}^n : \|x\|^2 > \|\bar{x}\|^2 + \frac{\tilde{\epsilon}}{3}\right\}$ , we have the lower bound

$$\begin{aligned} \langle x, \nabla \mathbb{E}(f(x)) \rangle &= \left(3\|x\|^2 - \|\bar{x}\|^2 - \tilde{\epsilon}\right) \|x\|^2 - 2|\bar{x}^\top x|^2, \\ &\geq \left(3\|x\|^2 - 3\|\bar{x}\|^2 - \tilde{\epsilon}\right) \|x\|^2 > 0. \end{aligned}$$

Hence,  $\|\nabla \mathbb{E}(f(x))\|$  is bounded away from zero on this region yielding the same conclusion.

- **Strict saddle points:** A point in the sphere  $\left\{x \in \mathbb{R}^n : \|x\|^2 = \frac{\|\bar{x}\|^2 + \tilde{\epsilon}}{3}\right\}$  is a critical point if it is orthogonal to the true vector  $\bar{x}$ . Indeed we have,

$$\nabla \mathbb{E}(f(x)) = 0 \iff \left(3\|x\|^2 - \|\bar{x}\|^2 - \tilde{\epsilon}\right) x = 2\bar{x}\bar{x}^\top x \iff \bar{x}^\top x = 0.$$

Besides, for any  $v \in \mathbb{R}^n$  we have

$$\begin{aligned} \langle v, \nabla^2 \mathbb{E}(f(x)) v \rangle &= 6|v^\top x|^2 + 3\|x\|^2 \|v\|^2 - 2|v^\top \bar{x}|^2 - \|\bar{x}\|^2 \|v\|^2 - \tilde{\epsilon} \|v\|^2 \\ &= 6|v^\top x|^2 - 2|v^\top \bar{x}|^2 + \|v\|^2 \left(3\|x\|^2 - \|\bar{x}\|^2 - \tilde{\epsilon}\right) \\ &= 6|v^\top x|^2 - 2|v^\top \bar{x}|^2. \end{aligned}$$

In the direction  $v = x$ , we deduce that

$$\langle x, \nabla^2 \mathbb{E}(f(x)) x \rangle = 6\|x\|^4 > 0,$$

and in the direction  $v = \bar{x}$  we have

$$\langle \bar{x}, \nabla^2 \mathbb{E}(f(x)) \bar{x} \rangle = -2\|\bar{x}\|^4 < 0,$$

where we have used orthogonality of  $x$  and  $\bar{x}$ . These facts show that the critical points in this region, *i.e.* points orthogonal to  $\bar{x}$ , are strict saddle points of  $\mathbb{E}(f)$ .

- **Global minimizers:** In view of the above, local/global minimizers can only occur on the sphere  $\left\{x \in \mathbb{R}^n : \|x\|^2 = \|\bar{x}\|^2 + \frac{\tilde{\epsilon}}{3}\right\}$ . Any point in his set is a critical point of  $\mathbb{E}(f)$  if and only if

$$\nabla \mathbb{E}(f(x)) = 0 \iff \left(3\|x\|^2 - \|\bar{x}\|^2 - \tilde{\epsilon}\right)x = 2\bar{x}\bar{x}^\top x \iff \left(\|\bar{x}\|^2 \text{Id} - \bar{x}\bar{x}^\top\right)x = 0.$$

Therefore,  $x$  is a critical point on this region if and only if  $x$  is an eigenvector of  $\bar{x}\bar{x}^\top$ , that is  $x \in \text{span}(\bar{x})$ , or equivalently,  $\exists \beta \in \mathbb{R}$  such that  $x = \beta\bar{x}$  with

$$\|x\|^2 = \beta^2 \|\bar{x}\|^2 = \|\bar{x}\|^2 + \frac{\tilde{\epsilon}}{3} \iff \beta = \pm \sqrt{1 + \frac{\tilde{\epsilon}}{3\|\bar{x}\|^2}}.$$

The set of critical points in this region is reduced to  $\bar{\mathcal{X}}_\epsilon \stackrel{\text{def}}{=} \left\{\pm \bar{x} \sqrt{1 + \frac{\tilde{\epsilon}}{3\|\bar{x}\|^2}}\right\}$ . For  $x \in \bar{\mathcal{X}}_\epsilon$  we have

$$\begin{aligned} \nabla^2 \mathbb{E}(f(x)) &= 6xx^\top - 2\bar{x}\bar{x}^\top + \left(3\|x\|^2 - \|\bar{x}\|^2 - \tilde{\epsilon}\right) \text{Id} \\ &= (6\beta^2 - 2)\bar{x}\bar{x}^\top + 2\|\bar{x}\|^2 \text{Id} \succeq 2\|\bar{x}\|^2 \text{Id}. \end{aligned}$$

Indeed, we have

$$6\beta^2 - 2 = 4 + \frac{2\tilde{\epsilon}}{\|\bar{x}\|^2} \geq 4 > 0,$$

where we use again the non-negativity of  $\tilde{\epsilon}$  in Assumption (A.1). We conclude that  $\bar{\mathcal{X}}_\epsilon$  is the set of global minimizers of  $\mathbb{E}(f)$ . □

## C.2 Main result: Critical points of $f$

In this section, we study the landscape of the objective function  $f$  for the Gaussian measurement model. Our main result hereafter characterizes the set of critical points of  $f$  for  $m$  large enough.

**Theorem C.3. (Critical points of  $f$ )** Fix  $\lambda \in \left] \frac{1}{9\sqrt{2}}, 1 \right[$ . Let us assume that the noise vector satisfies Assumption (A.1). If  $m \gtrsim n \log(n)^3$ , then

$$\text{crit}(f) = \text{Argmin}(f) \cup \text{strisad}(f) \tag{C.4}$$

where  $\text{Argmin}(f) = \{\pm x^*\}$ . This holds with probability of at least  $1 - \frac{c}{m}$  where  $c$  is a positive numerical constant.

### Remark C.4.

- In [26, Theorem 2.2], the authors study the geometry of  $f$  in the noiseless case here coined as  $f_{\text{NL}}$ . We aim with our result to extend it to the noisy case with small enough noise (see Assumption (A.1)).
- This result shows that when the number of measurements  $m$  is sufficiently large and the noise  $\epsilon$  is very small compared to the true vector which is entailed by a large SNR, then the set of critical points of the objective function  $f$  is reduced to the set of global minimizers  $\text{Argmin}(f)$  and the set of strict saddle points  $\text{strisad}(f)$ . The strict saddle avoidance of mirror descent will then imply that the sequence provided by mirror descent will always converge to global minimizers of the function  $f$ .

We recall the radius  $\rho = \frac{1-\lambda}{\sqrt{3}} \|\bar{x}\|$  defined in Lemma B.5. To prove Theorem C.3, we consider the following regions of  $\mathbb{R}^n$  which are helpful to characterize the landscape of  $f$ :

$$\mathcal{R}_1 = \left\{ x \in \mathbb{R}^n : \langle \bar{x}, \mathbb{E}(\nabla^2 f(x)) \bar{x} \rangle \leq -\frac{1}{100} \|x\|^2 \|\bar{x}\|^2 - \frac{1}{50} \|\bar{x}\|^4 \right\}, \quad (\text{C.5})$$

$$\mathcal{R}_3 = \{x \in \mathbb{R}^n : \text{dist}(x, \bar{\mathcal{X}}) \leq \rho\}, \quad (\text{C.6})$$

$$\mathcal{R}_2 = (\mathcal{R}_1 \cup \mathcal{R}_3)^c. \quad (\text{C.7})$$

We also define specific regions  $\mathcal{R}_2^x$  and  $\mathcal{R}_2^h$ ,

$$\mathcal{R}_2^x = \left\{ x \in \mathbb{R}^n : \langle x, \mathbb{E}(\nabla f(x)) \rangle \geq \frac{1}{500} \|x\|^2 \|\bar{x}\|^2 + \frac{1}{100} \|x\|^4 \right\},$$

$$\mathcal{R}_2^h = \left\{ x \in \mathbb{R}^n : \langle d_x, \mathbb{E}(\nabla f(x)) \rangle \geq \frac{1}{250} \|x\| \|d_x\| \|\bar{x}\|^2, \frac{11}{20} \|\bar{x}\| \leq \|x\| \leq \|\bar{x}\|, \text{dist}(x, \bar{\mathcal{X}}) \geq \frac{\|\bar{x}\|}{3} \right\}.$$

where  $d_x = \frac{\pm\bar{x}-x}{\|\pm\bar{x}-x\|}$  if  $x \neq \pm\bar{x}$  and any vector on the unit sphere otherwise.

Let us observe that these regions are similar to those defined in [26] replacing  $f$  by  $f_{\text{NL}}$ . Indeed, the idea behind our assumptions is the fact that small noise will introduce small perturbations in the function  $f$ , and therefore under our assumption of small noise, the latter has benign influence on the landscape of  $f$  (see Figure 5). Mainly, in the region  $\mathcal{R}_1$  the function  $f$  still has negative curvature. In the region  $\mathcal{R}_2$ ,  $f$  has a large gradient and in  $\mathcal{R}_3$  relative strong convexity with respect to our chosen entropy  $\psi$ . It is important to observe that in the noisy case, it is not true that the true vectors  $\bar{\mathcal{X}}$  are critical points of  $f$  or even  $\mathbb{E}(f)$ . However, we have already shown in Lemma A.2 that  $\pm\bar{x}$  are actually  $\frac{\|\epsilon\|^2}{m}$ -minimizers. Moreover, we have already given in Proposition C.1 a description of the set of critical points of  $\mathbb{E}(f)$ , providing a hint that in the large oversampling regime, the geometry of  $f$  is close to that of  $f_{\text{NL}}$ . This result shows that the set of critical points of  $\mathbb{E}(f)$  is also reduced to the set of strict saddle points with symmetric global minimizers of  $\mathbb{E}(f)$ . This set of minimizers, that we denoted  $\bar{\mathcal{X}}_\epsilon$  (see Proposition C.1), are direct perturbations of the true vectors  $\bar{\mathcal{X}}$  by the noise; see also Lemma B.3 which quantifies the distance of global minimizers of  $f$  to  $\bar{\mathcal{X}}$  in probability.

**Proof.** In the following, all assertions are to be understood in high probability sense. The proof consists in invoking properly the statements of Proposition C.5. In the region  $\mathcal{R}_1$ , Proposition C.5-(i) shows that

$$\forall x \in \mathcal{R}_1, \quad \langle \bar{x}, \nabla^2 f(x) \bar{x} \rangle \leq -\frac{1}{100} (1 - c_s) \|\bar{x}\|^4, \quad (\text{C.8})$$

*i.e.*  $f$  has a negative curvature in the direction of the true vectors  $\bar{\mathcal{X}}$  which means that any critical point in  $\mathcal{R}_1$  is a strict saddle point for  $f$ . From Proposition C.5-(iii) and (iv), we deduce that

$$\forall x \in \mathcal{R}_2^x \cup \mathcal{R}_2^h, \quad \|\nabla f(x)\| \geq \frac{1}{500} (1 - c_s) \|x\| \|\bar{x}\|^2. \quad (\text{C.9})$$

Moreover, Proposition C.5-(v) entails that  $\mathcal{R}_2 \subset \mathcal{R}_2^x \cup \mathcal{R}_2^h$  which means that (C.9) holds true for all  $x \in \mathcal{R}_2$ . Thus the gradient of function  $f$  is bounded away from zero on  $\mathcal{R}_2$  which means that there are no critical points in this region. Therefore, local/global minimizers of  $f$  are necessarily located in the region  $\mathcal{R}_3$ . It remains to show that the only critical points in the domain  $\mathcal{R}_3$  are just the elements of  $\text{Argmin}(f)$ <sup>2</sup> which contains only two points  $\pm x^*$ . This will be a consequence of  $\sigma$ -strong convexity of  $f$  on  $\mathcal{R}_3$ . In the following, since  $\mathcal{R}_3 = B(\bar{x}, \rho) \cup B(-\bar{x}, \rho)$ , we prove the claim only on  $B(\bar{x}, \rho)$  and the

<sup>2</sup>Remember that  $\text{Argmin}(f)$  is a nonempty compact set by injectivity of  $A$  under the assumed measurement bound.

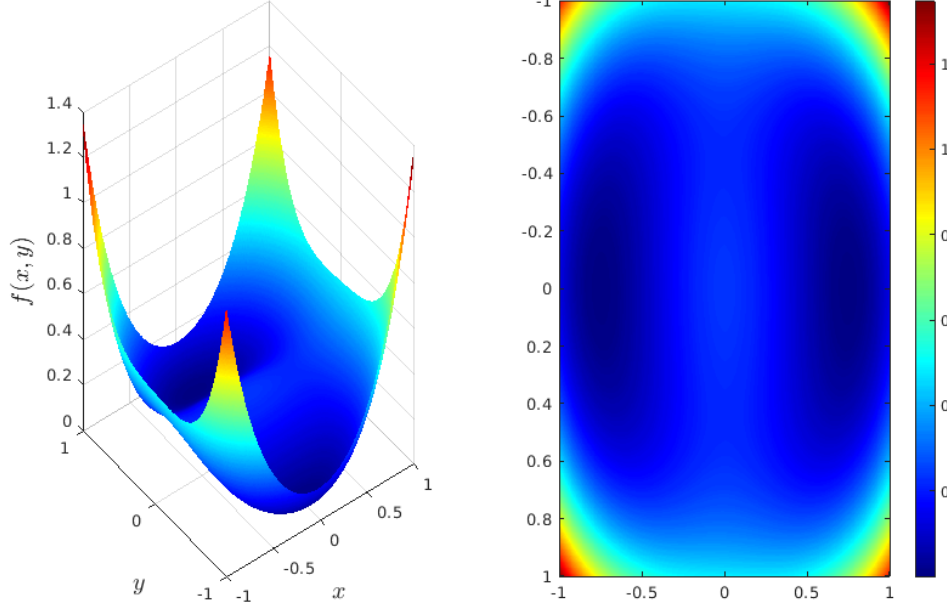


Figure 5: Landscape of the function  $f$  as  $m \rightarrow \infty$ ; we have  $(m, n) = (200, 2)$  and the true vectors are  $[\pm 3/4, 0]$ . The noise vector is generated at uniform in  $[-1, 1]$  such that  $\tilde{\epsilon} \approx 5.10^{-3}$ . One clearly sees that the geometry of the landscape of  $f$  is preserved and that the only minimizers of  $f$  are very close to the true vectors.

same holds for the symmetric case with  $-\bar{x}$ . Let  $x \in B(\bar{x}, \rho) \setminus \{x^*\}$ . In view of Proposition C.5-(ii), we have

$$D_f(x, x^*) = f(x) - \min f \geq \sigma D_\psi(x, x^*) \geq \frac{\sigma}{2} \|x - x^*\|^2.$$

The right hand side is positive since  $x \neq x^*$ , which means that  $f$  has a unique minimizer on  $B(\bar{x}, \rho)$ . Moreover,

$$D_f(x^*, x) = \min f - f(x) - \langle \nabla f(x), x^* - x \rangle \geq \sigma D_\psi(x^*, x) \geq \frac{\sigma}{2} \|x - x^*\|^2,$$

and thus

$$\langle \nabla f(x), x - x^* \rangle \geq D_f(x^*, x) \geq \frac{\sigma}{2} \|x - x^*\|^2.$$

Cauchy-Schwarz then entails

$$\|\nabla f(x)\| \geq \frac{\sigma}{2} \|x - x^*\| > 0$$

meaning that  $f$  has no other critical point than  $x^*$  on  $B(\bar{x}, \rho)$ . This completes the proof.  $\square$

The proof of the above result heavily relies on the behaviour of  $f$  on each region. This is the subject of the next proposition.

**Proposition C.5.** *If the number of samples obeys  $m \gtrsim n \log^3(n)$  then with probability  $1 - \frac{c}{m}$  where  $c$  is a positive numerical constant, we have the following statements.*

(i) *In the region  $\mathcal{R}_1$ , the objective  $f$  has a negative curvature i.e.,*

$$\forall x \in \mathcal{R}_1, \quad \langle \bar{x}, \nabla^2 f(x) \bar{x} \rangle \leq -\frac{1}{100} (1 - c_s) \|\bar{x}\|^4. \quad (\text{C.10})$$

(ii) *In  $\mathcal{R}_3$ ,  $f$  is  $\sigma$ -strongly convex where  $\sigma > 0$  is given in Proposition B.5.*

(iii) The gradient is bounded from away from zero in  $\mathcal{R}_2^x$ . More precisely,

$$\forall x \in \mathcal{R}_2^x, \quad \langle x, \nabla f(x) \rangle \geq \frac{1}{1000}(1 - c_s) \|x\|^2 \|\bar{x}\|^2. \quad (\text{C.11})$$

(iv) We have

$$\forall x \in \mathcal{R}_2^h, \quad \langle d_x, \nabla f(x) \rangle \geq \frac{1}{500}(1 - c_s) \|\bar{x}\|^2 \|x\| \|d_x\|. \quad (\text{C.12})$$

(v) We have  $\mathcal{R}_2 \subset \mathcal{R}_2^x \cup \mathcal{R}_2^h$ .

**Remark C.6.** The previous result extends the series of propositions ([26, Proposition 2.3-2.7]) to the noisy case. All the statements depend on the (inverse) signal-to-noise coefficient  $c_s$  which obviously less than 1 under our assumption. In the noiseless case, let us observe that we recover all the Propositions mentioned above.

**Proof.**

(i) For any  $x \in \mathcal{R}_1$ , we have

$$\langle \bar{x}, \nabla^2 f(x) \bar{x} \rangle = \frac{1}{m} \sum_{r=1}^m 3 \left| a_r^\top x \right|^2 \left| a_r^\top \bar{x} \right|^2 - \frac{1}{m} \sum_{r=1}^m \left| a_r^\top \bar{x} \right|^4 - \frac{1}{m} \sum_{r=1}^m \epsilon[r] \left| a_r^\top \bar{x} \right|^2.$$

By using similar concentration inequalities as in Lemma B.2 we have the following

$$\begin{aligned} \frac{1}{m} \sum_{r=1}^m 3 \left| a_r^\top x \right|^2 \left| a_r^\top \bar{x} \right|^2 &\leq \mathbb{E} \left( \frac{1}{m} \sum_{r=1}^m 3 \left| a_r^\top x \right|^2 \left| a_r^\top \bar{x} \right|^2 \right) + \varrho \|\bar{x}\|^2 \|x\|^2, \\ \frac{1}{m} \sum_{r=1}^m \left| a_r^\top \bar{x} \right|^4 &\geq \mathbb{E} \left( \frac{1}{m} \sum_{r=1}^m \left| a_r^\top \bar{x} \right|^4 \right) - \varrho \|\bar{x}\|^4, \\ \frac{1}{m} \sum_{r=1}^m \epsilon[r] \left| a_r^\top \bar{x} \right|^2 &\geq \mathbb{E} \left( \frac{1}{m} \sum_{r=1}^m \epsilon[r] \left| a_r^\top \bar{x} \right|^2 \right) - \varrho \|\epsilon\|_\infty \|x\|^2. \end{aligned}$$

After summing, we get

$$\langle \bar{x}, \nabla^2 f(x) \bar{x} \rangle \leq \langle \bar{x}, \mathbb{E}(\nabla^2 f(x)) \bar{x} \rangle + \varrho \|x\|^2 \|\bar{x}\|^2 + \varrho \|\bar{x}\|^4 + \varrho c_s \|\bar{x}\|^2 \|x\|^2.$$

We choose now  $\varrho = \frac{1}{100}$ , and since  $x \in \mathcal{R}_1$ , we finally obtain that

$$\begin{aligned} \langle \bar{x}, \nabla^2 f(x) \bar{x} \rangle &\leq -\frac{1}{100} \|x\|^2 \|\bar{x}\|^2 - \frac{1}{50} \|\bar{x}\|^4 + \frac{1}{100} \|x\|^2 \|\bar{x}\|^2 + \frac{1}{100} \|\bar{x}\|^4 + \frac{1}{100} c_s \|\bar{x}\|^2 \|x\|^2, \\ &= -\frac{1}{100} (1 - c_s) \|\bar{x}\|^4. \end{aligned}$$

(ii) Combine Lemma B.5 and 1-strong convexity of  $\psi$ .

(iii) Let  $x \in \mathcal{R}_2^x$ ,

$$\langle x, \nabla f(x) \rangle = \frac{1}{m} \sum_{r=1}^m \left| a_r^\top x \right|^4 - \frac{1}{m} \sum_{r=1}^m \left| a_r^\top \bar{x} \right|^2 \left| a_r^\top x \right|^2 - \frac{1}{m} \sum_{r=1}^m \epsilon[r] \left| a_r^\top x \right|^2.$$

using the same concentration arguments as in the proof of Lemma B.2, we get

$$\begin{aligned} \langle x, \nabla f(x) \rangle &\geq \langle x, \mathbb{E}(\nabla f(x)) \rangle - \frac{1}{100} \|x\|^4 - \frac{1}{1000} \|x\|^2 \|\bar{x}\|^2 - \frac{\|x\|^2 \|\epsilon\|_\infty}{1000}, \\ &\geq \frac{1}{500} \|x\|^2 \|\bar{x}\|^2 + \frac{1}{100} \|x\|^4 - \frac{1}{100} \|x\|^4 - \frac{1}{1000} \|x\|^2 \|\bar{x}\|^2 - \frac{\|x\|^2 \|\epsilon\|_\infty}{1000}, \\ &= \frac{1}{1000} (1 - c_s) \|\bar{x}\|^2 \|x\|^2, \end{aligned}$$

where we used Assumption (A.1) in the last inequality.



(iv) We have,

$$\langle d_x, \nabla f(x) \rangle = \langle d_x, \nabla f_{\text{NL}}(x) \rangle + \langle d_x, \nabla f_{\text{Ny}}(x) \rangle.$$

On the one hand, we have

$$\langle d_x, \nabla f_{\text{NL}}(x) \rangle = \frac{1}{m} \sum_{r=1}^m \left| a_r^\top \bar{x} \right|^2 (d_x^\top a_r a_r^\top x) - \frac{1}{m} \sum_{r=1}^m \left| a_r^\top \bar{x} \right|^2 (d_x^\top a_r a_r^\top x).$$

Therefore, the proof bears exactly the same path as the proof of [26, Proposition 2.6]. Indeed using Lemma B.2 more precisely [15, Lemma B.2], we have for  $m \gtrsim n \log(n)$

$$\frac{1}{m} \sum_{r=1}^m \left| a_r^\top \bar{x} \right|^2 (d_x^\top a_r a_r^\top x) \leq d_x^\top \mathbb{E} \left( \frac{1}{m} \sum_{r=1}^m \left| a_r^\top \bar{x} \right|^2 a_r a_r^\top \right) x + \frac{1}{1000} \|\bar{x}\|^2 \|d_x\| \|x\|,$$

with probability larger than  $1 - \frac{C}{m}$ ,  $C > 0$ .

To bound the first term, we use the line of reasoning used in the proof of [26, Proposition 2.6] to get that when  $m \gtrsim n \log(n)^3$ , with a probability at least  $1 - \frac{C'}{m}$ ,  $C' > 0$ , we have

$$\frac{1}{m} \sum_{r=1}^m \left| a_r^\top \bar{x} \right|^2 (d_x^\top a_r a_r^\top x) \geq d_x^\top \mathbb{E} \left( \frac{1}{m} \sum_{r=1}^m \left| a_r^\top \bar{x} \right|^2 a_r a_r^\top \right) x - \frac{1}{1000} \|\bar{x}\|^2 \|d_x\| \|x\|,$$

we refer to [26, Section 6.5] for the details of the proofs. Then, provided that  $m \gtrsim n \log(n)^3$ , with a union bound, we have with probability at least  $1 - \frac{C+C'}{m}$  that

$$\langle d_x, \nabla f_{\text{NL}}(x) \rangle \geq \langle d_x, \mathbb{E}(\nabla f_{\text{NL}}(x)) \rangle - \frac{1}{500} \|\bar{x}\|^2 \|x\| \|d_x\|.$$

On the other hand, with similar arguments as in the proof of Lemma B.2 and using again Assumption (A.1), we have with probability larger than  $1 - \frac{C+C'}{m}$

$$\langle d_x, \nabla f_{\text{Ny}}(x) \rangle = d_x^\top \left( \frac{1}{m} \sum_{r=1}^m \epsilon[r] a_r a_r^\top \right) x \geq \langle d_x, \mathbb{E}(\nabla f_{\text{Ny}}(x)) \rangle - \frac{c_s}{500} \|\bar{x}\|^2 \|x\| \|d_x\|.$$

We combine now the last two inequalities to get

$$\langle d_x, \nabla f(x) \rangle \geq \langle d_x, \mathbb{E}(\nabla f(x)) \rangle - \frac{1}{500} \|\bar{x}\|^2 \|x\| \|d_x\| - \frac{c_s}{500} \|\bar{x}\|^2 \|x\| \|d_x\|$$

with probability at least  $1 - 2\frac{C+C'}{m}$ . Using the definition of the region  $\mathcal{R}_2^h$  we obtain the claimed result.

(v) The proof is similar to that of [26, Proposition 2.7] which consists of showing that  $\mathbb{R}^n = \mathcal{R}_1 \cup \mathcal{R}_2^x \cup \mathcal{R}_2^h \cup \mathcal{R}_3$ . We then get our claim by definition of  $\mathcal{R}_2$  and that  $\mathbb{R}^n = \mathcal{R}_1 \cup \mathcal{R}_2 \cup \mathcal{R}_3$ . The idea is to divide the  $\mathbb{R}^n$  into several overlapping regions and show that we can cover them with our good partition. To achieve this task we will use the set

$$\mathcal{R}_2^{h'} = \left\{ x \in \mathbb{R}^n : \langle d_x, \mathbb{E}(\nabla f(x)) \rangle \geq \frac{1}{250} \|\bar{x}\|^2 \|x\| \|d_x\|, \|x\| \leq \|\bar{x}\| \right\}.$$

- We can cover the set  $\mathcal{R}_a \stackrel{\text{def}}{=} \{x \in \mathbb{R}^n : |x^\top \bar{x}| \leq \frac{1}{2} \|x\| \|\bar{x}\|\}$  with both  $\mathcal{R}_1$  and  $\mathcal{R}_2^x$ . If  $\|x\|^2 \leq \frac{298}{451} \|\bar{x}\|^2$

$$\begin{aligned}
\langle \bar{x}, \nabla^2 f(x) \bar{x} \rangle + \frac{1}{100} \|\bar{x}\|^2 \|x\|^2 &= 6(x^\top \bar{x})^2 + \frac{301}{300} \|x\|^2 \|\bar{x}\|^2 - 3 \|\bar{x}\|^4 - \tilde{\epsilon} \|\bar{x}\|^2 \\
&\leq \left( \frac{3}{2} + \frac{301}{100} \right) \|\bar{x}\|^2 \|x\|^2 - \frac{149}{50} \|\bar{x}\|^4 - \tilde{\epsilon} \|\bar{x}\|^2 - \frac{1}{50} \|\bar{x}\|^4 \\
&\leq \frac{298}{100} \|\bar{x}\|^4 - \frac{149}{50} \|\bar{x}\|^4 - \tilde{\epsilon} \|\bar{x}\|^2 - \frac{1}{50} \|\bar{x}\|^4 \\
&\leq \frac{1}{50} \|\bar{x}\|^4.
\end{aligned}$$

If  $\|x\|^2 \geq \frac{626}{995} \|\bar{x}\|^2$ ,

$$\begin{aligned}
\langle x, \mathbb{E}(\nabla f(x)) \rangle - \frac{1}{500} \|x\|^2 \|\bar{x}\|^2 &= 3 \|x\|^4 - 2(x^\top \bar{x})^2 - \frac{501}{500} \|\bar{x}\|^2 \|x\|^2 - \tilde{\epsilon} \|x\|^2 \\
&\geq 3 \|x\|^4 - \frac{1}{2} \|x\|^2 \|\bar{x}\|^2 - \frac{501}{500} \|\bar{x}\|^2 \|x\|^2 - \tilde{\epsilon} \|x\|^2 \\
&\geq \frac{1}{100} \|x\|^4 + \frac{299}{100} \|x\|^4 - \frac{751}{500} \|\bar{x}\|^2 \|x\|^2 - \tilde{\epsilon} \|x\|^2 \\
&\geq \frac{1}{100} \|x\|^4 + \frac{299}{100} \|x\|^4 - \left( \frac{751}{500} + \frac{1}{9\sqrt{2}} \right) \left( \frac{995}{626} \right) \|x\|^4 \\
&\geq \frac{1}{100} \|\bar{x}\|^4,
\end{aligned}$$

where we have used the fact that  $\tilde{\epsilon} \geq 0$  combined with the practical upper bound on  $c_s$  (B.13). Since  $\frac{298}{451} \geq \frac{626}{995}$ , we conclude that  $\mathcal{R}_a \subset \mathcal{R}_1 \cup \mathcal{R}_2^x$ .

- The set  $\mathcal{R}_b \stackrel{\text{def}}{=} \{x \in \mathbb{R}^n : |x^\top \bar{x}| \geq \frac{1}{2} \|x\| \|\bar{x}\|; \|x\| \leq \frac{57}{100} \|\bar{x}\|\}$  is covered by the set  $\mathcal{R}_1$ . Indeed for any  $x \in \mathcal{R}_b$  we have,

$$\begin{aligned}
\langle \bar{x}, \nabla^2 f(x) \bar{x} \rangle + \frac{1}{100} \|\bar{x}\|^2 \|x\|^2 &= 6(x^\top \bar{x})^2 + \frac{301}{300} \|x\|^2 \|\bar{x}\|^2 - 3 \|\bar{x}\|^4 - \tilde{\epsilon} \|\bar{x}\|^2 \\
&\leq \frac{901}{100} \|x\|^2 \|\bar{x}\|^2 - \frac{149}{50} \|\bar{x}\|^4 - \tilde{\epsilon} \|\bar{x}\|^2 - \frac{1}{50} \|\bar{x}\|^4 \\
&\leq \frac{901}{100} \left( \frac{57}{100} \right)^2 \|\bar{x}\|^2 - \frac{149}{50} \|\bar{x}\|^4 - \tilde{\epsilon} \|\bar{x}\|^2 - \frac{1}{50} \|\bar{x}\|^4 \\
&\leq -\tilde{\epsilon} \|\bar{x}\|^2 - \frac{1}{50} \|\bar{x}\|^4 \leq -\frac{1}{50} \|\bar{x}\|^4,
\end{aligned}$$

since  $\tilde{\epsilon} \geq 0$ .

- Let consider the set  $\mathcal{R}_c \stackrel{\text{def}}{=} \{x \in \mathbb{R}^n : \frac{1}{2} \|x\| \|\bar{x}\| \leq |x^\top \bar{x}| \leq \frac{99}{100} \|x\| \|\bar{x}\|\}$ , which is covered

by  $\mathcal{R}_2^x$  and  $\mathcal{R}_2^{h'}$ . For any  $x \in \mathcal{R}_c$  such that  $\|x\| \geq \sqrt{\frac{1996}{1973}} \|\bar{x}\|$  we have

$$\begin{aligned}
\langle x, \mathbb{E}(\nabla f(x)) \rangle &= \frac{1}{500} \|\bar{x}\|^2 \|x\|^2 + \frac{1}{100} \|x\|^4 \\
&= 3 \|x\|^4 - 2(x^\top \bar{x})^2 - \|\bar{x}\|^2 \|x\|^2 - \tilde{\epsilon} \|x\|^2 - \frac{1}{500} \|\bar{x}\|^2 \|x\|^2 \\
&\geq \frac{299}{100} \|x\|^4 + \frac{501}{500} \|\bar{x}\|^2 \|x\|^2 - 2(x^\top \bar{x})^2 - 2(x^\top \bar{x})^2 - \tilde{\epsilon} \|x\|^2 \\
&\geq \frac{299}{100} \|x\|^4 - \left( 2 \left( \frac{99}{100} \right)^2 + \frac{501}{500} + \frac{1}{9\sqrt{2}} \right) \|\bar{x}\|^2 \|x\|^2 \\
&\geq \left( \frac{299}{100} \left( \frac{1996}{1973} \right)^2 - \left( 2 \left( \frac{99}{100} \right)^2 + \frac{501}{500} + \frac{1}{9\sqrt{2}} \right) \right) \|x\|^4 \\
&\geq 0.
\end{aligned}$$

Therefore, we have  $\mathcal{R}_c \cap \left\{ x \in \mathbb{R}^n : \|x\| \geq \sqrt{\frac{1996}{1973}} \|\bar{x}\| \right\} \subset \mathcal{R}_2^x$ . To show the remaining inclusion, we use an  $(\alpha, \beta)$ -type argument. Let assume that  $\|x\| = \alpha \|\bar{x}\|$ ,  $|x^\top \bar{x}| = \beta = \|\bar{x}\| \|x\| = \alpha \beta \|\bar{x}\|$  and  $\tilde{\epsilon} = \varepsilon \|\bar{x}\|^2$  with  $\alpha \in \left[ \frac{11}{20}, \sqrt{\frac{1996}{1973}} \right]$ ,  $\beta \in \left[ \frac{1}{2}, \frac{99}{100} \right]$  and  $\varepsilon \in \left[ 0, \frac{1}{9\sqrt{2}} \right]$ . We have

$$\begin{aligned}
\langle x - \bar{x}, \mathbb{E}(\nabla f(x)) \rangle &= 3 \|x\|^4 + 3(x^\top \bar{x}) \left( \|\bar{x}\|^2 - \|x\|^2 \right) - 2(x^\top \bar{x})^2 - \|\bar{x}\|^2 \|x\|^2 + \tilde{\epsilon} \left( (x^\top \bar{x}) - \|x\| \right) \\
&= \|\bar{x}\|^4 \alpha \left( 3\alpha^3 + 3\beta(1 - \alpha^2) - 2\alpha\beta^2 - \alpha + \varepsilon(\beta - \alpha) \right).
\end{aligned}$$

Whence we have

$$\begin{aligned}
\frac{1}{\|\bar{x}\|^4 \alpha} \left( \langle x - \bar{x}, \mathbb{E}(\nabla f(x)) \rangle - \frac{1}{250} \|\bar{x}\|^2 \|x\| \|d_x\| \right) &= 3\alpha^3 + 3\beta(1 - \alpha^2) - 2\alpha\beta^2 \\
&\quad - \alpha + \varepsilon(\beta - \alpha) - \frac{1}{250} \sqrt{1 + \alpha^2 - 2\alpha\beta}.
\end{aligned}$$

It is straightforward that in this domain  $\frac{1}{250} \sqrt{1 + \alpha^2 - 2\alpha\beta} \leq \frac{1}{250} \sqrt{\frac{3969 - 2\sqrt{984527}}{1973}} \leq \frac{41}{10000}$ . Therefore we define the following function

$$p(\alpha, \beta, \varepsilon) \stackrel{\text{def}}{=} 3\alpha^3 + 3\beta(1 - \alpha^2) - 2\alpha\beta^2 - \alpha + \varepsilon(\beta - \alpha) - \frac{41}{10000}$$

Then  $p$  has a unique minimizer arising at  $\left( 0.998237, \frac{99}{100}, \frac{1}{9\sqrt{2}} \right)$  with a value  $\frac{87239}{2500000}$ . We deduce that

$$\langle x - \bar{x}, \mathbb{E}(\nabla f(x)) \rangle - \frac{1}{250} \|\bar{x}\|^2 \|x\| \|d_x\| \geq 0$$

Therefore  $\mathcal{R}_c \subset \mathcal{R}_2^x \cup \mathcal{R}_2^{h'}$ .

- We now cover  $\mathcal{R}_d \stackrel{\text{def}}{=} \left\{ x \in \mathbb{R}^n : \frac{99}{100} \|\bar{x}\| \|x\| \leq |x^\top \bar{x}| \leq \|x\| \|\bar{x}\|, \|x\| \geq \frac{11}{20} \|\bar{x}\| \right\}$  with  $\mathcal{R}_2^x, \mathcal{R}_3$

and  $\mathcal{R}_2^{h'}$ . For any  $x \in \mathcal{R}_d$ , with  $\|x\| \geq \sqrt{\frac{1031}{1000}} \|\bar{x}\|$ , we have

$$\begin{aligned} \langle x, \mathbb{E}(\nabla f(x)) \rangle - \frac{1}{500} \|x\|^2 \|\bar{x}\|^2 - \frac{1}{100} \|\bar{x}\|^4 &= \frac{299}{100} \|x\|^4 - 2(x^\top \bar{x})^2 - \frac{501}{500} \|\bar{x}\|^2 \|x\|^2 - \tilde{\epsilon} \|x\|^2 \\ &\geq \frac{299}{100} \|x\|^4 - \frac{1501}{500} \|\bar{x}\|^2 \|x\|^2 - \tilde{\epsilon} \|x\|^2 \\ &\geq \frac{299}{100} \|x\|^4 - \left( \frac{1501}{500} + \frac{1}{9\sqrt{2}} \right) \|\bar{x}\|^2 \|x\|^2 \\ &\geq \left( \frac{299}{100} - \left( \frac{1501}{500} + \frac{1}{9\sqrt{2}} \right) \frac{1000}{1031} \right) \|x\|^4 \geq 0. \end{aligned}$$

hence we have  $\mathcal{R}_d \cap \left\{ x \in \mathbb{R}^n : \|x\| \geq \sqrt{\frac{1031}{1000}} \|\bar{x}\| \right\} \subset \mathcal{R}_2^x$ . When  $\frac{23}{25} \leq \|x\| \leq \sqrt{\frac{1031}{1000}}$  we have,

$$\|d_x\|^2 = \|\bar{x}\|^2 + \|x\|^2 - 2x^\top \bar{x} = \|\bar{x}\|^2 (1 + \alpha^2 - 2\alpha\beta),$$

where we have  $\alpha \in \left[ \frac{23}{25}, \sqrt{\frac{1031}{1000}} \right]$  and  $\beta \in \left[ \frac{99}{100}, 1 \right]$ . Therefore, we consider

$$p(\alpha, \beta, \lambda) = 1 + \alpha^2 - 2\alpha\beta - \frac{(1-\lambda)^2}{3},$$

with  $\lambda \in \left] \frac{1}{9\sqrt{2}}, \frac{3}{5} \right]$ . The maximum value of  $p$  is taken at  $(\alpha, \beta, \lambda) = \left( \frac{23}{25}, \frac{99}{100}, \frac{3}{5} \right)$  thus  $p(\alpha, \beta, \lambda) \leq -\frac{107}{3750}$ . We deduce that  $\mathcal{R}_d \cap \left\{ x \in \mathbb{R}^n, \frac{23}{25} \leq \|x\| \leq \sqrt{\frac{1031}{1000}} \right\} \subset \mathcal{R}_3$ . When  $\frac{11}{20} \|\bar{x}\| \|x\| \leq \frac{24}{25} \|\bar{x}\|$ , we have

$$\begin{aligned} \frac{1}{\|\bar{x}\|^4 \alpha} \left( \langle x - \bar{x}, \mathbb{E}(\nabla f(x)) \rangle - \frac{1}{250} \|\bar{x}\|^2 \|x\| \|d_x\| \right) &= 3\alpha^3 + 3\beta(1 - \alpha^2) - 2\alpha\beta^2 \\ &\quad - \alpha + \varepsilon(\beta - \alpha) - \frac{1}{250} \sqrt{1 + \alpha^2 - 2\alpha\beta} \end{aligned}$$

where  $\alpha \in \left[ \frac{11}{20}, \sqrt{\frac{24}{25}} \right]$ ,  $\beta \in \left[ \frac{99}{100}, 1 \right]$  and  $\varepsilon \in \left[ 0, \frac{1}{9\sqrt{2}} \right]$ . One check easily that

$$p(\alpha, \beta, \varepsilon) = 3\alpha^3 + 3\beta(1 - \alpha^2) - 2\alpha\beta^2 - \alpha + \varepsilon(\beta - \alpha) - \frac{1}{250} \sqrt{1 + \alpha^2 - 2\alpha\beta} \geq 0.$$

Consequently, we have  $\mathcal{R}_d \cap \left\{ x \in \mathbb{R}^n : \frac{11}{20} \|\bar{x}\| \|x\| \leq \frac{24}{25} \|\bar{x}\| \right\} \subset \mathcal{R}_2^{h'}$ . Finally  $\mathcal{R}_d \subset \mathcal{R}_2^x \cup \mathcal{R}_3 \cup \mathcal{R}_2^{h'}$ .

By construction, we have  $\mathcal{R}_a \cup \mathcal{R}_b \cup \mathcal{R}_c \cup \mathcal{R}_d = \mathbb{R}^n$ , and therefore

$$\begin{aligned} \mathbb{R}^n &= \mathcal{R}_a \cup \mathcal{R}_b \cup \mathcal{R}_c \cup \mathcal{R}_d \\ &\subset \mathcal{R}_1 \cup \mathcal{R}_2^x \cup \mathcal{R}_2^{h'} \cup \mathcal{R}_3 \\ &= \mathcal{R}_1 \cup \mathcal{R}_2^x \cup \left( \mathcal{R}_2^{h'} \cap \left\{ x \in \mathbb{R}^n : \frac{11}{20} \|\bar{x}\| \leq \|x\| \right\} \right) \cup \mathcal{R}_3 \\ &= \mathcal{R}_1 \cup \mathcal{R}_2^x \cup \left( \mathcal{R}_2^{h'} \cap \left\{ x \in \mathbb{R}^n : \frac{11}{20} \|\bar{x}\| \leq \|x\| \right\} \cap \mathcal{R}_3^c \right) \cup \mathcal{R}_3 \\ &= \mathcal{R}_1 \cup \mathcal{R}_2^x \cup \mathcal{R}_2^h \cup \mathcal{R}_3. \end{aligned}$$

□

**Acknowledgments.** The authors thank the French National Research Agency (ANR) for funding the project FIRST (ANR-19-CE42-0009).

## References

- [1] C. Amra, M. Zerrad, S. Liukaityte, and M. Lequime. Instantaneous one-angle white-light scatterometer. *Opt. Express, OE*, 26(1):204–219, January 2018.
- [2] H. H. Bauschke, J. Bolte, J. Chen, M. Teboulle, and X. Wang. On linear convergence of non-euclidean gradient methods without strong convexity and Lipschitz gradient continuity. *Journal of Optimization Theory and Applications*, 182(3):1068–1087, 2019.
- [3] H. H. Bauschke, J. Bolte, and M. Teboulle. A descent lemma beyond Lipschitz gradient continuity: First-order methods revisited and applications. *Mathematics of Operations Research*, page 20, 2016.
- [4] B. Birnbaum, N. R. Devanur, and L. Xiao. Distributed algorithms via gradient descent for fisher markets. In *Proceedings of the 12th ACM conference on Electronic commerce*, pages 127–136, 2011.
- [5] J. Bolte, S. Sabach, M. Teboulle, and Y. Vaisbourd. First order methods beyond convexity and Lipschitz gradient continuity with applications to quadratic inverse problems. *SIAM J. Optim.*, 28(3):2131–2151, 2018.
- [6] E. Candès and X. Li. Solving quadratic equations via phaselift when there are about as many equations as unknowns. *Found. Comput. Math.*, 2014.
- [7] E. Candès, X. Li, and M. Soltanolkotabi. Phase retrieval via Wirtinger flow: Theory and algorithms. *IEEE Trans. Inform. Theory*, 61(4):1985–2007, 2015.
- [8] E. Candès, T. Strohmer, and V. Voroninski. PhaseLift: Exact and stable signal recovery from magnitude measurements via convex programming. *Communications on Pure and Applied Mathematics*, 66(8):1241–1274, 2013.
- [9] Y. Chen and E. Candès. Solving random quadratic systems of equations is nearly as easy as solving linear systems. *Comm. Pure Appl. Math.*, 70(5):822–883, May 2017.
- [10] C. Davis and W. M. Kahan. The rotation of eigenvectors by a perturbation. iii. *SIAM J. Numer. Anal.*, 7:1–4, 1970.
- [11] D. Davis, D. Drusvyatskiy, and C. Paquette. The nonsmooth landscape of phase retrieval. *IMA Journal of Numerical Analysis*, 40(4):2652–2695, October 2020.
- [12] L. Demanet and P. Hand. Stable Optimizationless Recovery from Phaseless Linear Measurements. *MIT web domain*, November 2013.
- [13] A. Fannjiang and T. Strohmer. The numerics of phase retrieval. *Acta Numerica*, 29:125–228, May 2020.
- [14] B. Gao, Y. Wang, and Z. Xu. Stable Signal Recovery from Phaseless Measurements. *J Fourier Anal Appl*, 22(4):787–808, August 2016.
- [15] J.-J. Godeme, J. Fadili, X. Buet, M. Zerrad, M. Lequime, and C. Amra. Provable phase retrieval with mirror descent. *SIAM J. Imaging Sci.*, 16(3):1106–1141, September 2023.
- [16] K. Jaganathan, Eldar Y. C., and B. Hassibi. Phase retrieval: An overview of recent developments. In A. Stern, editor, *Optical Compressive Imaging*. CRC Press, 2016.
- [17] F. Kraher and D. Stöger. Complex Phase Retrieval from Subgaussian Measurements. *J. Fourier Anal Appl*, 26(6):89, November 2020.

- [18] H. Lu, R. M. Freund, and Y. Nesterov. Relatively smooth convex optimization by first-order methods, and applications. *SIAM Journal on Optimization*, 28(1):333–354, 2018.
- [19] D. R. Luke. Phase Retrieval, What’s New? *SIAG/OPT Views and News*, 25(1):1–6, 2017.
- [20] Huang M. and Xu Z. Performance bound of the intensity-based model for noisy phase retrieval, 2021.
- [21] P. Netrapalli, P. Jain, and S. Sanghavi. Phase retrieval using alternating minimization. *IEEE Transactions on Signal Processing*, 63(18):4814–4826, 2015.
- [22] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, 1970.
- [23] H. Sahinoglou and S. Cabrera. On phase retrieval of finite-length sequences using the initial time sample. *IEEE Transactions on Circuits and Systems*, 38(5):954–958, 1991.
- [24] Y Shechtman, Y. C. Eldar, O. Cohen, H. N. Chapman, J. Miao, and M. Segev. Phase retrieval with application to optical imaging: A contemporary overview. *IEEE Signal Processing Magazine*, 32(3):87–109, May 2015.
- [25] A. Silveti-Falls, C. Molinari, and J. Fadili. A stochastic Bregman primal-dual splitting algorithm for composite optimization. *Pure and Applied Functional Analysis (special issue in honor of L. Bregman)*, 2022.
- [26] J. Sun, Q. Qu, and J. Wright. A geometric analysis of phase retrieval. *Found Comput Math*, 18(5):1131–1198, October 2018.
- [27] M. Teboulle. A simplified view of first order methods for optimization. *Mathematical Programming*, 170(1):67–96, 2018.
- [28] N. Vaswani. Non-convex structured phase retrieval. *arXiv:2006.13298 [cs, eess, math, stat]*, June 2020.
- [29] I. Waldspurger. Phase retrieval with random gaussian sensing vectors by alternating projections. *IEEE Transactions on Information Theory*, 64(5):3301–3312, May 2018.
- [30] G. Wang, G. B. Giannakis, and Y. C. Eldar. Solving systems of random quadratic equations via truncated amplitude flow. *arXiv:1605.08285 [cs, math, stat]*, August 2017.
- [31] Y. Xia and Z. Xu. The performance of the amplitude-based model for complex phase retrieval. *Information and Inference: A Journal of the IMA*, 13(1), 01 2024.
- [32] H. Zhang, Y. Liang, and Y. Chi. A nonconvex approach for phase retrieval: Reshaped Wirtinger flow and incremental algorithms. *Journal of Machine Learning Research*, 18(141):1–35, 2017.