# ON SMOOTHED AMPLITUDE FLOW MODELS FOR PHASE RETRIEVAL

JIANFENG CAI, MENG HUANG, DONG LI, AND YANG WANG

Abstract. The phase retrieval problem is concerned with recovering an unknown signal $\boldsymbol{x} \in \mathbb{R}^n$ from a set of magnitude-only measurements $y_i = |\langle \boldsymbol{a}_i, \boldsymbol{x} \rangle|, \ i = 1, \ldots, m$. If the measurement vectors $\boldsymbol{a}_i$ are i.i.d. standard Gaussian, Sun-Qu-Wright [26] showed that the landscape of an intensity flow model has a benign global geometry provided $m \geq Cn \log^3 n$. In the first part of this work we improve the sampling threshold to $m \geq Cn \log n$.

In the second part of this paper we introduce three smoothed amplitude flow models having well-tamed geometric landscapes under the optimal sampling complexity. In particular, we show that when the measurements $\boldsymbol{a}_i \in \mathbb{R}^n$ are Gaussian random vectors and the number of measurements $m \geq Cn$, our smoothed amplitude flow model admits no spurious local minimizers with high probability, i.e., the target solution $\boldsymbol{x}$ is the unique global minimizer (up to a global phase) and the loss function has a negative directional curvature around each saddle point. Due to this benign geometric landscape, the phase retrieval problem can be solved by the gradient descent algorithms without spectral initialization. Numerical experiments show that the gradient descent algorithm with random initialization outperforms state-of-the-art algorithms with spectral initialization in empirical success rate and convergence speed.

## Contents



J.F. Cai was supported in part by Hong Kong RGC grants GRF 16306317 and 16309219.
D. Li was supported in part by Hong Kong RGC grants GRF 16307317 and 16309518.
Y. Wang was supported in part by Hong Kong RGC grants GRF 16306415 and 16308518.

## 1. Introduction

1.1. **Background.** In a prototypical phase retrieval problem, one is interested in how to recover an unknown signal $\boldsymbol{x} \in \mathbb{R}^n$ from a series of magnitude-only measurements

$$y_i = |\langle \boldsymbol{a}_i, \boldsymbol{x} \rangle|, \ i = 1, \dots, m,$$

where $\boldsymbol{a}_i \in \mathbb{R}^n, i = 1, \dots, m$ are given vectors and $m$ is the number of measurements. This problem arises in many fields of science and engineering such as X-ray crystallography [16, 21], microscopy [20], astronomy [7], coherent diffractive imaging [15, 25] and optics [30] etc. In practical applications due to the physical limitations optical detectors can only record the magnitude of signals while losing the phase information. Despite its simple mathematical formulation, it has been shown that reconstructing a finite-dimensional discrete signal from the magnitude of its Fourier transform is generally an *NP-complete* problem [24].

Many algorithms have been designed to solve the phase retrieval problem. They generally fall into two categories: convex algorithms and non-convex ones. The convex algorithms usually rely on a "matrix-lifting" technique, which lifts the phase retrieval problem into a low rank matrix recovery problem. By using convex relaxation one can show that the matrix recovery problem under suitable conditions is equivalent to a convex optimization problem. The corresponding algorithms include PhaseLift [2, 4], PhaseCut [29] etc. It has been shown [2] that PhaseLift can achieve the exact recovery under the optimal sampling complexity with Gaussian random measurements.

Although convex methods have good theoretical guarantees to converge to the true solutions under some special conditions, they tend to be computationally inefficient for large scale problems. In contrast, many non-convex algorithms bypass the lifting step and operate directly on the lower-dimensional ambient space, making them much more computationally efficient. Early non-convex algorithms were mostly based on the technique of alternating projections, e.g. Gerchberg-Saxton [14] and Fineup [9]. The main drawback, however, is the lack of theoretical guarantee. Later Netrapalli et al [22] proposed the AltMinPhase algorithm based on a technique known as *spectral initialization*. They proved that the algorithm linearly converges to the true solution with $O(n\log^3 n)$ resampling Gaussian random measurements. This work led further to several other non-convex algorithms based on spectral initialization. A common thread is first choosing a good initial guess through spectral initialization, and then solving an optimization model through gradient descent. Two widely used optimization models are the intensity flow model

$$\min_{\boldsymbol{z}\in\mathbb{R}^n} \; F(\boldsymbol{z}) = \sum_{j=1}^{m} \left( |\langle \boldsymbol{a}_j, \boldsymbol{z}\rangle|^2 - y_j^2 \right)^2 ; \tag{1.1}$$

and the amplitude flow model

$$\min_{\boldsymbol{z}\in\mathbb{R}^n} \; F(\boldsymbol{z}) = \sum_{j=1}^{m} \left( |\langle \boldsymbol{a}_j, \boldsymbol{z}\rangle| - y_j \right)^2 . \tag{1.2}$$

Specifically in a seminal work [3] Candès et al developed the Wirtinger Flow (WF) method based on (1.1) and proved that the WF algorithm can achieve linear convergence with $O(n\log n)$ Gaussian random measurements. Chen and Candès in [6] improved the results to $O(n)$ Gaussian random measurements by incorporating a truncation which leads to a novel Truncated Wirtinger Flow (TWF) algorithm. Other methods based on (1.1) include the Gauss-Newton method [11], the trust-region method [26] and the like. On the other hand for the amplitude flow model (1.2), several algorithms have also been developed recently, such as the Truncated Amplitude Flow (TAF) algorithm [31], the Reshaped Wirtinger Flow (RWF) [32] algorithm and the Perturbed Amplitude Flow (PAF) [10] algorithm. These three algorithms have been shown to converge linearly to the true solution up to a global phase with $O(n)$ Gaussian random measurements.

Furthermore, there is ample evidence from numerical simulations showing that algorithms based on the amplitude flow model (1.2) tend to outperform algorithms based on model (1.1) when measured in empirical success rate and convergence speed.

1.2. **Prior arts and connections.** As was already mentioned earlier, producing a good initial guess using spectral initialization seems to be a prerequisite for prototypical non-convex algorithms such as TAF, RWF and PAF to succeed with good theoretical guarantee. A natural and fundamental question is:

*Is it possible for non-convex algorithms to achieve successful recovery with a random initialization (i.e. without spectral initialization or any additional truncation)*?

For the intensity-based model (1.1), the answer is affirmative. In a pioneering work [26], Ju Sun et al. carried out a deep study of the global geometric structure of the loss function of (1.1). They proved that the loss function $F(\boldsymbol{z})$ does not have any spurious local minima under $O(n\log^3 n)$ Gaussian random measurements. More specifically, it was shown in [26] that all minimizers coincide with the target signal $\boldsymbol{x}$ up to a global phase, and the loss function has a negative directional curvature around each saddle point. Thanks to this benign geometric landscape any algorithm which can avoid saddle points converges to the true solution with high probability. A trust-region method was employed in [26] to find the global minimizers with random initialization. To reduce the sampling complexity, it has been recently shown in [19] that a combination of the loss function (1.1) with a judiciously chosen activation function also possesses the benign geometry structure under $O(n)$ Gaussian random measurements.

The emerging concept of a benign geometric landscape has also recently been explored in many other applications of signal processing and machine learning, e.g. matrix sensing [1, 23], tensor decomposition [12], dictionary learning [27] and matrix completion [13]. For general optimization problems there exist a plethora of loss functions with well-behaved geometric landscapes such that all local optima are also global optima and each saddle point has a negative direction curvature in its vincinity. Correspondingly several techniques have been developed to guarantee that the standard gradient based optimization algorithms can escape such saddle points efficiently, see e.g. [8, 17, 18].

1.3. **Our contributions.** The contributions in this paper are as follows.

- We develop several new techniques and show that the landscape of the intensity model (1.1) has a benign global geometry under the sampling complexity $O(n\log n)$. This refines the Sun-Qu-Wright [26] bound $O(n\log^3 n)$.
- We introduce three novel smoothed amplitude flow models based a deep modification of the model (1.2). The first model employs a loss function which is Lipschitz at the origin, whereas the second and third models involve a globally smooth loss function. In all cases we introduce several new techniques and prove rigorously that the loss function has a benign geometric landscape under the optimal sampling complexity $O(n)$.

- For all models we solve the corresponding phase retrieval problem via vanilla gradient descent algorithms without spectral initialization. We carry out extensive numerical experiments and show that the gradient descent algorithm with random initialization outperforms the state-of-the-art algorithms with spectral initialization in empirical success rate and convergence speed.

We now give a slightly more detailed summary of the main theoretical results proved in this paper.

**Theorem A** (Informal version of Theorem 2.1). Consider the intensity model (1.1). Assume $\{\boldsymbol{a}_i\}_{i=1}^m$ are i.i.d. Gaussian random vectors and $\boldsymbol{x} \neq 0$. If $m \geq Cn\log n$, then with probability at least $1 - O(m^{-2})$, the loss function $F = F(\boldsymbol{z})$ has no spurious local minimizers. The only global minimizer is $\pm\boldsymbol{x}$, and all saddle points are strict saddles, i.e., each saddle point has a neighborhood where the function has negative directional curvature.

One may naturally wonder whether there exists a loss function akin to the model (1.2) whilst beating the $O(n\log n)$ barrier. This is indeed possible. We introduce the following:

$$\min_{\boldsymbol{z}\in\mathbb{R}^n} F(\boldsymbol{z}) = \frac{1}{m}\sum_{i=1}^m \left(\sqrt{\beta|\boldsymbol{z}|^2 + (\boldsymbol{a}_i^T\boldsymbol{z})^2} - \sqrt{\beta|\boldsymbol{z}|^2 + (\boldsymbol{a}_i^T\boldsymbol{x})^2}\right)^2. \tag{1.3}$$

**Theorem B** (Informal version of Theorem 3.1). Consider the smoothed amplitude flow model (1.3). Assume $\{\boldsymbol{a}_i\}_{i=1}^m$ are i.i.d. Gaussian random vectors and $\boldsymbol{x} \neq 0$. Let $0 < \beta < \infty$. If $m \geq Cn$, then with probability at least $1 - O(m^{-2})$, the loss function $F = F(\boldsymbol{z})$ has no spurious local minimizers. The only global minimizer is $\pm\boldsymbol{x}$, and all saddle points are strict saddles.

The avid reader should notice that the probability concentration in Theorem B is only $1 - O(m^{-2})$. Besides, the function is only Lipschitz continuous near the origin.[1] To remedy this and improve the probability concentration, we introduce the following genuinely globally smooth model:

$$\min_{\boldsymbol{z}\in\mathbb{R}^n} F(\boldsymbol{z}) = \frac{1}{m}\sum_{i=1}^m \left(\sqrt{\beta|\boldsymbol{z}|^2 + (\boldsymbol{a}_i^T\boldsymbol{z})^2 + (\boldsymbol{a}_i^T\boldsymbol{x})^2} - \sqrt{\beta|\boldsymbol{z}|^2 + 2(\boldsymbol{a}_i^T\boldsymbol{x})^2}\right)^2. \tag{1.4}$$

**Theorem C** (Informal version of Theorem 4.1). Consider the smoothed amplitude flow model (1.4). Assume $\{\boldsymbol{a}_i\}_{i=1}^m$ are i.i.d. Gaussian random vectors and $\boldsymbol{x} \neq 0$. Let $0 < \beta < \infty$. If $m \geq Cn$, then with probability at least $1 - e^{-cm}$, the loss function $F = F(\boldsymbol{z})$ has no spurious local minimizers. The only global minimizer is $\pm\boldsymbol{x}$, and all other critical points are strict saddles.

Our next model consists of three sub-models all of which are based on the following quotient loss function

$$F(\boldsymbol{z}) = \frac{1}{m}\sum_{i=1}^m \frac{((\boldsymbol{a}_i\cdot\boldsymbol{z})^2 - (\boldsymbol{a}_i\cdot\boldsymbol{x})^2)^2}{\beta|\boldsymbol{z}|^2 + \beta_1(\boldsymbol{a}_i\cdot\boldsymbol{z})^2 + \beta_2(\boldsymbol{a}_k\cdot\boldsymbol{x})^2},$$

where $\beta \geq 0$, $\beta_1 \geq 0$, $\beta_2 \geq 0$ are parameters.

[1] This is due to the function $\sqrt{\beta|\boldsymbol{z}|^2 + (\boldsymbol{a}_i^\top\boldsymbol{z})^2}$.

The first sub-model corresponds to $\beta = 0$, $\beta_1 = 0$, $\beta_2 = 1$. An attractive feature is that it does not contain any parameter. Consider

$$\min_{\boldsymbol{z}\in\mathbb{R}^n} F(\boldsymbol{z}) = \frac{1}{m}\sum_{i=1}^{m} \frac{((\boldsymbol{a}_i\cdot\boldsymbol{z})^2 - (\boldsymbol{a}_i\cdot\boldsymbol{x})^2)^2}{(\boldsymbol{a}_k\cdot\boldsymbol{x})^2}. \tag{1.5}$$

**Theorem D1** (Informal version of Theorem 5.1). Consider the smoothed amplitude model (1.5). Assume $\{\boldsymbol{a}_i\}_{i=1}^m$ are i.i.d. standard Gaussian random vectors and $\boldsymbol{x} \neq 0$. There exist positive absolute constants $c$, $C$, such that if $m \geq Cn$, then with probability at least $1 - e^{-cm}$ the loss function $F = F(\boldsymbol{z})$ has no spurious local minimizers. The only global minimizer is $\pm\boldsymbol{x}$. All other critical points are strict saddles.

The second case corresponds to $\beta > 0$, $\beta_1 = 0$, $\beta_2 = 1$. Namely consider for $\beta > 0$,

$$\min_{\boldsymbol{z}\in\mathbb{R}^n} F(\boldsymbol{z}) = \frac{1}{m}\sum_{i=1}^{m} \frac{((\boldsymbol{a}_i\cdot\boldsymbol{z})^2 - (\boldsymbol{a}_i\cdot\boldsymbol{x})^2)^2}{\beta|\boldsymbol{z}|^2 + (\boldsymbol{a}_k\cdot\boldsymbol{x})^2}. \tag{1.6}$$

**Theorem D2** (Informal version of Theorem 6.1). Consider the smoothed amplitude model (1.6). Let $0 < \beta < \infty$. Assume $\{\boldsymbol{a}_i\}_{i=1}^m$ are i.i.d. standard Gaussian random vectors and $\boldsymbol{x} \neq 0$. There exist positive constants $c$, $C$ depending only on $\beta$, such that if $m \geq Cn$, then with probability at least $1 - e^{-cm}$ the loss function $F = F(\boldsymbol{z})$ has no spurious local minimizers. The only global minimizer is $\pm\boldsymbol{x}$ and all other critical points are strict saddles.

*Remark* 1.1. There appears some subtle differences between model (1.5) and (1.6). Although the former looks more singular, one can prove full strong convexity in the neighborhood of the global minimizers. In the latter case, however, we only have certain restricted convexity.

The third sub-model corresponds to the choice $\beta = 1$, $\beta_1 > 0$, $\beta_2 > 0$ in the full model. Consider for $\beta_1, \beta_2 > 0$,

$$\min_{\boldsymbol{z}\in\mathbb{R}^n} F(\boldsymbol{z}) = \frac{1}{m}\sum_{i=1}^{m} \frac{((\boldsymbol{a}_i\cdot\boldsymbol{z})^2 - (\boldsymbol{a}_i\cdot\boldsymbol{x})^2)^2}{|\boldsymbol{z}|^2 + \beta_1(\boldsymbol{a}_k\cdot\boldsymbol{z})^2 + \beta_2(\boldsymbol{a}_k\cdot\boldsymbol{x})^2}. \tag{1.7}$$

**Theorem D3** (Informal version of Theorem 7.1). Consider the smoothed amplitude model (1.7). Let $0 < \beta_1, \beta_2 < \infty$. Assume $\{\boldsymbol{a}_i\}_{i=1}^m$ are i.i.d. standard Gaussian random vectors and $\boldsymbol{x} \neq 0$. There exist positive constants $c$, $C$ depending only on $\beta$, such that if $m \geq Cn$, then with probability at least $1 - e^{-cm}$ the loss function $F = F(\boldsymbol{z})$ has no spurious local minimizers. The only global minimizer is $\pm\boldsymbol{x}$ and all other critical points are strict saddles.

*Remark* 1.2. For this case, thanks to the strong damping, we have full strong convexity in the neighborhood of the global minimizers.

*Remark* 1.3. In a con-current work [5], we considered yet aother new smoothed amplitude flow model which is based on a piece-wise smooth modification of the amplitude model (1.2). The model takes the form

$$\min_{\boldsymbol{z}\in\mathbb{R}^n} F(\boldsymbol{z}) = \frac{1}{2m}\sum_{i=1}^{m} \left(\gamma\left(\frac{|\boldsymbol{a}_i^\top\boldsymbol{z}|}{|\boldsymbol{a}_i^\top\boldsymbol{x}|}\right) - 1\right)^2 \cdot |\boldsymbol{a}_i^\top\boldsymbol{x}|^2,$$

where the function $\gamma(t)$ is taken to be

$$\gamma(t) := \begin{cases} |t|, & |t| > \beta; \\ \frac{1}{2\beta}t^2 + \frac{\beta}{2}, & |t| \le \beta. \end{cases}$$

For $0 < \beta \le \frac{1}{2}$, we prove that the loss function has a benign landscape under the optimal sampling threshold $m = O(n)$. There are subtle technical difficulties in connection with the piecewise-smoothness of the loss function which make the overall proof therein quite special. On the other hand, there are exciting evidences that the machinery developed in this work can be generalized significantly in various directions (including complex-valued cases etc). We plan to address some of these important issues in forthcoming works.

The rest of this paper is organized as follows. In Section 2 we give the proof of the refined complexity bound $O(n \log n)$ for the intensity model (1.1). In Section 3–4 we introduce first two smoothed amplitude flow models and carry out an in-depth analysis of the corresponding geometric landscape under optimal sampling complexity $O(n)$. In Section 5, 6 and 7 we analyze the third group of smoothed amplitude flow models. Section 8 collects various numerical simulation results on all the aforementioned models. The appendices collect all the technical lemmas needed in the proof and some auxiliary estimates.

## 2. Refinement of the SQW bound

In this section we analyze the intensity model (1.1) and give the refined complexity bound. To alleviate the notation we shall write $\boldsymbol{a}_i$ as $a_i$, $\boldsymbol{x}$ as $x$ and so on.

Consider

$$f(u) = \frac{1}{m}\sum_{j=1}^{m} ((a_j \cdot u)^2 - (a_j \cdot x)^2)^2. \tag{2.1}$$

**Theorem 2.1.** *Assume $\{a_i\}_{i=1}^m$ are i.i.d. standard Gaussian random vectors and $x \neq 0$. There exist positive absolute constants $c$, $C$, such that if $m \ge Cn \log n$, then with probability at least $1 - cm^{-1}$ the loss function $f = f(u)$ defined by (2.1) has no spurious local minimizers. The only global minimizer is $\pm x$, and the loss function is strongly convex in a neighborhood of $\pm x$. The point $u = 0$ is a local maximum point where the the Hessian of the function is strictly negative-definite. All saddle points are strict saddles, i.e., each saddle point has a neighborhood where the function has negative directional curvature.*

*Remark* 2.1. This refines the SQW bound $m = O(n \log^3 n)$ to $O(n \log n)$. For simplicity we consider here only the real-valued case and will investigate the complex-valued case elsewhere. In Theorem 2.1 the probability bound $O(m^{-1})$ can be refined.

*Remark* 2.2. Another interesting issue is to show the measurements are non-adaptive, i.e., a single realization of measurement vectors $a_i$ can be used to reconstruct all $0 \neq x \in \mathbb{R}^n$. However we shall not dwell on this refinement here for simplicity.

In the rest of this section we shall carry out the proof of Theorem 2.1 in several steps.

**Notation**. In the proof below we shall adopt the following convention.

- We write $u \in \mathbb{S}^{n-1}$ if $u \in \mathbb{R}^n$ and $\|u\|_2 = \sqrt{\sum_j (u_j)^2} = 1$.
- We use $\chi$ to denote the usual characteristic function. For example $\chi_A(x) = 1$ if $x \in A$ and $\chi_A(x) = 0$ if $x \notin A$.
- We denote by $\epsilon$, $\eta$, $\eta_1$ various absolute constants whose value will be taken sufficiently small. The needed smallness will be clear from the context.
- For any quantity $X$, we shall write $X = O(Y)$ if $|X| \le CY$ for some constant $C > 0$. We shall write $X \ll Y$ if $X \le cY$ where the constant $c > 0$ will be sufficiently small.
- In our proof it is important for us to specify the precise dependence of the sampling size $m$ in terms of the dimension $n$. For this purpose we shall write $m \gtrsim n$ if $m \ge Cn$ where the constant $C$ is allowed to depend on the small absolute constants $\epsilon$, $\epsilon_i$ etc used in the argument. One can extract more explicit dependence of $C$ on the small constants but for simplicity we suppress this dependence here. In a similar vein we write $m \gtrsim n \ln n$ if $m \ge Cn \ln n$.
- We shall say an event $A$ happens with **good probability** if

$$\mathbb{P}(A^c) \le \frac{C}{m^2},$$

where $C > 0$ is a constant. Clearly if $A_i$, $i = 1, \cdots, k_0$, are such that

$$\mathbb{P}(A_i^c) \le \frac{C_i}{m^2},$$

Then $B = \bigcap_{i=1}^{k_0} A_i$ happens with good probability, since

$$\mathbb{P}(B^c) \le \frac{1}{m^2} \cdot (\sum_{i=1}^{k_0} C_i).$$

In our proof, $k_0$ will be finite, and the desired conclusion happens with probability at least $1 - O(m^{-2})$.

We now start the proof. Without loss of generality we can assume $x = e_1 = (1, 0, \cdots, 0)$. For $u \neq 0$, denote

$$u = \sqrt{R}\hat{u}, \qquad \hat{u} \in \mathbb{S}^{n-1};$$
$$t = \hat{u} \cdot e_1.$$

We denote the Hessian of the function $f(u)$ along the $\xi$ direction as

$$\begin{aligned} H_{\xi\xi}(u) &= \sum_{k,l} \xi_k \xi_l (\partial_{kl} f)(u) \\ &= 4\frac{1}{m} \sum_{j=1}^{m} (a_j \cdot \xi)^2 (3(a_j \cdot u)^2 - (a_j \cdot e_1)^2). \end{aligned}$$

First observe that $u = 0$ is a local maximum. By Corollary A.2 in the appendix, if $m \gtrsim n$, then it holds with good probability that

$$H_{\xi\xi}(0) = -\frac{4}{m}\sum_{j=1}^{m}(a_j \cdot \xi)^2(a_j \cdot e_1)^2 \le -c_1 < 0, \qquad \forall\, \xi \in \mathbb{S}^{n-1}.$$

In particular this shows that with good probability the Hessian $(\partial_{kl} f)(0)$ is strictly negative definite. Thus we do not need to consider $u = 0$ and can assume $u \neq 0$ in the rest of our argument.

Now for $u \neq 0$, we write

$$\frac{1}{4}H_{\xi\xi}(u) = 3R \cdot \frac{1}{m}\sum_{j=1}^{m}(a_j \cdot \xi)^2(a_j \cdot \hat{u})^2 - \frac{1}{m}\sum_{j=1}^{m}(a_j \cdot \xi)^2(a_j \cdot e_1)^2.$$

2.1. **The regime $|\hat{u} \cdot e_1| < \sqrt{\frac{\sqrt{3}-1}{2}} - \epsilon$ is fine.** We rewrite $f(u)$ as

$$\begin{aligned} f &= R^2\Big(\frac{1}{m}\sum_{j=1}^{m}(a_j \cdot \hat{u})^4\Big) - 2R\Big(\frac{1}{m}\sum_{j=1}^{m}(a_j \cdot \hat{u})^2(a_j \cdot e_1)^2\Big) + \frac{1}{m}\sum_{k=1}^{m}(a_j \cdot e_1)^4 \\ &=: R^2 A - 2RB + \frac{1}{m}\sum_{k=1}^{m}(a_j \cdot e_1)^4. \end{aligned}$$

If at some $R \neq 0$ we have a critical point, then apparently (by Lemma A.11, it holds with good probability that $A \ge 1$)

$$\partial_R f = 0 \Rightarrow\ R = \frac{B}{A}. \tag{2.2}$$

On the other hand, the Hessian at this point along the direction $e_1$ is

$$\frac{1}{4}H_{e_1e_1} = 3RB - \frac{1}{m}\sum_{j=1}^{m}(a_j \cdot e_1)^4 = \frac{1}{A}(3B^2 - A \cdot C),$$

where

$$C = \frac{1}{m}\sum_{j=1}^{m}(a_j \cdot e_1)^4, \quad B = \frac{1}{m}\sum_{j=1}^{m}(a_j \cdot e_1)^2(a_j \cdot \hat{u})^2, \quad A = \frac{1}{m}\sum_{j=1}^{m}(a_j \cdot \hat{u})^4.$$

Recall $t = \hat{u} \cdot e_1$ and note that $\mathbb{E}B = 2t^2 + 1, \quad \mathbb{E}A = \mathbb{E}C = 3.$

**Theorem 2.2.** *Assume $|t| \le \sqrt{\frac{\sqrt{3}-1}{2}} - \eta_1$ and $0 < \eta_1 < \frac{1}{8}$ is a fixed absolute constant which is taken to be sufficiently small. Then for $m \gtrsim n$, we have with good probability,*

$$3B^2 - A \cdot C < 0, \qquad \forall\, \hat{u} \in \mathbb{S}^{n-1}.$$

*Proof.* Take $\phi \in C_c^\infty(\mathbb{R})$, $0 \le \phi \le 1$ for all $z \in \mathbb{R}$, $\phi \equiv 1$ for $|z| \le 1$ and $\phi \equiv 0$ for $|z| > 2$. We write

$$\begin{aligned} B &= \frac{1}{m}\sum_{j=1}^{m}(a_j \cdot e_1)^2\phi\Big(\frac{a_j \cdot e_1}{N}\Big)(a_j \cdot \hat{u})^2\phi\Big(\frac{a_j \cdot \hat{u}}{N}\Big) + \frac{1}{m}\sum_{j=1}^{m}(a_j \cdot e_1)^2\Big(1 - \phi\Big(\frac{a_j \cdot e_1}{N}\Big)\Big)(a_j \cdot \hat{u})^2\phi\Big(\frac{a_j \cdot \hat{u}}{N}\Big) \\ &\quad + \frac{1}{m}\sum_{j=1}^{m}(a_j \cdot e_1)^2(a_j \cdot \hat{u})^2\Big(1 - \phi\Big(\frac{a_j \cdot \hat{u}}{N}\Big)\Big) =:\ B_1 + B_2 + B_3. \end{aligned}$$

Note that

$$\begin{aligned}
&|\mathbb{E}B_1 - \mathbb{E}B| \\
=&|\mathbb{E}(a_1 \cdot e_1)^2 (a_1 \cdot \hat{u})^2 \cdot ((\phi(\frac{a_1 \cdot e_1}{N}) - 1)\phi(\frac{a_1 \cdot \hat{u}}{N}) + \phi(\frac{a_1 \cdot \hat{u}}{N}) - 1)| \\
\leq&\mathbb{E}(a_1 \cdot e_1)^2 (a_1 \cdot \hat{u})^2 \cdot (\chi_{|a_1 \cdot e_1| \geq N} + \chi_{|a_1 \cdot \hat{u}| \geq N}).
\end{aligned}$$

Thus by Lemma A.6 in the appendix, we have for $N$ sufficiently large (depending only on $\epsilon_1$), it holds that

$$|\mathbb{E}B_1 - \mathbb{E}B| = |\mathbb{E}B_1 - (2t^2 + 1)| \leq \epsilon_1.$$

By Lemma A.10, for $N$ sufficiently large depending only on $\epsilon_1$, $m \gtrsim n$ and with good probability, it holds that

$$\begin{aligned}
&\frac{1}{m}\sum_{j=1}^{m}(a_j \cdot e_1)^4 \chi_{|a_j \cdot e_1| > N} \leq \epsilon_1; \\
&\frac{1}{m}\sum_{j=1}^{m}(a_j \cdot e_1)^4 \chi_{|a_j \cdot \hat{u}| > N} \leq \epsilon_1, \qquad \forall\, \hat{u} \in \mathbb{S}^{n-1}.
\end{aligned}$$

Thus for $B_2$ and $B_3$, we have (with good probability)

$$\begin{aligned}
B_2^2 \leq\ & (\frac{1}{m}\sum_{j=1}^{m}(a_j \cdot \hat{u})^4)(\frac{1}{m}\sum_{j=1}^{m}(a_j \cdot e_1)^4 \chi_{|a_j \cdot e_1| > N}) \leq \epsilon_1 A; \\
B_3^2 \leq\ & (\frac{1}{m}\sum_{j=1}^{m}(a_j \cdot \hat{u})^4)(\frac{1}{m}\sum_{j=1}^{m}(a_j \cdot e_1)^4 \chi_{|a_j \cdot \hat{u}| > N}) \leq \epsilon_1 A.
\end{aligned}$$

Thanks to the smooth cut-off, $B_1$ is OK for union bounds. By Lemma A.12, for $m \gtrsim n$, it holds with good probability that

$$|B_1 - \mathbb{E}B_1| \leq \epsilon_1, \qquad \forall\, \hat{u} \in \mathbb{S}^{n-1}.$$

Thus with good probability, it holds that

$$3B^2 \leq 3(2t^2 + 1 + \epsilon_1 + 2\sqrt{\epsilon_1 A})^2 = 3(2t^2 + 1)^2 + O(\epsilon_1) + O(\epsilon_1) \cdot A.$$

Therefore with good probability, we have

$$3B^2 - AC \leq 3(2t^2 + 1)^2 + O(\epsilon_1) - A \cdot (C - O(\epsilon_1)).$$

Note that the term $C$ only involves a fixed vector $e_1$. Clearly by Lemma A.9, if $m \gtrsim n$, then with good probability, it holds that

$$C \geq 3 - \epsilon_1.$$

On the other hand, for the term $A$ (for $m \gtrsim n$), by Lemma A.11, with good probability we have

$$A \geq 3 - \epsilon_1, \qquad \forall\, \hat{u} \in \mathbb{S}^{n-1}.$$

It follows that (below we assume $|t| \le \sqrt{\frac{\sqrt{3}-1}{2}} - \eta_1$ and $\alpha > 0$ is an absolute constant)

$$\begin{aligned} 3B^2 - A \cdot C &\le 3 \cdot (2t^2+1)^2 + 9 + O(\epsilon_1) \\ &\le -\alpha \cdot \eta_1 + O(\epsilon_1) < 0, \end{aligned}$$

if we take $\epsilon_1$ sufficiently small. $\square$

2.2. **Localization of $\hat{u}\cdot e_1$ and $R$.** The goal of this section is to localize (i.e. in good probability) $|\hat{u} \cdot e_1|$ to near 1 and $R$ to near 1. For a unit vector $\xi$, we have (below $\partial_\xi f = \xi \cdot \nabla f$)

$$\frac{1}{4}\partial_\xi f = \frac{1}{m}\sum_{j=1}^m (a_j \cdot u)^3 (a_j \cdot \xi) - \frac{1}{m}\sum_{j=1}^m (a_j \cdot u)(a_j \cdot \xi)(a_j \cdot e_1)^2.$$

At any potential critical point $u = \sqrt{R}\hat{u}$, we should have $\nabla f = 0$. Thus taking $\xi = \hat{u}$ and $\xi = e_1$ respectively gives

$$\begin{aligned} &R\frac{1}{m}\sum_{j=1}^m (a_j \cdot \hat{u})^4 - \frac{1}{m}\sum_{j=1}^m (a_j \cdot \hat{u})^2 (a_j \cdot e_1)^2 = 0; \\ &R\frac{1}{m}\sum_{j=1}^m (a_j \cdot \hat{u})^3 (a_j \cdot e_1) - \frac{1}{m}\sum_{j=1}^m (a_j \cdot \hat{u})(a_j \cdot e_1)^3 = 0. \end{aligned}$$

We then derive the fundamental relation for any critical point $u = \sqrt{R}\hat{u}$:

$$\Big(\frac{1}{m}\sum_{j=1}^m (a_j \cdot \hat{u})^2 (a_j \cdot e_1)^2\Big) \cdot \Big(\frac{1}{m}\sum_{j=1}^m (a_j \cdot e_1)(a_j \cdot \hat{u})^3\Big) - \Big(\frac{1}{m}\sum_{j=1}^m (a_j \cdot \hat{u})^4\Big) \cdot \underbrace{\Big(\frac{1}{m}\sum_{j=1}^m (a_j \cdot \hat{u})(a_j \cdot e_1)^3\Big)}_{=:\tilde{F}} = 0. \tag{2.3}$$

Write $\hat{u} = se_1^\perp \pm \sqrt{1-s^2}e_1$, where $e_1^\perp \in \mathbb{S}^{n-1}$ satisfies $e_1^\perp \cdot e_1 = 0$. By Theorem 2.2 we may confine ourselves to the regime where $|s|$ is away (say $|s| > 1 - 10^{-4}$) from the end-point $|s| = 1$ (which corresponds to $t = 0$). Note that by Corollary A.3 the term $\tilde{F}$ is well under control.

Define

$$\mathcal{C}_1 = \Big\{ u \in \mathbb{S}^{n-1} : \quad u \text{ satisfies (2.3) and } |u \cdot e_1| > 10^{-4} \Big\}.$$

Here the cut-off $10^{-4}$ is for convenience only. It can be replaced by any sufficiently small absolute constant. The choice of this number can affect other absolute constants used in the proof below. In order not to overburden the reader with too many constants we fix this particular number and will not pursue the most general case.

**Theorem 2.3** (Rigidity of the critical points)**.** *Let $0 < \eta \ll 1$. If $m \gtrsim n$, then the following hold with good probability:*

*For any $\hat{u} \in \mathcal{C}_1$, we have $|s| \le \eta$, and*

$$3 - \eta \le \frac{1}{m}\sum_{j=1}^{m} |a_j \cdot \hat{u}|^4 \le 3 + \eta;$$

$$3 - \eta \le \frac{1}{m}\sum_{j=1}^{m} (a_j \cdot \hat{u})^2 (a_j \cdot e_1)^2 \le 3 + \eta;$$

$$3 - \eta \le \frac{1}{m}\sum_{j=1}^{m} (a_j \cdot \hat{u})^3 (a_j \cdot e_1) \le 3 + \eta;$$

*Moreover,*

$$\frac{1}{m}\sum_{j=1}^{m} (|a_j \cdot \hat{u}|^4 + |a_j \cdot e_1|^4)\chi_{|a_j \cdot \hat{u}| > N} \le \eta,$$

*where $N > 0$ is a sufficiently large constant depending only on $\eta$.*

*Furthermore, if $0 \ne u = \sqrt{R}\hat{u}$ is a critical point with $\hat{u} \in \mathcal{C}_1$, then*

$$1 - \eta \le R \le 1 + \eta.$$

*Proof.* Observe that

$$\mathbb{E}\frac{1}{m}\sum_{j=1}^{m} (a_j \cdot \hat{u})^2 (a_j \cdot e_1)^2 = 2t^2 + 1 = 3 - 2s^2;$$

$$\mathbb{E}\frac{1}{m}\sum_{j=1}^{m} (a_j \cdot e_1)(a_j \cdot \hat{u})^3 = 3\sqrt{1 - s^2};$$

$$\mathbb{E}\frac{1}{m}\sum_{j=1}^{m} (a_j \cdot \hat{u})^4 = 3; \quad \mathbb{E}\frac{1}{m}\sum_{j=1}^{m} (a_j \cdot \hat{u})(a_j \cdot e_1)^3 = 3\sqrt{1 - s^2}.$$

Then

$$(\mathbb{E}\frac{1}{m}\sum_{j=1}^{m} (a_j \cdot \hat{u})^2 (a_j \cdot e_1)^2) \cdot (\mathbb{E}\frac{1}{m}\sum_{j=1}^{m} (a_j \cdot e_1)(a_j \cdot \hat{u})^3) - (\mathbb{E}\frac{1}{m}\sum_{j=1}^{m} (a_j \cdot \hat{u})^4) \cdot (\mathbb{E}\frac{1}{m}\sum_{j=1}^{m} (a_j \cdot \hat{u})(a_j \cdot e_1)^3)$$
$$= 3\sqrt{1 - s^2}(-2s^2).$$

Denote

$$A = \frac{1}{m}\sum_{j=1}^{m} (a_j \cdot \hat{u})^4, \quad B = \frac{1}{m}\sum_{j=1}^{m} (a_j \cdot \hat{u})^2 (a_j \cdot e_1)^2;$$

$$A_1 = \frac{1}{m}\sum_{j=1}^{m} (a_j \cdot \hat{u})^3 (a_j \cdot e_1).$$

Let $\phi \in C_c^\infty(\mathbb{R})$ be such that $0 \le \phi(x) \le 1$ for all $x \in \mathbb{R}$, $\phi(x) = 1$ for $|x| \le 1$ and $\phi(x) = 0$ for $|x| \ge 2$. Then

$$B = \frac{1}{m}\sum_{j=1}^{m} (a_j \cdot \hat{u})^2 (a_j \cdot e_1)^2 \phi(\frac{a_j \cdot \hat{u}}{N})\phi(\frac{a_j \cdot \hat{e}_1}{N}) + r_B,$$

$$A_1 = \frac{1}{m}\sum_{j=1}^{m} (a_j \cdot \hat{u})^3 (a_j \cdot e_1)\phi(\frac{a_j \cdot \hat{u}}{N}) + r_1,$$

where

$$r_B \leq \frac{1}{m}\sum_{j=1}^{m}(a_j\cdot\hat{u})^2(a_j\cdot e_1)^2(\chi_{|a_j\cdot u|\geq N}+\chi_{|a_j\cdot e_1|\geq N});$$

$$r_1 \leq \frac{1}{m}\sum_{j=1}^{m}|a_j\cdot\hat{u}|^3|a_j\cdot e_1|\cdot\chi_{|a_j\cdot\hat{u}|\geq N}.$$

Now note that

$$\begin{aligned}
\sum_{j=1}^{m}(a_j\cdot\hat{u})^2\chi_{|a_j\cdot\hat{u}|>N}(a_j\cdot e_1)^2 &= \sum_{j=1}^{m}|a_j\cdot\hat{u}|^{\frac{2}{3}}\chi_{|a_j\cdot\hat{u}|>N}\cdot(a_j\cdot e_1)^2\cdot|a_j\cdot\hat{u}|^{\frac{4}{3}}\\
&\leq (\sum_{j=1}^{m}|a_j\cdot\hat{u}|\chi_{|a_j\cdot\hat{u}|>N}|a_j\cdot e_1|^3)^{\frac{2}{3}}(\sum_{j=1}^{m}|a_j\cdot\hat{u}|^4)^{\frac{1}{3}};\\
\sum_{j=1}^{m}|a_j\cdot\hat{u}|^3\chi_{|a_j\cdot\hat{u}|>N}|a_j\cdot e_1| &= \sum_{j=1}^{m}|a_j\cdot\hat{u}|^{\frac{1}{3}}\chi_{|a_j\cdot\hat{u}|>N}|a_j\cdot e_1|\cdot|a_j\cdot\hat{u}|^{\frac{8}{3}}\\
&\leq (\sum_{j=1}^{m}|a_j\cdot\hat{u}|\chi_{|a_j\cdot\hat{u}|>N}|a_j\cdot e_1|^3)^{\frac{1}{3}}(\sum_{j=1}^{m}|a_j\cdot\hat{u}|^4)^{\frac{2}{3}};\\
\sum_{j=1}^{m}(a_j\cdot\hat{u})^2\chi_{|a_j\cdot e_1|>N}(a_j\cdot e_1)^2 &= \sum_{j=1}^{m}|a_j\cdot\hat{u}|^{\frac{2}{3}}\chi_{|a_j\cdot e_1|>N}(a_j\cdot e_1)^2\cdot|a_j\cdot\hat{u}|^{\frac{4}{3}}\\
&\leq (\sum_{j=1}^{m}|a_j\cdot\hat{u}|\chi_{|a_j\cdot e_1|>N}|a_j\cdot e_1|^3)^{\frac{2}{3}}(\sum_{j=1}^{m}|a_j\cdot\hat{u}|^4)^{\frac{1}{3}}.
\end{aligned}$$

Thus by Lemma A.15 and Lemma A.12, we can derive (for $m \gtrsim n$ and with good probability)

$$(3-2s^2+O(\epsilon)A^{\frac{1}{3}})\cdot(3\sqrt{1-s^2}+O(\epsilon)A^{\frac{2}{3}}) = A(3\sqrt{1-s^2}+O(\epsilon)). \tag{2.4}$$

By Lemma A.11, for $m \gtrsim n$ and with good probability, it holds that

$$A \geq 3-\epsilon.$$

In particular, we have $A \geq 1$ in good probability. Then we can simplify (2.4) as

$$(3-2s^2)\cdot 3\sqrt{1-s^2} = A\cdot(3\sqrt{1-s^2}+O(\epsilon)) \geq (3-\epsilon)(3\sqrt{1-s^2}+O(\epsilon)). \tag{2.5}$$

Since $\hat{u}\in\mathcal{C}_1$, we have $\sqrt{1-s^2} > 10^{-4}$. It then follows easily that $|s| \ll 1$ (by taking $\epsilon > 0$ sufficiently small). Also it follows from the equality in (2.5) that

$$A \leq 3-2s^2+\frac{1}{1-s^2}O(\epsilon) \leq 3+\eta_0.$$

Thus the desired two-way bound for $A$ follows.

Since

$$B = 3-2s^2+O(\epsilon)A^{\frac{1}{3}}, \quad A_1 = 3\sqrt{1-s^2}+O(\epsilon)A^{\frac{2}{3}},$$

the desired two-way bounds for $B$ and $A_1$ also hold.

Next the bound

$$\frac{1}{m}\sum_{j=1}^{m}|a_j\cdot e_1|^4\chi_{|a_j\cdot\hat{u}|>N} \ll 1$$

follows from Lemma A.10.

Next since

$$\Big|\frac{1}{m}\sum_{j=1}^{m}|a_j\cdot\hat{u}|^4-3|\ll 1,$$

and by Lemma A.8 and we have (taking $N$ large)

$$\Big|\frac{1}{m}\sum_{j=1}^{m}(a_j\cdot\hat{u})^4\phi(\frac{2a_j\cdot\hat{u}}{N})-3\Big|\ll 1,$$

it follows that

$$\frac{1}{m}\sum_{j=1}^{m}(a_j\cdot\hat{u})^4\chi_{|a_j\cdot\hat{u}|\geq N}\ll 1.$$

Finally if $0\neq u=\sqrt{R}\hat{u}$ is a critical point, then by (2.2), we have

$$R=\frac{B}{A}.$$

Since $\hat{u}\in\mathcal{C}_1$, and we have already shown that $B$ and $A$ are well-localized (i.e. close to their expectation in good probability), it follows that

$$1-\eta\leq R\leq 1+\eta.$$

□

2.3. **The regime $|R-1|\ll 1$, $||\hat{u}\cdot e_1|-1|\ll 1$.**

**Theorem 2.4.** *For $m\gtrsim n\ln n$, the following hold with good probability:*

*If $u=\sqrt{R}\hat{u}$ with $|R-1|\ll 1$ and $|\hat{u}\pm e_1|\ll 1$, then*

$$H_{\xi\xi}(u)\geq c_1>0,\qquad\forall\,\xi\in\mathbb{S}^{n-1},$$

*where $c_1>0$ is an absolute constant.*

*Proof of Theorem 2.4.* Recall that along any direction $\xi\in\mathbb{S}^{n-1}$,

$$\begin{aligned}H_{\xi\xi}(u)&=\sum_{k,l}\xi_k\xi_l(\partial_{kl}f)(u)\\&=4\frac{1}{m}\sum_{j=1}^{m}(a_j\cdot\xi)^2(3(a_j\cdot u)^2-(a_j\cdot e_1)^2).\end{aligned}$$

The idea is to work with localization. Since we are looking for a lower bound, the first quartic term (in $a_j$) has a favorable sign and is amenable to localization.

Step 1. Take $\phi\in C_c^\infty(\mathbb{R})$ such that $0\leq\phi(x)\leq 1$ for all $x$, $\phi(x)=1$ for $|x|\leq 1$ and $\phi(x)=0$ for $|x|\geq 2$. For any given small $\epsilon>0$, we claim that there exists $N>0$ sufficiently large (depending only on $\epsilon$) such that

$$\mathbb{E}((a_1\cdot\xi)^2\phi(\frac{a_1\cdot\xi}{N})(a_1\cdot\hat{u})^2\phi(\frac{a_1\cdot\hat{u}}{N})\geq\mathbb{E}(a_1\cdot\xi)^2(a_1\cdot\hat{u})^2-\epsilon,\qquad\forall\,\xi,\hat{u}\in S^{n-1}.$$

Indeed

$$\begin{aligned}
&\Big|\mathbb{E}((a_1\cdot\xi)^2\phi(\frac{a_1\cdot\xi}{N})(a_1\cdot\hat{u})^2\phi(\frac{a_1\cdot\hat{u}}{N})-\mathbb{E}(a_1\cdot\xi)^2(a_1\cdot\hat{u})^2\Big|\big|\\
=&|\mathbb{E}(a_1\cdot\xi)^2(a_1\cdot\hat{u})^2\cdot((\phi(\frac{a_1\cdot\xi}{N})-1)\phi(\frac{a_1\cdot\hat{u}}{N})+\phi(\frac{a_1\cdot\hat{u}}{N})-1)|\\
\le&\mathbb{E}(a_1\cdot\xi)^2(a_1\cdot\hat{u})^2\cdot(\chi_{|a_1\cdot\xi|\ge N}+\chi_{|a_1\cdot\hat{u}|\ge N}).
\end{aligned}$$

Thus by Lemma A.6 in the appendix the desired smallness easily follows.

Step 2. By Lemma A.17, we have (for $m\gtrsim n\ln n$ and with good probability)

$$\begin{aligned}
\frac{1}{4}H_{\xi\xi}(u)&=3R\cdot\frac{1}{m}\sum_{j=1}^m(a_j\cdot\xi)^2(a_j\cdot\hat{u})^2-\frac{1}{m}\sum_{j=1}^m(a_j\cdot\xi)^2(a_j\cdot e_1)^2\\
&\ge 3\cdot(1-\epsilon)\frac{1}{m}\sum_{j=1}^m(a_j\cdot\xi)^2\phi(\frac{a_j\cdot\xi}{N})(a_j\cdot\hat{u})^2\phi(\frac{a_j\cdot\hat{u}}{N})-\mathbb{E}(a_1\cdot\xi)^2(a_1\cdot e_1)^2-\epsilon,\quad\forall\,\xi,\hat{u}\in S^{n-1}.
\end{aligned}$$

By Lemma A.12, we have for $m\gtrsim n$ and with good probability, it holds that

$$\Big|\frac{1}{m}\sum_{j=1}^m(a_j\cdot\xi)^2\phi(\frac{a_j\cdot\xi}{N})(a_j\cdot\hat{u})^2\phi(\frac{a_j\cdot\hat{u}}{N})-\mathbb{E}((a_1\cdot\xi)^2\phi(\frac{a_1\cdot\xi}{N})(a_1\cdot\hat{u})^2\phi(\frac{a_1\cdot\hat{u}}{N})\Big|\le\epsilon,\qquad\forall\,\xi,\hat{u}\in\mathbb{S}^{n-1}.$$

Thus we obtain with good probability,

$$\frac{1}{4}H_{\xi\xi}(u)\ge 3\cdot(1-\epsilon)(\mathbb{E}(a_1\cdot\xi)^2(a_1\cdot\hat{u})^2-2\epsilon)-\mathbb{E}(a_1\cdot\xi)^2(a_1\cdot e_1)^2-\epsilon.$$

Without loss of generality we assume $|\hat{u}-e_1|\ll 1$. The case $|\hat{u}+e_1|\ll 1$ is similar.

Now

$$(a_1\cdot\hat{u})^2=(a_1\cdot e_1+a_1\cdot(\hat{u}-e_1))^2\ge(a_1\cdot e_1)^2+2(a_1\cdot e_1)(a_1\cdot(\hat{u}-e_1)).$$

Thus

$$\frac{1}{4}H_{\xi\xi}(u)\ge(2-3\epsilon)\mathbb{E}(a_1\cdot\xi)^2(a_1\cdot e_1)^2+O(\epsilon)+6(1-\epsilon)\mathbb{E}(a_1\cdot\xi)^2(a_1\cdot e_1)(a_1\cdot(\hat{u}-e_1)).$$

Obviously since $|\hat{u}-e_1|\ll 1$, we have

$$|6(1-\epsilon)\mathbb{E}(a_1\cdot\xi)^2(a_1\cdot e_1)(a_1\cdot(\hat{u}-e_1))|\ll 1.$$

On the other hand, we have

$$\mathbb{E}(a_1\cdot\xi)^2(a_1\cdot e_1)^2=1+2(\xi\cdot e_1)^2\ge 1.$$

The desired result then easily follows by taking $\epsilon>0$ sufficiently small. $\square$

Our next result shows that one can control the Hessian along the $e_1$-direction under the mere assumption that $m\gtrsim n$.

**Theorem 2.5.** *For $m\gtrsim n$, the following hold with good probability:*

*If $u=\sqrt{R}\hat{u}$ with $|R-1|\ll 1$ and $|\hat{u}\pm e_1|\ll 1$, then*

$$H_{e_1e_1}(u)\ge c_1>0,$$

*where $c_1>0$ is an absolute constant.*

*Proof of Theorem 2.5.* We use the parametrization $\hat{u} = se_1^{\perp} \pm \sqrt{1-s^2}e_1$. Without loss of generality we consider the case $\hat{u} = se_1^{\perp} + \sqrt{1-s^2}e_1$. The argument for the other case is similar and therefore omitted.

Now

$$\begin{aligned}\frac{1}{4}H_{e_1e_1}(u) &= R\frac{3}{m}\sum_{j=1}^{m}(a_j\cdot e_1)^2(a_j\cdot\hat{u})^2 - \frac{1}{m}\sum_{j=1}^{m}(a_j\cdot e_1)^4\\ &\geq 3(1-\epsilon)\frac{1}{m}\sum_{j=1}^{m}(a_j\cdot e_1)^2(sa_j\cdot e_1^{\perp}+\sqrt{1-s^2}a_j\cdot e_1)^2 - \frac{1}{m}\sum_{j=1}^{m}(a_j\cdot e_1)^4.\end{aligned}$$

Now observe

$$(sa_j\cdot e_1^{\perp}+\sqrt{1-s^2}a_j\cdot e_1)^2 \geq (1-s^2)(a_j\cdot e_1)^2 + 2s\sqrt{1-s^2}(a_j\cdot e_1^{\perp})(a_j\cdot e_1).$$

By Lemma A.14, we have for $m \gtrsim n$ and with high probability,

$$\frac{1}{m}\sum_{j=1}^{m}(a_j\cdot e_1)^3(a_j\cdot e_1^{\perp}) \geq -\frac{\epsilon}{100}.$$

Thus

$$\frac{1}{4}H_{e_1e_1}(u) \geq (3(1-\epsilon)(1-s^2)-1)\frac{1}{m}\sum_{j=1}^{m}(a_j\cdot e_1)^4 - \epsilon.$$

The desired result then easily follows from Lemma A.11. □

Finally we complete the proof of Theorem 2.1.

*Proof of Theorem 2.1.* 1. By the discussion at the beginning part of Section 2, we have $u=0$ is a local maximum point and the Hessian is strictly negative-definite.

2. For any $u \neq 0$ being a possible critical point and satisfying $|\hat{u}\cdot e_1| \leq \sqrt{\frac{\sqrt{3}-1}{2}} - 0.01$, Theorem 2.2 shows that

$$H_{e_1e_1}(u) < 0.$$

3. For any $0 \neq u = \sqrt{R}\hat{u}$ being a possible critical point, $\hat{u}$ must satisfy the equation (2.3). Theorem 2.3 shows that if $|\hat{u}\cdot e_1| > 10^{-4}$, then we must have

$$||\hat{u}\cdot e_1|-1| \ll 1, \quad |R-1| \ll 1.$$

4. Theorem 2.4 shows that in the neighborhood $||\hat{u}\cdot e_1|-1| \ll 1$, $|R-1| \ll 1$, it holds that the Hessian is strictly positive definite therein. Hence the loss function is strongly convex with $u = \pm e_1$ being a minimum. □

## 3. The first model

Consider for $\beta > 0$,

$$f(u) = \frac{1}{m}\sum_{j=1}^{m}\Big(\sqrt{\beta|u|^2+(a_j\cdot u)^2} - \sqrt{\beta|u|^2+(a_j\cdot x)^2}\Big)^2, \tag{3.1}$$

where for convenience we have denoted $|u| = \|u\|_2 = \sqrt{\sum_j (u_j)^2}$.

**Theorem 3.1.** *Let $0 < \beta < \infty$. Assume $\{a_i\}_{i=1}^m$ are i.i.d. standard Gaussian random vectors and $x \neq 0$. There exist positive constants $C$, $C_1$ depending only on $\beta$, such that if $m \geq Cn$, then with probability at least $1 - \frac{C_1}{m^2}$ the loss function $f = f(u)$ defined by (3.1) has no spurious local minimizers. The only global minimizer is $\pm x$, and the loss function is strongly convex in a neighborhood of $\pm x$. At the point $u = 0$ the loss function has non-vanishing directional gradient along any direction $\xi \in \mathbb{S}^{n-1}$. All other critical points are strict saddles, i.e., each saddle point has a neighborhood where the function has negative directional curvature.*

*Remark* 3.1. We shall show that most of the statements can be proved with high probability $1 - e^{-cm}$. The only part where the weaker probability $1 - O(m^{-2})$ is used comes in the analysis of the strong convexity near the global minimizer $u = \pm x$ (see e.g. Lemma C.9). This can be refined but we shall not dwell on it here.

We list below a few notation to be used in the proof.

**Notation**.

- Throughout this proof we fix $\beta > 0$ as a constant and do not study the precise dependence of other parameters on $\beta$.
- We write $u \in \mathbb{S}^{n-1}$ if $u \in \mathbb{R}^n$ and $\|u\|_2 = \sqrt{\sum_j (u_j)^2} = 1$.
- We use $\chi$ to denote the usual characteristic function. For example $\chi_A(x) = 1$ if $x \in A$ and $\chi_A(x) = 0$ if $x \notin A$.
- We denote by $\delta_1$, $\epsilon$, $\eta$, $\eta_1$ various constants whose value will be taken sufficiently small. The needed smallness will be clear from the context.
- For any quantity $X$, we shall write $X = O(Y)$ if $|X| \leq CY$ for some constant $C > 0$. We write $X \lesssim Y$ if $X \leq CY$ for some constant $C > 0$. We shall write $X \ll Y$ if $X \leq cY$ where the constant $c > 0$ will be sufficiently small.
- In our proof it is important for us to specify the precise dependence of the sampling size $m$ in terms of the dimension $n$. For this purpose we shall write $m \gtrsim n$ if $m \geq Cn$ where the constant $C$ is allowed to depend on $\beta$ and the small constants $\epsilon$, $\epsilon_i$ etc used in the argument. One can extract more explicit dependence of $C$ on the small constants and $\beta$ but for simplicity we suppress this dependence here.
- We shall say an event $A$ happens with **high probability** if

$$\mathbb{P}(A^c) \leq Ce^{-cm},$$

where $c > 0$, $C > 0$ are constants. The constants $c$ and $C$ are allowed to depend on $\beta$ and the small constants $\epsilon$, $\delta$ mentioned before.

If we denote

$$f_j(u, x) = \Big( \sqrt{\beta|u|^2 + (a_j \cdot u)^2} - \sqrt{\beta|u|^2 + (a_j \cdot x)^2} \,\Big)^2,$$

then clearly for $\lambda > 0$,

$$f_j(\lambda u, \lambda x) = \lambda^2 f_j(u, x).$$

In view of this homogeneity and the rotation invariance of the Gaussian distribution, we may assume without loss of generality that $x = e_1$ when studying the landscape of $f(u)$. Thus throughout the rest of the proof we shall assume $x = e_1$.

3.1. **The regimes $\|u\|_2 \le \frac{\sqrt{\beta}}{4(1+\beta)}$ and $\|u\|_2 \ge 3\sqrt{1+\beta}$ are fine.**

Write $u = \rho \hat{u}$ where $\hat{u} \in S^{n-1}$. Then

$$\begin{aligned}
&\Big( \sqrt{\beta |u|^2 + (a_j \cdot u)^2} - \sqrt{\beta |u|^2 + (a_j \cdot e_1)^2} \Big)^2 \\
&= \rho^2((a_j \cdot \hat{u})^2 + 2\beta) + (a_j \cdot e_1)^2 - 2\rho \sqrt{\beta + (a_j \cdot \hat{u})^2} \sqrt{\beta \rho^2 + (a_j \cdot e_1)^2}.
\end{aligned}$$

Note that

$$\begin{aligned}
\partial_\rho f = \frac{1}{m} \sum_{j=1}^m \Big( & 2\rho((a_j \cdot \hat{u})^2 + 2\beta) \\
& - 2\sqrt{\beta + (a_j \cdot \hat{u})^2} \sqrt{\beta \rho^2 + (a_j \cdot e_1)^2} - 2\sqrt{\beta + (a_j \cdot \hat{u})^2} \frac{\beta \rho^2}{\sqrt{\beta \rho^2 + (a_j \cdot e_1)^2}} \Big).
\end{aligned}$$

**Lemma 3.1** (The regime $\rho \ge 3\sqrt{1+\beta}$ is OK)**.** *For $m \gtrsim n$, with high probability it holds that*

$$\partial_\rho f > 0, \quad \forall \rho \ge 3\sqrt{1+\beta}, \ \forall \hat{u} \in \mathbb{S}^{n-1}.$$

*Proof.* By using Bernstein's inequality (Lemma A.3), we have with high probability,

$$\Big| \frac{1}{m} \sum_{j=1}^m (a_j \cdot \hat{u})^2 - 1 \Big| \le \delta_1 \ll 1, \qquad \forall \hat{u} \in \mathbb{S}^{n-1}.$$

Then

$$\frac{1}{m} \sum_{j=1}^m \Big( 2\rho((a_j \cdot \hat{u})^2 + 2\beta) \Big) \ge 2\rho((1 - \delta_1) + 2\beta).$$

On the other hand,

$$\begin{aligned}
&\frac{2}{m} \sum_{j=1}^m \Big( \sqrt{\beta + (a_j \cdot \hat{u})^2} \sqrt{\beta \rho^2 + (a_j \cdot e_1)^2} \Big) \\
&\le \frac{2}{m} \Big( \sum_{j=1}^m (\beta + (a_j \cdot \hat{u})^2) \Big)^{\frac{1}{2}} \cdot \Big( \sum_{j=1}^m (\beta \rho^2 + (a_j \cdot e_1)^2) \Big)^{\frac{1}{2}} \\
&\le 2(\beta + 1 + \delta_1)^{\frac{1}{2}} (\beta \rho^2 + 1 + \delta_1)^{\frac{1}{2}}.
\end{aligned}$$

Also clearly,

$$\begin{aligned}
&\frac{2}{m} \sum_{j=1}^m \Big( \sqrt{\beta + (a_j \cdot \hat{u})^2} \frac{\beta \rho^2}{\sqrt{\beta \rho^2 + (a_j \cdot e_1)^2}} \Big) \\
&\le \frac{2}{m} \sum_{j=1}^m \Big( \sqrt{\beta + (a_j \cdot \hat{u})^2} \sqrt{\beta \rho^2 + (a_j \cdot e_1)^2} \Big) \\
&\le 2(\beta + 1 + \delta_1)^{\frac{1}{2}} (\beta \rho^2 + 1 + \delta_1)^{\frac{1}{2}}.
\end{aligned}$$

And thus

$$\partial_\rho f \geq 2\rho((1-\delta_1)+2\beta) - 4(\beta+1+\delta_1)^{\frac{1}{2}}(\beta\rho^2+1+\delta_1)^{\frac{1}{2}}$$
$$= 2\frac{\rho^2(1-(2+8\beta)\delta_1) - 4(1+\beta) + (-8-4\beta)\delta_1 + (\rho^2-4)\delta_1^2}{\rho((1-\delta_1)+2\beta) + 2(\beta+1+\delta_1)^{\frac{1}{2}}(\beta\rho^2+1+\delta_1)^{\frac{1}{2}}}.$$

Clearly if $\delta_1 > 0$ is sufficiently small and $\rho \geq 3\sqrt{1+\beta}$, then $\partial_\rho f > 0$. □

**Lemma 3.2** (The regime $0 < \rho \leq \frac{\sqrt{\beta}}{4(1+\beta)}$ is OK)**.** *For $m \gtrsim n$, with high probability it holds that*

$$\partial_\rho f < 0, \quad \forall\, 0 < \rho \leq \frac{\sqrt{\beta}}{4(1+\beta)},\ \forall\, \hat{u} \in \mathbb{S}^{n-1}.$$

*Proof.* By Bernstein's inequality Lemma A.3, we have with high probability,

$$\Big|\frac{1}{m}\sum_{j=1}^m |a_j \cdot e_1| - \sqrt{\frac{2}{\pi}}\Big| \leq \delta_1 \ll 1,$$
$$\Big|\frac{1}{m}\sum_{j=1}^m (a_j \cdot \hat{u})^2 - 1\Big| \leq \delta_1 \ll 1, \qquad \forall\, \hat{u} \in \mathbb{S}^{n-1}.$$

Thus

$$\partial_\rho f \leq \frac{1}{m}\sum_{j=1}^m \Big(2\rho((a_j\cdot\hat{u})^2 + 2\beta)\Big) - \frac{1}{m}\sum_{j=1}^m 2\sqrt{\beta}|a_j \cdot e_1|$$
$$\leq 2\rho(1+\delta_1+2\beta) - 2\sqrt{\beta}(\sqrt{\frac{2}{\pi}} - \delta_1).$$

Since $\sqrt{\frac{2}{\pi}} \approx 0.797885$, the desired result clearly follows by taking $\delta_1$ sufficiently small. □

The point $u = 0$ needs to be treated with care since our loss function $f(u)$ is only Lipschitz at this point. To this end, we define the one-sided directional derivative of $f$ along a direction $\xi \in \mathbb{S}^{n-1}$ as

$$D_\xi f(0) = \lim_{t \to 0^+} \frac{f(t\xi)}{t}. \tag{3.2}$$

It is easy to check that

$$D_\xi f(0) = -\frac{2}{m}\sum_{j=1}^m \sqrt{\beta + (a_j \cdot \xi)^2}|a_j \cdot e_1|.$$

**Lemma 3.3** (The point $u = 0$ is OK)**.** *For $m \gtrsim n$, with high probability it holds that*

$$D_\xi f(0) < -\sqrt{\beta}, \quad \forall\, \xi \in \mathbb{S}^{n-1}.$$

*Proof.* Clearly with high probability and uniformly in $\xi \in \mathbb{S}^{n-1}$,

$$D_\xi f(0) \leq -\frac{2}{m}\sum_{j=1}^m \sqrt{\beta}|a_j \cdot e_1| < -2\sqrt{\beta}(\sqrt{\frac{2}{\pi}} - 0.01) < -\sqrt{\beta}.$$

□

**Theorem 3.2** (Non-vanishing gradient when $\|u\|_2 \leq \frac{\sqrt{\beta}}{4(1+\beta)}$ or $\|u\|_2 \geq 3(1+\beta)$)**.** *For $m \gtrsim n$, with high probability the following hold:*

*(1) We have*

$$\partial_\rho f < 0, \qquad \forall\, 0 < \rho \le \frac{\sqrt{\beta}}{4(1+\beta)}, \quad \forall\, \hat{u} \in \mathbb{S}^{n-1};$$
$$\partial_\rho f > 0, \qquad \forall\, \rho \ge 3(1+\beta), \quad \forall\, \hat{u} \in \mathbb{S}^{n-1}.$$

*(2) For $u = 0$, we have*

$$D_\xi f(0) < -\sqrt{\beta}, \qquad \forall\, \xi \in \mathbb{S}^{n-1},$$

*where $D_\xi f(0)$ was defined in (3.2).*

*Proof of Theorem 3.2.* This follows from Lemma 3.1, 3.2 and 3.3. □

### 3.2. Analysis of the regime $\rho \sim 1$, $||\hat{u} \cdot e_1| - 1| \ge \epsilon_0 > 0$.

In this section we consider the regime $0 < c_1 < \rho < c_2 < \infty$, $|\hat{u} \cdot e_1| < 1 - \epsilon_0$, where $0 < \epsilon_0 \ll 1$. The choice of the constants $c_1$ and $c_2$ can be quite flexible. For example, we can take $c_1 = \frac{\sqrt{\beta}}{4(1+\beta)}$, $c_2 = 3(1+\beta)$. For this reason we write $\rho \sim 1$.

To simplify the discussion, we need to employ a new coordinate system. Write

$$\begin{aligned}\hat{u} &= (\hat{u} \cdot e_1)e_1 + \tilde{u},\\ &= te_1 + \sqrt{1-t^2}e^{\perp},\end{aligned}$$

where $e^{\perp} \in \mathbb{S}^{n-1}$ satisfies $e^{\perp} \cdot e_1 = 0$. Clearly in the regime $\rho \sim 1$, $|t| < 1$, we have a smooth representation

$$u = \rho\hat{u} = \rho(te_1 + \sqrt{1-t^2}e^{\perp}) =: \psi(\rho, t, e^{\perp}).$$

The following pedestrian proposition shows that the landscape of a smooth function undergoes mild changes under smooth change of variables.

**Proposition 3.1** (Criteria for no local minimum)**.** *In the regime $\rho \sim 1$, $|t| < 1$, consider*

$$f(u) = f(\psi(\rho, t, e^{\perp})) =: g(\rho, t, e^{\perp}).$$

*Then the following hold:*

*(1) If at some point $|\partial_t g| > 0$, then $\|\nabla f\|_2 > 0$ at the corresponding point.*

*(2) If at some point $\partial_{tt} g < 0$, then either $\nabla f \neq 0$ at the corresponding point, or $\nabla f = 0$ and $f$ has a negative curvature at the corresponding point (i.e. a strict saddle).*

*Proof.* These easily follow from the formulae:

$$\begin{aligned}\partial_t g &= \nabla f \cdot \partial_t \psi,\\ \partial_{tt} g &= \nabla f \cdot \partial_{tt}\psi + (\partial_t \psi)^T \nabla^2 f \partial_t \psi,\end{aligned}$$

where $\nabla^2 f = (\partial_{ij} f)$ denotes the Hessian matrix of $f$. □

Proposition 3.1 allows us to simplify the computation greatly by looking only at the derivatives $\partial_t$ and $\partial_{tt}$. We shall use these in the regime $|t| < 1 - \epsilon_0$ where $0 < \epsilon_0 \ll 1$.

Now observe that

$$\frac{1}{2}\mathbb{E}f = \frac{1}{2}(1+2\beta)\rho^2 + \frac{1}{2} - \rho\mathbb{E}\sqrt{\beta + (a\cdot\hat{u})^2}\sqrt{\beta\rho^2 + (a\cdot e_1)^2},$$

where $a \sim \mathcal{N}(0, \mathrm{I_n})$.

Denote $X_1 = a \cdot e_1$ and $Y_1 = a \cdot e^{\perp}$ so that $a \cdot \hat{u} = tX_1 + \sqrt{1-t^2}Y_1 =: X_t$.

We focus on the term

$$\begin{aligned}&\mathbb{E}\sqrt{\beta + (a\cdot\hat{u})^2}\sqrt{\beta\rho^2 + (a\cdot e_1)^2}\\ &= \mathbb{E}\sqrt{\beta + X_t^2}\sqrt{\beta\rho^2 + X_1^2} =: h_\infty(\rho, t).\end{aligned}$$

**Lemma 3.4** (The limiting profile). *For any $0 < \eta_0 \ll 1$, the following hold:*

*(1) We have*

$$\sup_{|t|\le 1-\eta_0, \rho\sim 1} (|\partial_t h_\infty(\rho,t)| + |\partial_{tt} h_\infty(\rho,t)| + |\partial_{ttt} h_\infty(\rho,t)|) \lesssim 1.$$

*(2) $|\partial_t h_\infty(\rho,t)| \gtrsim |t|$ for $0 < |t| < 1$, $\rho \sim 1$.*

*(3) $\partial_{tt} h_\infty(\rho,t) \gtrsim 1$ for $|t| \ll 1$, $\rho \sim 1$.*

*Proof.* See appendix. □

**Theorem 3.3** (The regime $\|u\|_2 \sim 1$, $||\hat{u}\cdot e_1| - 1| > \eta_0$ is fine). *For any given $0 < \eta_0 \ll 1$ and $0 < c_1 < c_2 < \infty$, if $m \gtrsim n$, then the following hold with high probability:*

*In the regime $c_1 < \rho = \|u\|_2 < c_2$, $||\hat{u}\cdot e_1| - 1| > \eta_0$, there are only two possibilities:*

*(1) $\|\nabla f(u)\|_2 > 0$;*

*(2) $\nabla f(u) = 0$, and $f$ has a negative directional curvature at this point.*

*Proof of Theorem 3.3.* Denote

$$\begin{aligned}g(\rho, t, e^{\perp}) &= 2\beta\rho^2 + \rho^2\frac{1}{m}\sum_{j=1}^m (a_j\cdot\hat{u})^2 - 2\rho\frac{1}{m}\sum_{j=1}^m \sqrt{\beta + (a_j\cdot\hat{u})^2}\cdot\sqrt{\beta\rho^2 + X_j^2}\\ &=: 2\beta\rho^2 + \rho^2 h_0(\rho, t, e^{\perp}) - 2\rho h(\rho, t, e^{\perp}),\end{aligned}$$

where $X_j = a_j \cdot e_1$, and

$$a_j \cdot \hat{u} = tX_j + \sqrt{1-t^2}Y_j, \qquad Y_j = a_j\cdot e^{\perp}.$$

Clearly

$$\partial_t g = \rho^2\partial_t h_0 - 2\rho\partial_t h;$$

$$\partial_{tt} g = \rho^2\partial_{tt} h_0 - 2\rho\partial_{tt} h.$$

Observe that

$$
\begin{aligned}
h_0 &= \frac{1}{m}\sum_{j=1}^m (tX_j + \sqrt{1-t^2}Y_j)^2 \\
&= t^2\frac{1}{m}\sum_{j=1}^m X_j^2 + 2t\sqrt{1-t^2}\frac{1}{m}\sum_{j=1}^m X_jY_j + (1-t^2)\frac{1}{m}\sum_{j=1}^m Y_j^2.
\end{aligned}
$$

Clearly then for any small $\epsilon > 0$ and $m \gtrsim n$, it holds with high probability that

$$
|\partial_t h_0 - \mathbb{E}\partial_t h_0| + |\partial_{tt} h_0 - \mathbb{E}\partial_{tt} h_0| \le \epsilon, \qquad \forall\, |t| \le 1-\epsilon_0, e^\perp \cdot e_1 = 0, e^\perp \in \mathbb{S}^{n-1}.
$$

Note that we actually have $\mathbb{E}\partial_t h_0 = 0$ and $\mathbb{E}\partial_{tt} h_0 = 0$.

By Lemma C.3 and C.4, for any small $\epsilon > 0$ and $m \gtrsim n$, it also holds with high probability that

$$
\begin{aligned}
&|\partial_t h - \mathbb{E}\partial_t h| \le \epsilon, \qquad \forall\, |t| \le 1-\epsilon_0, e^\perp \cdot e_1 = 0, e^\perp \in \mathbb{S}^{n-1}, \forall\, c_1 \le \rho \le c_2; \\
&\partial_{tt} h \ge \mathbb{E}\partial_{tt} h - \epsilon, \qquad \forall\, |t| \le 1-\epsilon_0, e^\perp \cdot e_1 = 0, e^\perp \in \mathbb{S}^{n-1}, \forall\, c_1 \le \rho \le c_2.
\end{aligned}
$$

We then obtain for small $\epsilon > 0$, if $m \gtrsim n$, it holds with high probability that

$$
\begin{aligned}
&|\partial_t g - \mathbb{E}\partial_t g| \le 2\epsilon; \\
&\partial_{tt} g \le \mathbb{E}\partial_{tt} g + 2\epsilon.
\end{aligned}
$$

Clearly

$$
\begin{aligned}
&\mathbb{E}\partial_t g = -2\rho\partial_t h_\infty(\rho, t); \\
&\mathbb{E}\partial_{tt} g = -2\rho\partial_{tt} h_\infty(\rho, t).
\end{aligned}
$$

By Lemma 3.4, we can take $t_0 \ll 1$ such that

$$
\begin{aligned}
&\partial_{tt} h_\infty(\rho, t) \ge \epsilon_1 > 0, \qquad \forall\, |t| \le t_0, c_1 \le \rho \le c_2; \\
&|\partial_t h_\infty(\rho, t)| \ge \epsilon_2 > 0, \qquad \forall\, t_0 \le |t| \le 1-\epsilon_0, c_1 \le \rho \le c_2.
\end{aligned}
$$

By taking $\epsilon > 0$ sufficiently small, we can then guarantee that

$$
\begin{aligned}
&|\partial_t g| > \epsilon_3 > 0, \qquad \forall\, |t| \le t_0, c_1 \le \rho \le c_2; \\
&\partial_{tt} g \le -\epsilon_4 < 0, \qquad \forall\, t_0 \le |t| \le 1-\epsilon_0, c_1 \le \rho \le c_2.
\end{aligned}
$$

The desired result then follows from Proposition 3.1. □

### 3.3. Localization of $\rho$, the regime $||\hat{u} \cdot e_1| - 1| \ll 1$.

In this section we shall localize $\rho$ under the assumption that $||\hat{u} \cdot e_1| - 1| \ll 1$, i.e., we shall show that if $||\hat{u} \cdot e_1| - 1| \le \epsilon_0 \ll 1$, then with high probability that $|\rho - 1| \le \eta(\epsilon_0) \ll 1$.

In the lemma below we assume $\rho \ge c_1$ since by Theorem 3.2 the regime $\rho \ll 1$ is already treated.

**Lemma 3.5.** *Let $0 < \beta \le \frac{1}{4}$ and consider the regime $0 < c_1 \le \rho < 1$. If $0 < \eta_0 \ll 1$ is sufficiently small, then for $m \gtrsim n$, it holds with high probability that*

$$\partial_\rho f < 0, \qquad \forall \rho \le 1 - c(\eta_0), \forall \hat{u} \in \mathbb{S}^{n-1} \text{ with } ||\hat{u} \cdot e_1| - 1| \le \eta_0,$$

*where $c(\eta_0) \to 0$ as $\eta_0 \to 0$.*

*Remark.* In Theorem 3.4 we shall remove the constraint $0 < \beta \le \frac{1}{4}$ and prove the result for all $0 < \beta < \infty$.

*Proof.* Recall

$$\begin{aligned}\frac{1}{2}\partial_\rho f = \frac{1}{m}\sum_{j=1}^{m}\Big(&\rho((a_j \cdot \hat{u})^2 + 2\beta) \\ &- \sqrt{\beta + (a_j \cdot \hat{u})^2}\sqrt{\beta\rho^2 + (a_j \cdot e_1)^2} - \sqrt{\beta + (a_j \cdot \hat{u})^2}\frac{\beta\rho^2}{\sqrt{\beta\rho^2 + (a_j \cdot e_1)^2}}\Big).\end{aligned}$$

Without loss of generality we assume

$$\|\hat{u} - e_1\|_2 \le r \ll 1.$$

The other case $\|\hat{u} + e_1\|_2 \ll 1$ is similar and therefore omitted.

Clearly

$$\begin{aligned}&|\sqrt{\beta + (a_j \cdot \hat{u})^2} - \sqrt{\beta + (a_j \cdot e_1)^2}| \\ =&\Big|\frac{(a_j \cdot (\hat{u} - e_1))(a_j \cdot (\hat{u} + e_1))}{\sqrt{\beta + (a_j \cdot \hat{u})^2} + \sqrt{\beta + (a_j \cdot e_1)^2}}\Big| \\ \le& |a_j \cdot (\hat{u} - e_1)|.\end{aligned}$$

Thus

$$\begin{aligned}\frac{1}{2}\partial_\rho f \le \frac{1}{m}\sum_{j=1}^{m}\Big(&\rho((a_j \cdot \hat{u})^2 + 2\beta) \\ &- \sqrt{\beta + (a_j \cdot e_1)^2}\sqrt{\beta\rho^2 + (a_j \cdot e_1)^2} - \sqrt{\beta + (a_j \cdot e_1)^2} \cdot \frac{\beta\rho^2}{\sqrt{\beta\rho^2 + (a_j \cdot e_1)^2}}\Big) + H,\end{aligned}$$

where

$$\begin{aligned}H &\lesssim \frac{1}{m}\sum_{j=1}^{m}|a_j \cdot (\hat{u} - e_1)|(1 + |a_j \cdot e_1|) \\ &\lesssim \frac{1}{m}\sum_{j=1}^{m}\Big(a_j \cdot (\hat{u} - e_1))^2 \cdot \frac{1}{2r} + r(1 + (a_j \cdot e_1)^2)\Big).\end{aligned}$$

Clearly it holds with high probability that

$$H \le B_1 r,$$

where $B_1 > 0$ is a constant.

For $\rho \le 1$, we have

$$
\begin{aligned}
-\sqrt{\beta+(a_j\cdot e_1)^2}\sqrt{\beta\rho^2+(a_j\cdot e_1)^2} &\le -(\beta\rho^2+(a_j\cdot e_1)^2);\\
-\sqrt{\beta+(a_j\cdot e_1)^2}\cdot\frac{\beta\rho^2}{\sqrt{\beta\rho^2+(a_j\cdot e_1)^2}} &\le -\beta\rho^2.
\end{aligned}
$$

Then assuming $\rho < 1$, it holds with high probability that

$$
\frac{1}{2}\partial_\rho f \le \rho(1+2\beta) - (\beta\rho^2+1) - \beta\rho^2 + B_1 r + \delta_1,
$$

where $\delta_1 > 0$ is a small constant which accounts for the deviation from the mean value used in the Bernstein's inequality. For $0 < \beta \le \frac{1}{4}$ (actually $0 < \beta < \frac{1}{2}$ suffices) the desired conclusion then clearly follows by taking $\delta_1 = O(\eta_0)$ and $r = O(\eta_0)$. $\square$

**Lemma 3.6** ($\partial_{\rho\rho} f$ is good)**.** *We have almost surely it holds that*

$$
\partial_{\rho\rho} f > 0, \qquad \forall\, 0<\rho<\infty,\ \forall\, \hat{u}\in\mathbb{S}^{n-1}.
$$

*Furthermore, for any fixed two constants $0 < c_1 < c_2 < \infty$, if $m \gtrsim n$, then it holds with high probability that*

$$
\partial_{\rho\rho} f \ge \alpha > 0, \qquad \forall\, c_1 \le \rho \le c_2,\ \forall\, \hat{u}\in\mathbb{S}^{n-1},
$$

*where $\alpha > 0$ is a constant depending only on $(c_1, c_2, \beta)$.*

*Proof.* Recall

$$
\begin{aligned}
\frac{1}{2}\partial_\rho f = \frac{1}{m}\sum_{j=1}^m \Big(&\rho((a_j\cdot\hat{u})^2+2\beta)\\
&-\sqrt{\beta+(a_j\cdot\hat{u})^2}\sqrt{\beta\rho^2+(a_j\cdot e_1)^2} - \sqrt{\beta+(a_j\cdot\hat{u})^2}\frac{\beta\rho^2}{\sqrt{\beta\rho^2+(a_j\cdot e_1)^2}}\Big).
\end{aligned}
$$

Then

$$
\begin{aligned}
\frac{1}{2}\partial_{\rho\rho} f &= \frac{1}{m}\sum_{j=1}^m \Big(((a_j\cdot\hat{u})^2+2\beta)\\
&\qquad -\sqrt{\beta+(a_j\cdot\hat{u})^2}\frac{3\beta\rho}{\sqrt{\beta\rho^2+(a_j\cdot e_1)^2}} + \sqrt{\beta+(a_j\cdot\hat{u})^2}\frac{\beta^2\rho^3}{(\beta\rho^2+(a_j\cdot e_1)^2)^{\frac{3}{2}}}\Big)\\
&= \frac{1}{m}\sum_{j=1}^m \Bigg(((a_j\cdot\hat{u})^2+2\beta)\\
&\qquad -\sqrt{\beta+(a_j\cdot\hat{u})^2}\cdot\sqrt{\beta}\cdot\left(3\left(\frac{a_j\cdot e_1}{\sqrt{\beta}\rho}\right)^2+2\right)\cdot\left(\left(\frac{a_j\cdot e_1}{\sqrt{\beta}\rho}\right)^2+1\right)^{-\frac{3}{2}}\Bigg).
\end{aligned}
$$

For $0 \le x < \infty$, denote

$$
h_0(x) = \frac{3x+2}{(1+x)^{\frac{3}{2}}}.
$$

It is not difficult to check that

$$
h_0(x) \le 2, \qquad \forall\, 0 \le x < \infty
$$

and the equality holds if and only if $x = 0$.

Now define $h_1(x) = h_0(x^2)$. Then clearly

$$\frac{1}{2}\partial_{\rho\rho} f = \frac{1}{m}\sum_{j=1}^m (\sqrt{\beta + (a_j \cdot \hat{u})^2} - \sqrt{\beta})^2 + \frac{1}{m}\sum_{j=1}^m \sqrt{\beta + (a_j \cdot \hat{u})^2} \cdot \sqrt{\beta} \cdot \Big(2 - h_1(\frac{a_j \cdot e_1}{\sqrt{\beta}\rho})\Big).$$

Obviously then

$$\frac{1}{2}\partial_{\rho\rho} f \geq \beta \frac{1}{m}\sum_{j=1}^m \Big(2 - h_1(\frac{a_j \cdot e_1}{\sqrt{\beta}\rho})\Big) > 0,$$

which holds almost surely since the event $\bigcap_{j=1}^m \{a_j \cdot e_1 = 0\}$ has zero probability.

By using the Bernstein's inequality, we have with high probability that

$$\frac{1}{m}\sum_{j=1}^m \Big(2 - h_1(\frac{a_j \cdot e_1}{\sqrt{\beta}\rho})\Big) \gtrsim 1, \qquad \forall\, c_1 \leq \rho \leq c_2.$$

Thus

$$\partial_{\rho\rho} f \gtrsim 1, \qquad \forall\, c_1 \leq \rho \leq c_2, \quad \forall\, \hat{u} \in \mathbb{S}^{n-1}.$$

□

**Theorem 3.4** (Localization of $\rho$ when $||\hat{u} \cdot e_1| - 1| \ll 1$)**.** *Consider the regime $0 < c_1 \leq \rho \leq c_2$. If $0 < \eta_0 \ll 1$ is sufficiently small, then for $m \gtrsim n$, it holds with high probability that*

$$\partial_\rho f < 0, \qquad \forall\, \rho \leq 1 - c(\eta_0), \forall\, \hat{u} \in \mathbb{S}^{n-1} \text{ with } ||\hat{u} \cdot e_1| - 1| \leq \eta_0;$$
$$\partial_\rho f > 0, \qquad \forall\, \rho \geq 1 + c(\eta_0), \forall\, \hat{u} \in \mathbb{S}^{n-1} \text{ with } ||\hat{u} \cdot e_1| - 1| \leq \eta_0;$$

*where $c(\eta_0) \to 0$ as $\eta_0 \to 0$.*

*Proof.* We shall sketch the proof.

We first consider the regime $\rho \geq 1$.

Without loss of generality we assume $\|\hat{u} - e_1\|_2 \leq \eta_0$. The other case $\|\hat{u} + e_1\|_2 \leq \eta_0$ can be similarly treated.

First observe that

$$(\partial_\rho f)(\rho = 1, \hat{u} = e_1) = 0.$$

Then by a calculation similar to the estimate of $H$ term in Lemma 3.5, we have with high probability that

$$|\partial_\rho h(\rho = 1, \hat{u})| = c_1(\eta_0) \ll 1, \qquad \forall\, |\hat{u} - e_1| \leq \eta_0,$$

where $c_1(\eta_0) \to 0$ as $\eta_0 \to 0$.

Now by Lemma 3.6, it holds with high probability that

$$\partial_{\rho\rho} f \geq \alpha > 0, \quad \forall\, c_1 \leq \rho \leq c_2, \forall\, \hat{u} \in \mathbb{S}^{n-1}.$$

It then implies that for $\rho \geq 1 + \frac{2c_1(\eta_0)}{\alpha_0}$, we have

$$(\partial_\rho f)(\rho, \hat{u}) \geq \alpha_0 \cdot \frac{2c_1(\eta_0)}{\alpha_0} - c_1(\eta_0) = c_1(\eta_0) > 0.$$

Redefining $c(\eta_0)$ suitably then yields the result. The argument for $\rho \le 1 - c(\eta_0)$ is similar. We omit the details. $\square$

### 3.4. **Strong convexity near the global minimizers $u = \pm e_1$: analysis of the limiting profile.**

In this section we shall show that in the small neighborhood of $u = \pm e_1$ where

$$||\hat{u} \cdot e_1| - 1| \ll 1, \quad |\rho - 1| \ll 1,$$

the Hessian of the expectation of the loss function must be strictly positive definite. In yet other words $\mathbb{E}f$ must be strictly convex in this neighborhood so that $u = \pm e_1$ are the unique minimizers. To this end consider

$$h(u) = \frac{1}{2}(\mathbb{E}f - 1) = \frac{1}{2}(1 + 2\beta)|u|^2 - \mathbb{E}\sqrt{\beta|u|^2 + (a \cdot u)^2}\sqrt{\beta|u|^2 + (a \cdot e_1)^2}, \tag{3.3}$$

where $a \sim \mathcal{N}(0, \mathrm{I}_n)$.

**Theorem 3.5** (Strong convexity of $\mathbb{E}f$ when $\|u \pm e_1\| \ll 1$)**.** *Consider $h$ defined by (3.3). There exists $0 < \epsilon_0 \ll 1$ such that the following hold:*

*(1) If $\|u - e_1\|_2 \le \epsilon_0$, then for any $\xi \in \mathbb{S}^{n-1}$, we have*

$$\sum_{i,j=1}^{n} \xi_i \xi_j \mathbb{E}(\partial_i \partial_j h)(u) \ge \gamma_1 > 0,$$

*where $\gamma_1$ is a constant.*

*(2) If $\|u + e_1\|_2 \le \epsilon_0$, then for any $\xi \in \mathbb{S}^{n-1}$, we have*

$$\sum_{i,j=1}^{n} \xi_i \xi_j \mathbb{E}(\partial_i \partial_j h)(u) \ge \gamma_1 > 0.$$

*Proof of Theorem 3.5.* We shall only consider the case $\|u - e_1\|_2 \ll 1$. The other case $\|u + e_1\|_2 \ll 1$ is similar and therefore omitted. Note that

$$\|u - e_1\|_2^2 = \|\rho\hat{u} - e_1\|_2^2 = (\rho - 1)^2 + 2\rho(1 - t) \le \epsilon_0^2,$$

where $t = \hat{u} \cdot e_1$. Thus for $0 < \epsilon_0 \ll 1$, we have

$$|\rho - 1| \le \epsilon_0, \qquad 1 - \epsilon_0^2 \le t \le 1.$$

We now need to make a change of variable. The representation $u = te_1 + \sqrt{1 - t^2}e^{\perp}$ is not so suitable since the derivatives blow up as $t \to 1-$. This is an artificial singularity due to the non-smoothness of the representation $\sqrt{1 - t^2}$ as $t \to 1-$. To resolve this, we use a different representation (recall $1 - \epsilon_0^2 \le \hat{u} \cdot e_1 \to 1$),

$$\hat{u} = \sqrt{1 - s^2}e_1 + se^{\perp}, \qquad e^{\perp} \cdot e_1 = 0,\ e^{\perp} \in \mathbb{S}^{n-1},$$

where we assume $0 \le s \ll 1$. Note that $s = \frac{|u'|}{\rho}$, and $u' = u - (u \cdot e_1)e_1 = (0, u_2, \cdots, u_n)$.

*Remark.* This "switch of representation" reminds us that on a sphere we need to use at least two charts! (One near the north pole, and another near the south pole.)

To calculate $\partial^2 h$ we need to compute the Hessian expressed in the $(\rho, s)$ coordinate. It is not difficult to check that by (3.3), the value of $h(u)$ depends only on $(\rho, s)$. Thus by a slight abuse of notation we write $h = h(|u|, \frac{|u'|}{|u|}) = h(\rho, s)$ (we denote $|u| = \|u\|_2$, $|u'| = \|u'\|_2$) and compute (below we assume $s > 0$ so that $|u'| > 0$)

$$
\begin{aligned}
&\partial_i h = \partial_\rho h \frac{u_i}{\rho} + \partial_s h \cdot (-\frac{u_i}{\rho^3}|u'| + 1_{i\neq 1}\frac{1}{\rho} \cdot \frac{u_i}{|u'|});\\
&\partial_{ij} h = \partial_{\rho\rho} h \frac{u_i u_j}{\rho^2} + \frac{u_i}{\rho}\partial_{\rho s} h \cdot (-\frac{u_j}{\rho^3}|u'| + 1_{j\neq 1}\frac{1}{\rho} \cdot \frac{u_j}{|u'|})\\
&\quad + \partial_\rho h \cdot (\frac{\delta_{ij}}{\rho} - \frac{u_i u_j}{\rho^3})\\
&\quad + \frac{u_j}{\rho}\partial_{\rho s} h \cdot (-\frac{u_i}{\rho^3}|u'| + 1_{i\neq 1}\frac{1}{\rho} \cdot \frac{u_i}{|u'|})\\
&\quad + \partial_{ss} h \cdot (-\frac{u_i}{\rho^3}|u'| + 1_{i\neq 1}\frac{1}{\rho} \cdot \frac{u_i}{|u'|}) \cdot (-\frac{u_j}{\rho^3}|u'| + 1_{j\neq 1}\frac{1}{\rho} \cdot \frac{u_j}{|u'|})\\
&\quad + \partial_s h \cdot (-\frac{\delta_{ij}}{\rho^3}|u'| + \frac{3u_i u_j}{\rho^5}|u'| - \frac{u_i}{\rho^3}\frac{u_j}{|u'|}1_{j\neq 1} - \frac{u_j}{\rho^3}\frac{u_i}{|u'|}1_{i\neq 1} + 1_{i\neq 1}\frac{1}{\rho}\frac{\delta_{ij}}{|u'|} - 1_{i\neq 1}1_{j\neq 1}\frac{1}{\rho} \cdot \frac{u_i u_j}{|u'|^3}).
\end{aligned}
$$

Then denoting $a = \xi \cdot \hat{u}$, $b = \sum_{j\neq 1} \xi_j \cdot \frac{u_j}{|u'|}$, we have

$$
\begin{aligned}
\sum_{i,j=1}^n \xi_i \xi_j \partial_{ij} h &= \partial_{\rho\rho} h \cdot a^2 + 2a\partial_{\rho s} h \cdot (-\frac{as}{\rho} + \frac{b}{\rho})\\
&\quad + \partial_\rho h \cdot (\frac{|\xi|^2 - |\xi \cdot \hat{u}|^2}{\rho})\\
&\quad + \partial_{ss} h \cdot (-\frac{as}{\rho} + \frac{b}{\rho})^2\\
&\quad + \partial_s h \cdot (-\frac{|\xi|^2}{\rho^2}s + 3\frac{a^2 s}{\rho^2} - 2\frac{ab}{\rho^2} + \frac{|\xi'|^2}{\rho|u'|} - \frac{b^2}{\rho|u'|})\\
&= \partial_{\rho\rho} h \cdot a^2 + 2a\partial_{\rho s} h \cdot (-\frac{as}{\rho} + \frac{b}{\rho})\\
&\quad + \partial_\rho h \cdot (\frac{|\xi|^2 - |\xi \cdot \hat{u}|^2}{\rho})\\
&\quad + \partial_{ss} h \cdot (\frac{a^2 s^2 - 2abs}{\rho^2}) + (\partial_{ss} h - \frac{1}{s}\partial_s h)\frac{b^2}{\rho^2}\\
&\quad + \partial_s h \cdot (-\frac{|\xi|^2}{\rho^2}s + 3\frac{a^2 s}{\rho^2} - 2\frac{ab}{\rho^2} + \frac{|\xi'|^2}{\rho^2 s})
\end{aligned}
$$

We should point it out that, in the above computation, one does not need to worry about the formal singularity caused by $\frac{1}{s}$. Since $\partial_s h(\rho, s = 0) = 0$, we write

$$
(\partial_s h)(\rho, s) \cdot \frac{1}{s} = \frac{(\partial_s h)(\rho, s) - (\partial_s h)(\rho, 0)}{s} = \int_0^1 (\partial_{ss} h)(\rho, \theta s) d\theta, \quad s > 0.
$$

In particular we have

$$
\begin{aligned}
&\lim_{s\to 0^+} (\partial_s h)(\rho, s) \cdot \frac{1}{s} = (\partial_{ss} h)(\rho, 0);\\
&\Big|(\partial_{ss} h)(\rho, s) - \frac{1}{s}\partial_s h_1(\rho, s)\Big| = O(s) \to 0, \quad \text{as } s \to 0.
\end{aligned}
$$

By using this observation and Lemma C.6, we obtain

$$\sum_{i,j=1}^{n} \xi_i \xi_j (\partial_{ij} h)(e_1) = (\partial_{\rho\rho} h)(1,0) \cdot a^2 \Big|_{a=\xi_1} + (\partial_{ss} h)(1,0) \cdot |\xi'|^2$$

$$\geq \gamma_0 \cdot |\xi|^2, \qquad \forall\, \xi \in \mathbb{S}^{n-1},$$

where $\gamma_0 > 0$ is a constant. Now for $\|u - e_1\|_2 \ll 1$, by using Lemma C.5 and Lemma C.6, we have

$$\begin{aligned}
&\Big| \sum_{i,j=1}^{n} \xi_i \xi_j \Big( (\partial_{ij} h)(e_1) - (\partial_{ij} h)(u) \Big) \Big| \\
&\lesssim |(\partial_{\rho\rho} h)(\rho, s) - (\partial_{\rho\rho} h)(1,0)| + |(\partial_{\rho\rho} h)(\rho, s)| \cdot |(\xi \cdot \hat{u})^2 - (\xi \cdot e_1)^2| \\
&\quad + |(\partial_{\rho s} h)(\rho, s)| \cdot (1 + s) + |(\partial_{\rho} h)(\rho, s)| + |\partial_{ss} h(\rho, s)| \cdot (s + s^2) \\
&\quad + |\partial_s h(\rho, s)| + \Big| \frac{1}{\rho^2} \int_0^1 (\partial_{ss} h)(\rho, \theta s) d\theta - (\partial_{ss} h)(1,0) \Big| \\
&\quad + \Big| (\partial_{ss} h)(\rho, s) - (\partial_s h)(\rho, s) \cdot \frac{1}{s} \Big| \cdot \frac{b^2}{\rho^2} \\
&\lesssim O(|\rho - 1| + |s| + \|\hat{u} - e_1\|_2).
\end{aligned}$$

It follows that if $\|u - e_1\|_2$ is sufficiently small, we then have

$$\sum_{i,j=1}^{n} \xi_i \xi_j (\partial_{ij} h)(u) \geq \frac{\gamma_0}{2} |\xi|^2, \qquad \forall\, \xi \in \mathbb{S}^{n-1}.$$

□

### 3.5. Near the global minimizer: strong convexity.

In this section we show strong convexity of the loss function $f(u)$ near the global minimizer $u = \pm e_1$.

**Theorem 3.6** (Strong convexity near the global minimizer). *There exists $0 < \epsilon_0 \ll 1$ and a constant $\beta_1 > 0$ such that if $m \gtrsim n$, then the following hold with probability at least $1 - \frac{\beta_1}{m^2}$:*

*(1) If $\|u - e_1\|_2 \leq \epsilon_0$, then*

$$\sum_{i,j=1}^{n} \xi_i \xi_j (\partial_{ij} f)(u) \geq \gamma > 0, \qquad \forall\, \xi \in \mathbb{S}^{n-1},$$

*where $\gamma$ is a constant.*

*(2) If $\|u + e_1\|_2 \leq \epsilon_0$, then*

$$\sum_{i,j=1}^{n} \xi_i \xi_j (\partial_{ij} f)(u) \geq \gamma > 0, \qquad \forall\, \xi \in \mathbb{S}^{n-1},$$

*where $\gamma$ is a constant.*

*In yet other words, $f(u)$ is strongly convex in a sufficiently small neighborhood of $\pm e_1$.*

*Proof of Theorem 3.6.* Recall

$$f(u) = 2f_0(u) + \frac{1}{m}\sum_{k=1}^{m}((a_k \cdot u)^2 + 2\beta|u|^2),$$

where

$$f_0(u) = -\frac{1}{m}\sum_{k=1}^{m}\sqrt{\beta|u|^2 + (a_k \cdot u)^2} \cdot \sqrt{\beta|u|^2 + (a_k \cdot e_1)^2}.$$

Clearly

$$\sum_{i,j=1}^{n} \xi_i\xi_j \partial_{ij} f(u) = 2\Big(\frac{1}{m}\sum_{k=1}^{m}|a_k \cdot \xi|^2\Big) + 4\beta|\xi|^2 + 2\sum_{i,j=1}^{n}\xi_i\xi_j\partial_{ij}f_0(u).$$

Obviously we have for $m \gtrsim n$, it holds with high probability that

$$\Big|\frac{1}{m}\sum_{k=1}^{m}|a_k \cdot \xi|^2 - 1\Big| \le \frac{\epsilon}{100}, \qquad \forall\, \xi \in \mathbb{S}^{n-1}.$$

By Lemma C.9, we have

$$\Big|\sum_{i,j=1}^{n}\xi_i\xi_j(\partial_{ij}f_0(u) - \mathbb{E}\partial_{ij}f_0(u))\Big| \le \frac{\epsilon}{100}, \qquad \forall\, \xi \in \mathbb{S}^{n-1},\ \forall\, \frac{1}{3} \le \|u\|_2 \le 3.$$

Thus we have

$$\Big|\sum_{i,j=1}^{n}\xi_i\xi_j(\partial_{ij}f(u) - \mathbb{E}\partial_{ij}f(u))\Big| \le \frac{\epsilon}{100}, \qquad \forall\, \xi \in \mathbb{S}^{n-1},\ \forall\, \frac{1}{3} \le \|u\|_2 \le 3.$$

The desired result then follows from Theorem 3.5 by taking $\epsilon > 0$ sufficiently small. ◻

We now complete the proof of the main theorem.

*Proof of Theorem 3.1.* We proceed in several steps.

(1) By Theorem 3.2, we see that with high probability the function $f(u)$ has non-vanishing gradient in the regimes

$$0 < \|u\|_2 \le \frac{\sqrt{\beta}}{4(1+\beta)} = c_1$$

and

$$\|u\|_2 \ge 3(1+\beta) = c_2.$$

Moreover at the point $u = 0$, we have the directional gradient is strictly less than $-\sqrt{\beta}$ along any direction $\xi \in \mathbb{S}^{n-1}$.

(2) By Theorem 3.6, there exists $\epsilon_0 > 0$ sufficiently small, such that with probability at least $1 - O(m^{-2})$, $f(u)$ is strongly convex in the neighborhood $\|u \pm e_1\|_2 \le \epsilon_0$.

(3) By Theorem 3.4, we have that with high probability

$$\|\nabla f\|_2 > 0,$$

if $|\rho - 1| \ge c(\eta_0)$ and $||\hat{u} \cdot e_1| - 1| \le \eta_0$. Here we recall $\rho = \|u\|_2$ and $u = \rho\hat{u}$. Observe that

$$\|u \pm e_1\|_2^2 = (\rho - 1)^2 + 2\rho(1 \pm \hat{u} \cdot e_1).$$

By taking $\eta_0 = \epsilon_0^2/100$, we see that $\|u \pm e_1\|_2 > \epsilon_0$, $||\hat{u} \cdot e_1| - 1| \le \eta_0$ must imply

$$|\rho - 1| > \frac{\epsilon_0}{10}.$$

Thus it remains for us to treat the regime $\||\hat{u} \cdot e_1| - 1| > \eta_0$, $c_1 \le \|u\|_2 \le c_2$.

(4) In the regime $||\hat{u} \cdot e_1| - 1| > \eta_0$, $\|u\|_2 \sim 1$, we have by Theorem 3.3, with high probability it holds that either the function has a non-vanishing gradient at the point $u$, or the gradient vanishes at $u$, but $f$ has a negative directional curvature at this point.

$\square$

## 4. The second model

Consider for $\beta > 0$,

$$f(u) = \frac{1}{m} \sum_{j=1}^{m} \Big( \sqrt{\beta|u|^2 + (a_j \cdot u)^2 + (a_j \cdot x)^2} - \sqrt{\beta|u|^2 + 2(a_j \cdot x)^2} \ \Big)^2. \tag{4.1}$$

We shall adopt similar notation as in Section 3.

**Theorem 4.1.** *Let $0 < \beta < \infty$. Assume $\{a_i\}_{i=1}^m$ are i.i.d. standard Gaussian random vectors and $x \neq 0$. There exist positive constants $c$, $C$ depending only on $\beta$, such that if $m \ge Cn$, then with probability at least $1 - e^{-cm}$ the loss function $f = f(u)$ defined by (4.1) has no spurious local minimizers. The only global minimizer is $\pm x$, and the loss function is strongly convex in a neighborhood of $\pm x$. The point $u = 0$ is a local maximum point with strictly negative-definite Hessian. All other critical points are strict saddles, i.e., each saddle point has a neighborhood where the function has negative directional curvature.*

*Remark* 4.1. One should note that the set $\bigcup_{j=1}^m \{a_j \cdot x = 0\}$ has measure zero. Therefore for a typical realization, $a_j \cdot x$ is always non-zero for all $j$ and the function

$$\tilde{f}_j(y) = \sqrt{y^2 + (a_j \cdot x)^2}$$

is smooth. In particular, we can compute (for each realization) the derivatives of the summands in (4.1) without any problem.

*Remark* 4.2. Thanks to the regularization term $(a_j \cdot x)^2$, our new model (4.1) enjoys a better probability concentration bound $1 - e^{-cm}$ than the model (3.1) where the weaker probability concentration $1 - O(m^{-2})$ is proved.

Wthout loss of generality we shall assume $x = e_1$ throughout the rest of the proof.

### 4.1. The regimes $\|u\|_2 \ll 1$ and $\|u\|_2 \gg 1$ are fine.

Write $u = \rho\hat{u}$ where $\hat{u} \in S^{n-1}$. Then

$$\begin{aligned} &\Big( \sqrt{\beta|u|^2 + (a_j \cdot u)^2 + |a_j \cdot e_1|^2} - \sqrt{\beta|u|^2 + 2(a_j \cdot e_1)^2} \ \Big)^2 \\ = \ & \rho^2((a_j \cdot \hat{u})^2 + 2\beta) + 3(a_j \cdot e_1)^2 - 2\sqrt{\beta\rho^2 + \rho^2(a_j \cdot \hat{u})^2 + (a_j \cdot e_1)^2}\sqrt{\beta\rho^2 + 2(a_j \cdot e_1)^2}. \end{aligned}$$

Note that

$$\partial_\rho f = 2\rho \frac{1}{m}\sum_{j=1}^m \Big( ((a_j\cdot\hat{u})^2 + 2\beta) \\ - \frac{\beta + (a_j\cdot\hat{u})^2}{\sqrt{\beta\rho^2+\rho^2(a_j\cdot\hat{u})^2+(a_j\cdot e_1)^2}}\sqrt{\beta\rho^2+2(a_j\cdot e_1)^2} \\ - \sqrt{\beta\rho^2+\rho^2(a_j\cdot\hat{u})^2+(a_j\cdot e_1)^2}\frac{\beta}{\sqrt{\beta\rho^2+2(a_j\cdot e_1)^2}}\Big).$$

**Lemma 4.1** (The regime $\rho \gg 1$ is OK)**.** *There exist constants $R_1 = R_1(\beta) > 0$, $d_1 = d_1(\beta) > 0$ such that the following hold:*

*For $m \gtrsim n$, with high probability it holds that*

$$\partial_\rho f \ge d_1\rho, \quad \forall\, \rho \ge R_1,\ \forall\, \hat{u} \in \mathbb{S}^{n-1}.$$

*Proof.* We only sketch the proof. Denote $X_j = a_j \cdot e_1$ and $Z_j = a_j\cdot\hat{u}$. Observe that

$$\begin{aligned} &\frac{\beta\rho^2+2X_j^2}{(\beta+Z_j^2)\rho^2+X_j^2} - \frac{\beta}{\beta+Z_j^2} \\ =& \frac{\beta X_j^2 + 2X_j^2 Z_j^2}{\Big((\beta+Z_j^2)\rho^2+X_j^2\Big)\cdot(\beta+Z_j^2)} \\ \le& \frac{2X_j^2}{(\beta+Z_j^2)\rho^2+X_j^2} \le \frac{1}{\rho^2}\cdot\frac{2X_j^2}{\beta+Z_j^2}. \end{aligned}$$

Thus

$$\begin{aligned} \frac{\beta+Z_j^2}{\sqrt{\rho^2(\beta+Z_j^2)+X_j^2}}\sqrt{\beta\rho^2+2X_j^2} &\le (\beta+Z_j^2)\Big(\sqrt{\frac{\beta}{\beta+Z_j^2}} + \frac{1}{\rho}\frac{\sqrt{2}|X_j|}{\sqrt{\beta+Z_j^2}}\Big) \\ &\le \sqrt{\beta}\sqrt{\beta+Z_j^2} + \frac{1}{\rho}\sqrt{2}|X_j|\sqrt{\beta+Z_j^2}. \end{aligned}$$

On the other hand,

$$\sqrt{\rho^2(\beta+Z_j^2)+X_j^2}\cdot\frac{\beta}{\sqrt{\beta\rho^2+2X_j^2}} \le \sqrt{\beta}\cdot\sqrt{\beta+Z_j^2}.$$

Thus

$$\begin{aligned} \frac{1}{2\rho}\partial_\rho f &\ge \frac{1}{m}\sum_{j=1}^m \Big(Z_j^2 + 2\beta - 2\sqrt{\beta}\sqrt{\beta+Z_j^2}\Big) - \frac{1}{\rho}\sqrt{2}\frac{1}{m}\sum_{j=1}^m |X_j|\cdot\sqrt{\beta+Z_j^2} \\ &\ge \frac{1}{m}\sum_{j=1}^m (\sqrt{Z_j^2+\beta}-\sqrt{\beta})^2 - \frac{1}{\rho}\cdot\frac{1}{m}\sum_{j=1}^m (X_j^2+Z_j^2+\beta). \end{aligned}$$

By Bernstein's inequality (and simple union bound arguments, cf. Lemma A.12), we clearly have with high probability,

$$\begin{aligned} &\Big|\frac{1}{m}\sum_{j=1}^m (\sqrt{Z_j^2+\beta}-\sqrt{\beta})^2 - \text{mean}\Big| \le \epsilon, \quad \forall\, \hat{u}\in\mathbb{S}^{n-1}; \\ &\frac{1}{m}\sum_{j=1}^m (X_j^2+Z_j^2+\beta) \le 3+\beta, \qquad \forall\, \hat{u}\in\mathbb{S}^{n-1}. \end{aligned}$$

The desired result then clearly follows. $\square$

**Lemma 4.2** (The regime $\|u\|_2 \ll 1$ is OK)**.** *There exist constants $R_2 = R_2(\beta) > 0$, $d_2 = d_2(\beta) > 0$ such that the following hold:*

*For $m \gtrsim n$, with high probability it holds that*

$$\partial_\rho f \le -d_2\rho < 0, \quad \forall\, 0 < \rho \le R_2,\ \forall\, \hat{u} \in \mathbb{S}^{n-1}.$$

*Moreover, at $u = 0$, we have $\nabla f(0) = 0$, and*

$$\sum_{k,l=1}^{n} \xi_k \xi_l (\partial_{kl} f)(0) \le -d_2 \|\xi\|_2^2, \qquad \forall\, \xi \in \mathbb{S}^{n-1}.$$

*In yet other words, $u = 0$ is a strict local maximum point with strictly negative definite Hessian.*

*Proof.* We only sketch the proof. Again denote $X_j = a_j \cdot e_1$ and $Z_j = a_j \cdot \hat{u}$. Observe that

$$\frac{\beta\rho^2 + 2X_j^2}{\rho^2(\beta + Z_j^2) + X_j^2} = 2 - \frac{\rho^2(\beta + 2Z_j^2)}{\rho^2(\beta + Z_j^2) + X_j^2}.$$

and

$$\begin{aligned}&(\beta + Z_j^2)\sqrt{\frac{\beta\rho^2 + 2X_j^2}{\rho^2(\beta + Z_j^2) + X_j^2}} \\ &\ge \sqrt{2}(\beta + Z_j^2) - (\beta + Z_j^2)\sqrt{\frac{\rho^2(\beta + 2Z_j^2)}{\rho^2(\beta + Z_j^2) + X_j^2}}.\end{aligned}$$

On the other hand,

$$\frac{\rho^2(\beta + Z_j^2) + X_j^2}{\beta\rho^2 + 2X_j^2} - \frac{1}{2} = \frac{1}{2} \cdot \frac{\rho^2(\beta + 2Z_j^2)}{\beta\rho^2 + 2X_j^2} \ge 0.$$

Thus

$$\begin{aligned}\frac{1}{2\rho}\partial_\rho f \le{}& \frac{1}{m}\sum_{j=1}^{m}\Big(Z_j^2 + 2\beta - \sqrt{2}\cdot(\beta + Z_j^2) - \frac{1}{\sqrt{2}}\beta\Big) \\ &+ \frac{1}{m}\sum_{j=1}^{m}(\beta + Z_j^2)\sqrt{\frac{\rho^2(\beta + 2Z_j^2)}{\rho^2(\beta + Z_j^2) + X_j^2}}.\end{aligned}$$

Since

$$Z_j^2 + 2\beta - \sqrt{2}\cdot(\beta + Z_j^2) - \frac{1}{\sqrt{2}}\beta = -(\sqrt{2} - 1)Z_j^2 - (\sqrt{2} + \frac{1}{\sqrt{2}} - 2)\beta,$$

the first summand clearly gives a nontrivial negative lower bound. The desired result then follows from Lemma D.1. We note that the result for $u = 0$ follows by taking $u = t\xi$ and re-run the above argument taking $t \to 0+$. $\square$

**Theorem 4.2** (The regimes $\|u\|_2 \ll 1$ and $\|u\|_2 \gg 1$ are OK)**.** *For $m \gtrsim n$, with high probability the following hold:*

*(1) We have*

$$\begin{aligned}&\partial_\rho f \ge d_1\rho > 0, \qquad \forall\, \rho \ge R_1, \quad \forall\, \hat{u} \in \mathbb{S}^{n-1};\\ &\partial_\rho f \le -d_2\rho < 0, \qquad \forall\, 0 < \rho \le R_2, \quad \forall\, \hat{u} \in \mathbb{S}^{n-1},\end{aligned}$$

*where $d_1$, $d_2$, $R_1$, $R_2$ are constants depending only on $\beta$.*

*(2) The point $u=0$ is a local maximum point with strictly negative-definite Hessian,*

$$\sum_{k,l=1}^{n} \xi_k \xi_l (\partial_{kl} f)(0) \le -d_2 < 0, \qquad \forall\, \xi \in \mathbb{S}^{n-1}.$$

*Proof of Theorem 4.2.* This follows from Lemma 4.1 and 4.2. $\square$

**Theorem 4.3** (The regime $\|u\|_2 \sim 1$, $||\hat{u}\cdot e_1| - 1| \le \epsilon_0$, $|\|u\|_2 - 1| \ge c(\epsilon_0)$ is OK)**.** *Let $R_1$, $R_2$ be the same as in Lemma 4.1 and 4.2. Let $0 < \epsilon_0 \ll 1$ be given and consider the regime $\Big||\hat{u}\cdot e_1| - 1\Big| \le \epsilon_0$ with $R_1 \le \|u\|_2 \le R_2$. There exists a constant $c(\epsilon_0) > 0$ ($c(\epsilon_0)$ also depends on $\beta$ but we suppress this dependence) which tends to zero as $\epsilon_0 \to 0$ such that the following hold:*

*For $m \gtrsim n$, with high probability it holds that*

$$\partial_\rho f < 0, \quad \forall\, R_2 \le \rho \le 1 - c(\epsilon_0), \ \ \forall\, \hat{u} \in \mathbb{S}^{n-1} \text{with } ||\hat{u}\cdot e_1| - 1| \le \epsilon_0;$$
$$\partial_\rho f > 0, \quad \forall\, 1 + c(\epsilon_0) \le \rho \le R_1, \ \ \forall\, \hat{u} \in \mathbb{S}^{n-1} \text{with } ||\hat{u}\cdot e_1| - 1| \le \epsilon_0.$$

*Proof of Theorem 4.3.* We shall work with the variable $R = \rho^2$. Write

$$f(u) = \frac{1}{m}\sum_{j=1}^{m} F(\rho^2, (a_j\cdot\hat{u})^2, (a_j\cdot e_1)^2),$$

where

$$F(R,s,t) = R(s+2\beta) + 3t - 2\sqrt{R(\beta+s)+t}\sqrt{\beta R + 2t}.$$

Observe that

$$\begin{aligned}\partial_R F &= s + 2\beta - \Big(\beta\sqrt{\frac{R(\beta+s)+t}{\beta R+2t}} + (\beta+s)\sqrt{\frac{\beta R+2t}{R(\beta+s)+t}}\Big) \\ &= s + 2\beta - F_0(\sqrt{z(R,s,t)}),\end{aligned} \tag{4.2}$$

where

$$F_0(y) = \beta y + (\beta+s)y^{-1}, \quad z(R,s,t) = \frac{R(\beta+s)+t}{\beta R + 2t}.$$

Note that $F_0'(y) < 0$ for any $0 < y < \sqrt{\frac{\beta+s}{\beta}}$. It is easy to check that for $t > 0$, $s \ge 0$ (note that $t = (a_j\cdot e_1)^2 > 0$ for all $j$ almost surely)

$$z(R,s,t) < \frac{\beta+s}{\beta}.$$

Furthermore, since

$$z(R,s,t) = \frac{\beta+s}{\beta} - \frac{t\cdot\frac{\beta+2s}{\beta}}{\beta R + 2t},$$

we have $\partial_R z(R,s,t) > 0$ for $R > 0$, $t > 0$, $s \ge 0$. Thus

$$\partial_{RR} F > 0, \qquad \forall\, R > 0, \forall\, s \ge 0, t > 0. \tag{4.3}$$

On the other hand, by directly using (4.2), it is not difficult to check that for $R \sim 1$,

$$|(\partial_R F)(R,s,t) - (\partial_R F)(R,t,t)| \lesssim |s-t|. \tag{4.4}$$

Also observe that for $R \sim 1$, we have

$$
\begin{aligned}
& z(R,t,t) \sim 1, \quad \partial_R z(R,t,t) \sim \frac{t}{1+t}; \\
& -F_0'(\sqrt{z(R,t,t)}) = -\beta + \frac{\beta+t}{z(R,t,t)} = \frac{1}{z(R,t,t)} \cdot \frac{t(\beta+2t)}{\beta R+2t} \sim t; \\
& \partial_{RR} F(R,t,t) = -F_0'(\sqrt{z(R,t,t)}) \frac{1}{2} z(R,t,t)^{-\frac{1}{2}} \cdot \partial_R z(R,t,t) \sim \frac{t^2}{1+t}.
\end{aligned} \tag{4.5}
$$

Note that $z(1,t,t)=1$ and $\partial_R F(1,t,t)=0$. By using (4.3), (4.4) and (4.5), we obtain for $R \ge 1+\eta_0$ ($0<\eta_0 \ll 1$ will be specified later)

$$
\begin{aligned}
(\partial_R F)(R,s,t) &> (\partial_R F)(1+\eta_0,s,t) \\
&\ge (\partial_R F)(1+\eta_0,t,t) - B_1|s-t| \\
&\ge B_2 \cdot \frac{t^2}{1+t} \cdot \eta_0 - B_1|s-t|,
\end{aligned}
$$

where $B_1>0$, $B_2>0$ are constants depending only on $\beta$. Consequently we have for $R \ge 1+\eta_0$,

$$
\partial_R f \ge B_2 \eta_0 \frac{1}{m} \sum_{j=1}^m \frac{(a_j \cdot e_1)^4}{1+(a_j\cdot e_1)^2} - B_1 \frac{1}{m}\sum_{j=1}^m |(a_j\cdot \hat{u})^2 - (a_j \cdot e_1)^2|.
$$

Clearly with high probability,

$$
\begin{aligned}
& \frac{1}{m}\sum_{j=1}^m \frac{(a_j\cdot e_1)^4}{1+(a_j\cdot e_1)^2} \ge B_3 > 0, \\
& \frac{1}{m}\sum_{j=1}^m |(a_j\cdot\hat{u})^2 - (a_j\cdot e_1)^2| \le \frac{1}{m}\sum_{j=1}^m |a_j\cdot(\hat{u}-e_1)| \cdot |a_j \cdot (\hat{u}+e_1)| \\
& \qquad \le B_4 \min\{\|\hat{u}-e_1\|_2, \|\hat{u}+e_1\|_2\}, \qquad \forall\, \hat{u} \in \mathbb{S}^{n-1},
\end{aligned}
$$

where $B_3>0$, $B_4>0$ are absolute constants. Clearly then for $R \ge 1+\eta_0$, we have (below $B_5>0$, $B_6>0$ are constants depending only on $\beta$)

$$
\partial_R f \ge B_5 \eta_0 - B_6 \sqrt{1-|\hat{u}\cdot e_1|} > 0
$$

if $\eta_0$ is chosen suitably small. The case for $R \le 1-\eta_0$ is similar. We omit the details. $\square$

4.2. **Analysis of the regime $\rho \sim 1$, $||\hat{u}\cdot e_1|-1| \ge \epsilon_0 > 0$.**

In this section we consider the regime $\rho\sim 1$, $|\hat{u}\cdot e_1| < 1-\epsilon_0$, where $0<\epsilon_0 \ll 1$.

To simplify the discussion, we use the coordinate system

$$
\begin{aligned}
\hat{u} &= (\hat{u}\cdot e_1)e_1 + \tilde{u}, \\
&= te_1 + \sqrt{1-t^2}e^{\perp},
\end{aligned}
$$

where $e^{\perp} \in \mathbb{S}^{n-1}$ satisfies $e^{\perp}\cdot e_1 = 0$. Clearly in the regime $\rho \sim 1$, $|t|<1$, we have a smooth representation

$$
u = \rho\hat{u} = \rho\cdot(te_1 + \sqrt{1-t^2}e^{\perp}) = \psi(\rho,t,e^{\perp}).
$$

By Proposition 3.1, we can simplify the computation by examining only at the derivatives $\partial_t$ and $\partial_{tt}$. We shall use these in the regime $|t|<1-\epsilon_0$ where $0<\epsilon_0\ll 1$.

Now observe

$$\frac{1}{2}\mathbb{E}f = \frac{1}{2}(1+2\beta)\rho^2 + \frac{3}{2} - \mathbb{E}\sqrt{\beta\rho^2 + \rho^2(a\cdot\hat{u})^2 + (a\cdot e_1)^2}\sqrt{\beta\rho^2 + 2(a\cdot e_1)^2},$$

where $a \sim \mathcal{N}(0, \mathrm{I_n})$.

Denote $X_1 = a\cdot e_1$ and $Y_1 = a\cdot e^{\perp}$ so that $a\cdot\hat{u} = tX_1 + \sqrt{1-t^2}Y_1 =: X_t$.

We focus on the term

$$\begin{aligned}&\mathbb{E}\sqrt{\beta\rho^2 + \rho^2(a\cdot\hat{u})^2 + (a\cdot e_1)^2}\sqrt{\beta\rho^2 + 2(a\cdot e_1)^2}\\ &= \mathbb{E}\sqrt{\beta\rho^2 + \rho^2 X_t^2 + X_1^2}\sqrt{\beta\rho^2 + 2X_1^2} =: h_\infty(\rho, t).\end{aligned}$$

**Lemma 4.3** (The limiting profile)**.** *For any $0 < \eta_0 \ll 1$, the following hold:*

*(1) We have*

$$\sup_{|t|\le 1-\eta_0, \rho\sim 1} (|\partial_t h_\infty(\rho,t)| + |\partial_{tt} h_\infty(\rho,t)| + |\partial_{ttt} h_\infty(\rho,t)|) \lesssim 1.$$

*(2) $|\partial_t h_\infty(\rho,t)| \gtrsim |t|$ for $0 < |t| < 1$, $\rho \sim 1$.*

*(3) $\partial_{tt} h_\infty(\rho,t) \gtrsim 1$ for $|t| \ll 1$, $\rho \sim 1$.*

*Proof.* See appendix. □

**Theorem 4.4** (The regime $\|u\|_2 \sim 1$, $||\hat{u}\cdot e_1| - 1| > \eta_0$ is fine)**.** *For any given $0 < \eta_0 \ll 1$ and $0 < c_1 < c_2 < \infty$, if $m \gtrsim n$, then the following hold with high probability:*

*In the regime $c_1 < \rho = \|u\|_2 < c_2$, $||\hat{u}\cdot e_1| - 1| > \epsilon_0$, there are only two possibilities:*

*(1) $\|\nabla f(u)\|_2 > 0$;*

*(2) $\nabla f(u) = 0$, and $f$ has a negative directional curvature at this point.*

*Proof of Theorem 4.4.* Denote $X_j = a_j \cdot e_1$ and

$$\begin{aligned}g(\rho, t, e^{\perp}) &= 2\beta\rho^2 + \rho^2\frac{1}{m}\sum_{j=1}^m (a_j\cdot\hat{u})^2 - 2\frac{1}{m}\sum_{j=1}^m \sqrt{\beta\rho^2 + \rho^2(a_j\cdot\hat{u})^2 + X_j^2}\sqrt{\beta\rho^2 + 2X_j^2}\\ &=: 2\beta\rho^2 + \rho^2 h_0(\rho, t, e^{\perp}) - 2h(\rho, t, e^{\perp}),\end{aligned}$$

where

$$a_j\cdot\hat{u} = tX_j + \sqrt{1-t^2}Y_j, \quad X_j = a_j\cdot e_1,\ Y_j = a_j\cdot e^{\perp}.$$

By the same argument as in the proof of Theorem 3.3, we have for any small $\epsilon > 0$ and $m \gtrsim n$, it holds with high probability that

$$|\partial_t h_0 - \mathbb{E}\partial_t h_0| + |\partial_{tt} h_0 - \mathbb{E}\partial_{tt} h_0| \le \epsilon, \qquad \forall\, |t| \le 1-\epsilon_0, e^{\perp}\cdot e_1 = 0, e^{\perp}\in\mathbb{S}^{n-1}.$$

Note also $\mathbb{E}\partial_t h_0 = 0$ and $\mathbb{E}\partial_{tt} h_0 = 0$.

By Lemma D.4, for any small $\epsilon > 0$ and $m \gtrsim n$, it also holds with high probability that

$$|\partial_t h - \mathbb{E}\partial_t h| + |\partial_{tt} h - \mathbb{E}\partial_{tt} h| \le \epsilon, \qquad \forall\, |t| \le 1-\epsilon_0, e^{\perp}\cdot e_1 = 0, e^{\perp}\in\mathbb{S}^{n-1}, \forall\, c_1 \le \rho \le c_2.$$

We then obtain for small $\epsilon > 0$, if $m \gtrsim n$, it holds with high probability that

$$|\partial_t g - \mathbb{E}\partial_t g| \leq 2\epsilon;$$
$$\partial_{tt} g \leq \mathbb{E}\partial_{tt} g + 2\epsilon.$$

Clearly

$$\mathbb{E}\partial_t g = -\partial_t h_\infty(\rho, t);$$
$$\mathbb{E}\partial_{tt} g = -\partial_{tt} h_\infty(\rho, t).$$

By Lemma 4.3, we can take $t_0 \ll 1$ such that

$$\partial_{tt} h_\infty(\rho, t) \geq \epsilon_1 > 0, \qquad \forall\, |t| \leq t_0, c_1 \leq \rho \leq c_2;$$
$$|\partial_t h_\infty(\rho, t)| \geq \epsilon_2 > 0, \qquad \forall\, t_0 \leq |t| \leq 1 - \epsilon_0, c_1 \leq \rho \leq c_2.$$

By taking $\epsilon > 0$ sufficiently small, we can then guarantee that

$$|\partial_t g| > \epsilon_3 > 0, \qquad \forall\, |t| \leq t_0, c_1 \leq \rho \leq c_2;$$
$$\partial_{tt} g \leq -\epsilon_4 < 0, \qquad \forall\, t_0 \leq |t| \leq 1 - \epsilon_0, c_1 \leq \rho \leq c_2.$$

The desired result then follows from Proposition 3.1. □

4.3. **Strong convexity near the global minimizers $u = \pm e_1$: analysis of the limiting profile.**

In this section we shall show that in the small neighborhood of $u = \pm e_1$ where

$$||\hat{u} \cdot e_1| - 1| \ll 1, \quad |\rho - 1| \ll 1,$$

that the Hessian of the expectation of the loss function must be strictly positive definite. In yet other words $\mathbb{E}f$ must be strictly convex in this neighborhood so that $u = \pm e_1$ are the unique minimizers. To this end consider

$$h(u) = \frac{1}{2}(\mathbb{E}f - 3) = \frac{1}{2}(1 + 2\beta)\rho^2 - \mathbb{E}\sqrt{\beta\rho^2 + \rho^2(a \cdot \hat{u})^2 + (a \cdot e_1)^2}\sqrt{\beta\rho^2 + 2(a \cdot e_1)^2}, \tag{4.6}$$

where $a \sim \mathcal{N}(0, \mathrm{I}_n)$.

**Theorem 4.5** (Strong convexity of $\mathbb{E}f$ when $\|u \pm e_1\| \ll 1$)**.** *Consider $h$ defined by (4.6). There exists $0 < \epsilon_0 \ll 1$ such that the following hold:*

*(1) If $\|u - e_1\|_2 \leq \epsilon_0$, then for any $\xi \in \mathbb{S}^{n-1}$, we have*

$$\sum_{i,j=1}^n \xi_i \xi_j (\partial_i \partial_j h)(u) \geq \gamma_1 > 0,$$

*where $\gamma_1$ is a constant.*

*(2) If $\|u + e_1\|_2 \leq \epsilon_0$, then for any $\xi \in \mathbb{S}^{n-1}$, we have*

$$\sum_{i,j=1}^n \xi_i \xi_j (\partial_i \partial_j h)(u) \geq \gamma_1 > 0.$$

*Proof of Theorem 4.5.* We shall only consider the case $\|u-e_1\|_2 \ll 1$. The other case $\|u+e_1\|_2 \ll 1$ is similar and therefore omitted. Note that

$$\|u-e_1\|_2^2 = \|\rho\hat{u}-e_1\|_2^2 = (\rho-1)^2 + 2\rho(1-t) \le \epsilon_0^2,$$

where $t = \hat{u}\cdot e_1$. Thus for $0<\epsilon_0 \ll 1$, we have

$$|\rho-1| \le \epsilon_0, \qquad 1-\epsilon_0^2 \le t \le 1.$$

We make a change of variable and write (recall $1-\epsilon_0^2 \le \hat{u}\cdot e_1 \to 1$),

$$\hat{u} = \sqrt{1-s^2}e_1 + se^{\perp}, \qquad e^{\perp}\cdot e_1 = 0,\ e^{\perp} \in \mathbb{S}^{n-1},$$

where we assume $0 \le s \ll 1$. Note that $s = \frac{|u'|}{\rho}$, and $u' = u-(u\cdot e_1)e_1 = (0, u_2, \cdots, u_n)$.

To calculate $\partial^2 h$ we need to compute the Hessian expressed in the $(\rho, s)$ coordinate. It is not difficult to check that by (4.6), the value of $h(u)$ depends only on $(\rho, s)$. Thus by a slight abuse of notation we write $h(u) = h(|u|, \frac{|u'|}{|u|}) = h(\rho, s)$ (we denote $|u| = \|u\|_2$, $|u'| = \|u'\|_2$) and compute (below we assume $s>0$ so that $|u'|>0$)

$$\begin{aligned}\sum_{i,j=1}^n \xi_i\xi_j\partial_{ij}h = {} & \partial_{\rho\rho}h\cdot a^2 + 2a\partial_{\rho s}h\cdot(-\frac{as}{\rho}+\frac{b}{\rho}) \\ & + \partial_\rho h\cdot(\frac{|\xi|^2-|\xi\cdot\hat{u}|^2}{\rho}) \\ & + \partial_{ss}h\cdot(\frac{a^2s^2-2abs}{\rho^2}) + (\partial_{ss}h - \frac{1}{s}\partial_s h)\frac{b^2}{\rho^2} \\ & + \partial_s h\cdot(-\frac{|\xi|^2}{\rho^2}s + 3\frac{a^2 s}{\rho^2} - 2\frac{ab}{\rho^2} + \frac{|\xi'|^2}{\rho^2 s})\end{aligned}$$

In the above computation, one does not need to worry about the formal singularity caused by $\frac{1}{s}$. Since (by Lemma D.6) $\partial_s h(\rho, s=0) = 0$ for any $\rho>0$, we write

$$(\partial_s h)(\rho,s)\cdot\frac{1}{s} = \frac{(\partial_s h)(\rho,s)-(\partial_s h)(\rho,0)}{s} = \int_0^1 (\partial_{ss}h)(\rho,\theta s)d\theta, \quad s>0.$$

In particular we have

$$\begin{aligned}&\lim_{s\to 0^+}(\partial_s h)(\rho,s)\cdot\frac{1}{s} = (\partial_{ss}h)(\rho,0); \\ &\Big|(\partial_{ss}h)(\rho,s) - \frac{1}{s}\partial_s h_1(\rho,s)\Big| = O(s) \to 0, \quad \text{as } s\to 0.\end{aligned}$$

By using this observation and Lemma D.6, we obtain

$$\begin{aligned}\sum_{i,j=1}^n \xi_i\xi_j(\partial_{ij}h)(e_1) &= (\partial_{\rho\rho}h)(1,0)\cdot a^2\Big|_{a=\xi_1} + (\partial_{ss}h)(1,0)\cdot|\xi'|^2 \\ &\ge \gamma_0\cdot|\xi|^2, \qquad \forall\,\xi\in\mathbb{S}^{n-1},\end{aligned}$$

where $\gamma_0 > 0$ is a constant. Now for $\|u - e_1\|_2 \ll 1$, by using Lemma D.5 and Lemma D.6, we have

$$\begin{aligned}
&\Big|\sum_{i,j=1}^{n} \xi_i \xi_j \Big( (\partial_{ij} h)(e_1) - (\partial_{ij} h)(u) \Big) \Big| \\
&\lesssim |(\partial_{\rho\rho} h)(\rho, s) - (\partial_{\rho\rho} h)(1,0)| + |(\partial_{\rho\rho} h)(\rho, s)| \cdot |(\xi \cdot \hat{u})^2 - (\xi \cdot e_1)^2| \\
&\quad + |(\partial_{\rho s} h)(\rho, s)| \cdot (1+s) + |(\partial_{\rho} h)(\rho, s)| + |\partial_{ss} h(\rho, s)| \cdot (s + s^2) \\
&\quad + |\partial_s h(\rho, s)| + \Big| \frac{1}{\rho^2} \int_0^1 (\partial_{ss} h)(\rho, \theta s) d\theta - (\partial_{ss} h)(1,0) \Big| \\
&\quad + \Big| (\partial_{ss} h)(\rho, s) - (\partial_s h)(\rho, s) \cdot \frac{1}{s} \Big| \cdot \frac{b^2}{\rho^2} \\
&\lesssim O(|\rho - 1| + |s| + \|\hat{u} - e_1\|_2).
\end{aligned}$$

It follows that if $\|u - e_1\|_2$ is sufficiently small, we then have

$$\sum_{i,j=1}^{n} \xi_i \xi_j (\partial_{ij} h)(u) \ge \frac{\gamma_0}{2} |\xi|^2, \qquad \forall\, \xi \in \mathbb{S}^{n-1}.$$

□

### 4.4. **Near the global minimizer: strong convexity.**

In this section we show strong convexity of the loss function $f(u)$ near the global minimizer $u = \pm e_1$.

**Theorem 4.6** (Strong convexity near the global minimizer). *There exists $0 < \epsilon_0 \ll 1$ and a constant $\beta_1 > 0$ such that if $m \gtrsim n$, then the following hold with high probability:*

*(1) If $\|u - e_1\|_2 \le \epsilon_0$, then*

$$\sum_{i,j=1}^{n} \xi_i \xi_j (\partial_{ij} f)(u) \ge \gamma > 0, \qquad \forall\, \xi \in \mathbb{S}^{n-1},$$

*where $\gamma$ is a constant.*

*(2) If $\|u + e_1\|_2 \le \epsilon_0$, then*

$$\sum_{i,j=1}^{n} \xi_i \xi_j (\partial_{ij} f)(u) \ge \gamma > 0, \qquad \forall\, \xi \in \mathbb{S}^{n-1},$$

*where $\gamma$ is a constant.*

*In yet other words, $f(u)$ is strongly convex in a sufficiently small neighborhood of $\pm e_1$.*

*Proof of Theorem 4.6.* Recall

$$f(u) = 2 f_0(u) + \frac{1}{m} \sum_{k=1}^{m} \Big( (a_k \cdot u)^2 + 2\beta |u|^2 + 3 (a_k \cdot e_1)^2 \Big),$$

where

$$f_0(u) = -\frac{1}{m} \sum_{k=1}^{m} \sqrt{\beta |u|^2 + (a_k \cdot u)^2 + (a_k \cdot e_1)^2} \cdot \sqrt{\beta |u|^2 + 2 (a_k \cdot e_1)^2}.$$

Clearly

$$\sum_{i,j=1}^{n} \xi_i \xi_j \partial_{ij} f(u) = 2\Big(\frac{1}{m}\sum_{k=1}^{m} |a_k \cdot \xi|^2\Big) + 4\beta|\xi|^2 + 2\sum_{i,j=1}^{n} \xi_i\xi_j \partial_{ij} f_0(u).$$

Obviously we have for $m \gtrsim n$, it holds with high probability that

$$\Big|\frac{1}{m}\sum_{k=1}^{m} |a_k \cdot \xi|^2 - 1\Big| \le \frac{\epsilon}{100}, \qquad \forall\, \xi \in \mathbb{S}^{n-1}.$$

By Lemma D.7, we have

$$\Big|\sum_{i,j=1}^{n} \xi_i\xi_j(\partial_{ij} f_0(u) - \mathbb{E}\partial_{ij} f_0(u))\Big| \le \frac{\epsilon}{100}, \qquad \forall\, \xi \in \mathbb{S}^{n-1},\ \forall\, \frac{1}{3} \le \|u\|_2 \le 3.$$

Thus we have

$$\Big|\sum_{i,j=1}^{n} \xi_i\xi_j(\partial_{ij} f(u) - \mathbb{E}\partial_{ij} f(u))\Big| \le \frac{\epsilon}{100}, \qquad \forall\, \xi \in \mathbb{S}^{n-1},\ \forall\, \frac{1}{3} \le \|u\|_2 \le 3.$$

The desired result then follows from Theorem 4.5 by taking $\epsilon > 0$ sufficiently small. □

We now complete the proof of the main theorem.

*Proof of Theorem 4.1.* We proceed in several steps.

(1) By Theorem 4.2, we see that with high probability the function $f(u)$ has non-vanishing gradient in the regimes

$$0 < \|u\|_2 \le R_2 = R_2(\beta)$$

and

$$\|u\|_2 \ge R_1 = R_1(\beta),$$

where $R_1 > 0$, $R_2 > 0$ depend only on $\beta$. Moreover the point $u = 0$ is a local maximum point with strictly negative-definite Hessian.

(2) By Theorem 4.6, there exists $\epsilon_0 > 0$ sufficiently small, such that with high probability, $f(u)$ is strongly convex in the neighborhood $\|u \pm e_1\|_2 \le \epsilon_0$.

(3) By Theorem 4.3, we have that with high probability

$$\|\nabla f\|_2 > 0,$$

if $|\rho - 1| \ge c(\eta_0)$ and $||\hat{u} \cdot e_1| - 1| \le \eta_0$. Here we recall $\rho = \|u\|_2$ and $u = \rho\hat{u}$. Observe that

$$\|u \pm e_1\|_2^2 = (\rho - 1)^2 + 2\rho(1 \pm \hat{u} \cdot e_1).$$

By taking $\eta_0 = \epsilon_0^2/100$, we see that $\|u \pm e_1\|_2 > \epsilon_0$, $||\hat{u} \cdot e_1| - 1| \le \eta_0$ must imply

$$|\rho - 1| > \frac{\epsilon_0}{10}.$$

Thus it remains for us to treat the regime $||\hat{u} \cdot e_1| - 1| > \eta_0$, $R_1 \le \|u\|_2 \le R_2$.

(4) In the regime $||\hat{u}\cdot e_1|-1|>\eta_0$, $\|u\|_2\sim 1$, we have by Theorem 4.4, with high probability it holds that either the function has a non-vanishing gradient at the point $u$, or the gradient vanishes at $u$, but $f$ has a negative directional curvature at this point.

$\square$

## 5. The third model: 3A

Consider

$$f(u)=\frac{1}{m}\sum_{k=1}^{m}\frac{((a_k\cdot u)^2-(a_k\cdot x)^2)^2}{(a_k\cdot x)^2}. \tag{5.1}$$

We shall adopt similar notation as in Section 3.

**Theorem 5.1.** *Assume $\{a_k\}_{k=1}^m$ are i.i.d. standard Gaussian random vectors and $x\neq 0$. There exist positive absolute constants c, C, such that if $m\geq Cn$, then with probability at least $1-e^{-cm}$ the loss function $f=f(u)$ defined by (5.1) has no spurious local minimizers. The only global minimizer is $\pm x$, and the loss function is strongly convex in a neighborhood of $\pm x$. The point $u=0$ is a local maximum point with strictly negative-definite Hessian. All other critical points are strict saddles, i.e., each saddle point has a neighborhood where the function has negative directional curvature.*

Without loss of generality we shall assume $x=e_1$ throughout the rest of the proof. Note that the set $\bigcup_{k=1}^m\{a_k\cdot e_1=0\}$ has measure zero. Thus for typical realization we have $a_k\cdot e_1\neq 0$ for all $k$. This means that the loss function $f(u)$ defined by (5.1) is smooth almost surely.

We denote the Hessian of the function $f(u)$ along the $\xi$-direction ($\xi\in\mathbb{S}^{n-1}$) as

$$\begin{aligned}H_{\xi\xi}(u)&=\sum_{i,j=1}^{n}\xi_i\xi_j(\partial_{ij}f)(u)\\&=4\frac{1}{m}\sum_{k=1}^{m}\Big(3\frac{(a_k\cdot\xi)^2(a_k\cdot u)^2}{(a_k\cdot e_1)^2}-(a_k\cdot\xi)^2\Big).\end{aligned} \tag{5.2}$$

### 5.1. Strong convexity near the global minimizers $u=\pm e_1$.

**Theorem 5.2** (Strong convexity near $u=\pm e_1$)**.** *There exists an absolute constant $0<\epsilon_0\ll 1$ such that the following hold. For $m\gtrsim n$, it holds with high probability that*

$$H_{\xi\xi}(u)\geq 1,\qquad \forall\,\xi\in\mathbb{S}^{n-1},\quad \forall\,u \text{ with } \|u\pm e_1\|_2\leq\epsilon_0.$$

*Proof of Theorem 5.2.* By Lemma E.1, we can take $\epsilon>0$ sufficiently small, $N$ sufficiently large such that

$$\mathbb{E}\frac{(a_k\cdot\xi)^2(a_k\cdot e_1)^2}{\epsilon+(a_k\cdot e_1)^2}\phi(\frac{a_k\cdot\xi}{N})\geq 0.99,\qquad \forall\,\xi\in\mathbb{S}^{n-1},\ \forall\,1\leq k\leq m.$$

In the above $\phi \in C_c^\infty(\mathbb{R})$ satisfies $0 \le \phi(x) \le 1$ for all $x$, $\phi(x) = 1$ for $|x| \le 1$ and $\phi(x) = 0$ for $|x| \ge 2$. Clearly then if $\|u \pm e_1\|_2 \le \epsilon_0$ and $\epsilon_0$ is sufficiently small (depending on $\epsilon$ and $N$), then

$$\mathbb{E}\frac{(a_k\cdot\xi)^2(a_k\cdot u)^2}{\epsilon+(a_k\cdot e_1)^2}\phi(\frac{a_k\cdot\xi}{N}) \ge 0.98, \qquad \forall\,\xi\in\mathbb{S}^{n-1},\ \forall\,1\le k\le m.$$

The above term inside the expectation is clearly OK for union bounds. Thus for $\|u \pm e_1\| \le \epsilon_0$ and $m \gtrsim n$, it holds with high probability that

$$\begin{aligned}\frac{1}{4}H_{\xi\xi}(u) &\ge \frac{1}{m}\sum_{k=1}^m\Bigl(\frac{(a_k\cdot\xi)^2(a_k\cdot u)^2}{\epsilon+(a_k\cdot e_1)^2}\phi(\frac{a_k\cdot\xi}{N})-(a_k\cdot\xi)^2\Bigr)\\ &\ge 3\cdot 0.97-1.01, \qquad \forall\,\xi\in\mathbb{S}^{n-1}.\end{aligned}$$

Thus the desired inequality follows. □

5.2. **The regimes $\|u\|_2 \ll 1$ and $\|u\|_2 \gg 1$ are fine.**

We first investigate the point $u = 0$. It is trivial to verify that $\nabla f(0) = 0$ since $a_k \cdot e_1 \ne 0$ for all $k$ almost surely.

**Lemma 5.1** ($u = 0$ has strictly negative-definite Hessian)**.** *We have $u = 0$ is a local maximum point with strictly negative-definite Hessian. More precisely, for $m \gtrsim n$, it holds with high probability that*

$$\sum_{k,l=1}^n \xi_k\xi_l(\partial_{kl}f)(0) \le -1, \quad \forall\,\xi\in\mathbb{S}^{n-1}.$$

*Proof of Lemma 5.1.* By (5.2), we have

$$H_{\xi\xi}(0) = -4\frac{1}{m}\sum_{k=1}^m (a_k\cdot\xi)^2.$$

The desired conclusion then easily follows from Bernstein. □

Write $u = \sqrt{R}\hat{u}$ where $\hat{u} \in S^{n-1}$ and $R > 0$. Then

$$f(u) = \frac{1}{m}\sum_{k=1}^m \frac{\Bigl(R(a_k\cdot\hat{u})^2-(a_k\cdot e_1)^2\Bigr)^2}{(a_k\cdot e_1)^2}.$$

Clearly

$$\partial_R f = 2R\frac{1}{m}\sum_{k=1}^m\frac{(a_k\cdot\hat{u})^4}{(a_k\cdot e_1)^2} - 2\frac{1}{m}\sum_{k=1}^m (a_k\cdot\hat{u})^2; \tag{5.3}$$

$$\partial_{RR} f = 2\frac{1}{m}\sum_{k=1}^m\frac{(a_k\cdot\hat{u})^4}{(a_k\cdot e_1)^2}. \tag{5.4}$$

**Lemma 5.2** (The regime $\|u\|_2 \ge 1+\epsilon_0$ is OK)**.** *Let $0 < \epsilon_0 \ll 1$ be any given small constant. Then the following hold:*

*For $m \gtrsim n$, with high probability it holds that*

$$\partial_R f > 0, \quad \forall\,R \ge 1+\epsilon_0,\ \forall\,\hat{u}\in\mathbb{S}^{n-1}.$$

*Proof.* Denote $X_k = a_k \cdot e_1$ and $Z_k = a_k \cdot \hat{u}$. By (5.3) and Cauchy-Schwartz, we have

$$\begin{aligned}\partial_R f &\geq 2R \frac{1}{m} \frac{\left(\sum_{k=1}^m (a_k \cdot \hat{u})^2\right)^2}{\sum_{k=1}^m (a_k \cdot e_1)^2} - 2\frac{1}{m}\sum_{k=1}^m (a_k \cdot \hat{u})^2 \\ &\geq\ 2R \cdot (1-\delta_1) - 2(1+\delta_1), \qquad \forall\, \hat{u} \in \mathbb{S}^{n-1},\end{aligned}$$

where $0 < \delta_1 \ll 1$ is an absolute constant which we can take to be sufficiently small, and in the last inequality we have used Bernstein. The desired result then easily follows by taking $R \geq R_1 = \frac{1+2\delta_1}{1-\delta_1}$ and choosing $\delta_1$ such that $R_1 \leq 1+\epsilon_0$. $\square$

From (5.3), due to the highly irregular coefficients near $R$, it is difficult to control the upper bound of $\partial_R f$ in the regime $R \ll 1$. To resolve this difficulty, we shall examine the Hessian in this regime.

**Lemma 5.3** (The regime $\|u\|_2 \leq \frac{1}{3}$ is OK)**.** *For $m \gtrsim n$, with high probability it holds that*

$$H_{e_1 e_1}(u) \leq -1 < 0, \quad \forall\, 0 < \|u\|_2 \leq \frac{1}{3},$$

*where $H_{e_1e_1}$ is defined in (5.2).*

*Proof.* By (5.2), we have for $m \gtrsim n$ with high probability,

$$\begin{aligned}\frac{1}{4} H_{e_1 e_1}(u) &= \frac{1}{m}\sum_{k=1}^m \Big(3(a_k \cdot u)^2 - (a_k \cdot e_1)^2\Big) \\ &\leq \|u\|_2^2 \cdot 3 \cdot \frac{10}{9} - \frac{8}{9}.\end{aligned}$$

The desired result then easily follows. $\square$

**Theorem 5.3** (The regimes $\|u\|_2 \leq \frac{1}{3}$ and $\|u\|_2 \geq 1+\epsilon_0$ are OK)**.** *Let $0 < \epsilon_0 \ll 1$ be a given small constant. For $m \gtrsim n$, with high probability the following hold:*

*(1) We have*

$$\partial_R f > 0, \qquad \forall\, R \geq 1+\epsilon_0, \quad \forall\, \hat{u} \in \mathbb{S}^{n-1}.$$

*(2) The point $u = 0$ is a local maximum point with strictly negative-definite Hessian,*

$$\sum_{k,l=1}^n \xi_k \xi_l (\partial_{kl} f)(0) \leq -1, \qquad \forall\, \xi \in \mathbb{S}^{n-1}.$$

*(3) We have*

$$H_{e_1 e_1}(u) \leq -1, \qquad \forall\, \|u\|_2 \leq \frac{1}{3}.$$

*Proof of Theorem 5.3.* This follows from Lemma 5.1, 5.2 and 5.3. $\square$

**Theorem 5.4** (The regime $\|u\|_2 \sim 1$, $||\hat{u} \cdot e_1| - 1| \geq \eta_0$)**.** *Let $0 < \eta_0 \ll 1$ be given. Then for $m \gtrsim n$, the following hold with high probability:*

*Suppose $u = \sqrt{R}\hat{u}$, $\frac{1}{9} \leq R \leq 2$, and*

$$\Big| |\hat{u} \cdot e_1| - 1 \Big| \geq \eta_0.$$

*If $(\partial_R f)(u) = 0$, then we must have*

$$H_{e_1 e_1}(u) < 0.$$

*Proof of Theorem 5.4.* By (5.3), we have if $\partial_R f(u) = 0$, then

$$R \frac{1}{m} \sum_{k=1}^{m} \frac{(a_k \cdot \hat{u})^4}{(a_k \cdot e_1)^2} = \frac{1}{m} \sum_{k=1}^{m} (a_k \cdot \hat{u})^2.$$

By Lemma E.2, we have for $m \gtrsim n$, it holds with high probability that

$$\frac{1}{m} \sum_{k=1}^{m} \frac{(a_k \cdot \hat{u})^4}{(a_k \cdot e_1)^2} \geq 100, \quad \forall\, \hat{u} \in \mathbb{S}^{n-1} \text{ with } ||\hat{u} \cdot e_1| - 1| \geq \eta_0.$$

Clearly then $R \leq \frac{1}{50}$ with high probability. Thus it follows easily that $H_{e_1 e_1}(u) < 0$ also with high probability. $\square$

**Theorem 5.5** (Localization of $R$ when $||\hat{u} \cdot e_1| - 1| \leq \eta_0$, $R \leq 1 + \eta_0$ and $u$ is a critical point)**.** *Let $0 < \eta_0 \ll 1$ be given. For $m \gtrsim n$, the following hold with high probability:*

*Assume $u = \sqrt{R}\hat{u}$ is a critical point with $\frac{1}{9} \leq R \leq 1 + \eta_0$, and $||\hat{u} \cdot e_1| - 1| \leq \eta_0$. Then we must have*

$$|R - 1| \leq c(\eta_0),$$

*where $c(\eta_0) \to 0$ as $\eta_0 \to 0$.*

*Proof of Theorem 5.5.* Denote $\partial_\xi f = \xi \cdot \nabla f$ for $\xi \in \mathbb{S}^{n-1}$. It is not difficult to check that

$$\frac{1}{4} \partial_\xi f = \frac{1}{m} \sum_{k=1}^{m} \frac{(a_k \cdot u)^3 (a_k \cdot \xi)}{X_k^2} - \frac{1}{m} \sum_{k=1}^{m} (a_k \cdot u)(a_k \cdot \xi) = 0,$$

where $X_k = a_k \cdot e_1$.

Setting $\xi = \hat{u}$ and $\xi = e_1$ respectively give us two equations:

$$R \cdot \Big( \frac{1}{m} \sum_{k=1}^{m} \frac{(a_k \cdot \hat{u})^4}{X_k^2} \Big) - \frac{1}{m} \sum_{k=1}^{m} (a_k \cdot \hat{u})^2 = 0, \tag{5.5}$$

$$R \cdot \Big( \frac{1}{m} \sum_{k=1}^{m} \frac{(a_k \cdot \hat{u})^3}{X_k} \Big) - \frac{1}{m} \sum_{k=1}^{m} (a_k \cdot \hat{u}) X_k = 0. \tag{5.6}$$

We then obtain

$$\Big( \frac{1}{m} \sum_{k=1}^{m} (a_k \cdot \hat{u})^2 \Big) \cdot \Big( \frac{1}{m} \sum_{k=1}^{m} \frac{(a_k \cdot \hat{u})^3}{X_k} \Big) = \Big( \frac{1}{m} \sum_{k=1}^{m} \frac{(a_k \cdot \hat{u})^4}{X_k^2} \Big) \cdot \Big( \frac{1}{m} \sum_{k=1}^{m} (a_k \cdot \hat{u}) X_k \Big). \tag{5.7}$$

Without loss of generality we assume $\|\hat{u} - e_1\| \leq \eta \ll 1$. Then with high probability we have

$$\frac{1}{m} \sum_{k=1}^{m} (a_k \cdot \hat{u}) X_k = 1 + O(\eta),$$

$$\frac{1}{m} \sum_{k=1}^{m} (a_k \cdot \hat{u})^2 = 1 + O(\eta).$$

Observe that by Cauchy-Schwartz,

$$\sum_{k=1}^m \frac{|a_k \cdot \hat{u}|^3}{|X_k|} \le \Big(\sum_{k=1}^m \frac{|a_k \cdot \hat{u}|^4}{X_k^2}\Big)^{\frac12} \cdot \Big(\sum_{k=1}^m (a_k \cdot \hat{u})^2\Big)^{\frac12}.$$

Plugging the above estimates into (5.7), we obtain

$$\sqrt{\frac{1}{m}\sum_{k=1}^m \frac{(a_k \cdot \hat{u})^4}{X_k^2}} \le 1 + O(\eta).$$

Using (5.5), we then get

$$R \ge 1 + O(\eta).$$

The desired result then easily follows. □

We now complete the proof of the main theorem.

*Proof of Theorem 5.1.* We proceed in several steps.

(1) By Theorem 5.2, the function $f(u)$ is strongly convex when $\|u \pm e_1\|_2 \ll 1$.

(2) By Theorem 5.3, $f$ has non-vanishing gradient when $R \ge 1 + \epsilon_0$. Also $H_{e_1e_1}(u) \le -1$ when $\|u\|_2 \le \frac13$. The point $u = 0$ is a strict local maximum point with strictly negative-definite Hessian.

(3) By Theorem 5.4, we have $H_{e_1e_1}(u) < 0$ if $\|u\|_2 \sim 1$ and $||\hat{u} \cdot e_1| - 1| \ge \epsilon_0$.

(4) Theorem 5.5 shows that if $R \le 1 + \epsilon_0$, $||\hat{u} \cdot e_1| - 1| \le \epsilon_0$ and $u$ is a critical point, then we must have $|R - 1| \le c(\epsilon_0) \ll 1$. In yet other words we must have $\|u \pm e_1\|_2 \ll 1$. This regime is then treated by Step 1.

□

## 6. Model 3B

Consider for $\beta > 0$,

$$f(u) = \frac{1}{m}\sum_{k=1}^m \frac{((a_k \cdot u)^2 - (a_k \cdot x)^2)^2}{\beta|u|^2 + (a_k \cdot x)^2}. \tag{6.1}$$

We shall adopt similar notation as in Section 3.

**Theorem 6.1.** *Let $0 < \beta < \infty$. Assume $\{a_k\}_{k=1}^m$ are i.i.d. standard Gaussian random vectors and $x \ne 0$. There exist positive constants $c$, $C$ depending only on $\beta$, such that if $m \ge Cn$, then with probability at least $1 - e^{-cm}$ the loss function $f = f(u)$ defined by (6.1) has no spurious local minimizers. The only global minimizer is $\pm x$, and the loss function is restrictively convex in a neighborhood of $\pm x$. The point $u = 0$ is a local maximum point with strictly negative-definite Hessian. All other critical points are strict saddles, i.e., each saddle point has a neighborhood where the function has negative directional curvature.*

*Remark* 6.1. See Theorem 6.4 for the precise statement concerning restrictive convexity.

Without loss of generality we shall assume $x = e_1$ throughout the rest of the proof.

## 6.1. The regimes $\|u\|_2 \ll 1$ and $\|u\|_2 \gg 1$ are fine.

We first investigate the point $u = 0$. It is trivial to verify that $\nabla f(0) = 0$ since $a_k \cdot e_1 \neq 0$ for all $k$ almost surely.

**Lemma 6.1** ($u = 0$ has strictly negative-definite Hessian). *We have $u = 0$ is local maximum point with strictly negative-definite Hessian. More precisely, it holds (almost surely) that*

$$\sum_{k,l=1}^{n} \xi_k \xi_l (\partial_{kl} f)(0) \leq -d_1, \quad \forall\, \xi \in \mathbb{S}^{n-1},$$

*where $d_1 > 0$ is a constant depending only on $\beta$. Also for $m \gtrsim n$, it holds with high probability that*

$$\sum_{k,l=1}^{n} \xi_k \xi_l (\partial_{kl} f)(0) \leq -d_2, \quad \forall\, \xi \in \mathbb{S}^{n-1},$$

*where $d_2 > 0$ is an absolute constant.*

*Proof of Lemma 5.1.* We begin by noting that since almost surely $a_k \cdot e_1 \neq 0$ for all $k$, the function $f$ is smooth at $u = 0$. It suffices for us to consider (write $u = \sqrt{t}\xi$)

$$G(t) = \frac{1}{m} \sum_{k=1}^{m} \frac{(t(a_k \cdot \xi)^2 - (a_k \cdot e_1)^2)^2}{\beta t + (a_k \cdot e_1)^2}.$$

Clearly

$$G'(0) = -\beta - 2\frac{1}{m} \sum_{k=1}^{m} (a_k \cdot \xi)^2.$$

The desired conclusion then easily follows (for the high probability statement, use Bernstein). □

Write $u = \sqrt{R}\hat{u}$ where $\hat{u} \in S^{n-1}$ and $R > 0$. Then

$$f(u) = \frac{1}{m} \sum_{k=1}^{m} \frac{\Big(R(a_k \cdot \hat{u})^2 - (a_k \cdot e_1)^2\Big)^2}{\beta R + (a_k \cdot e_1)^2}.$$

Clearly

$$\partial_R f = \frac{1}{m} \sum_{k=1}^{m} \frac{R^2(\beta (a_k \cdot \hat{u})^4) + 2R(a_k \cdot \hat{u})^4 (a_k \cdot e_1)^2 - \beta (a_k \cdot e_1)^4 - 2(a_k \cdot e_1)^4 (a_k \cdot \hat{u})^2}{(\beta R + (a_k \cdot e_1)^2)^2}; \tag{6.2}$$

$$\partial_{RR} f = 2\frac{1}{m} \sum_{k=1}^{m} \frac{(a_k \cdot e_1)^4 (\beta + (a_k \cdot \hat{u})^2)^2}{(\beta R + (a_k \cdot e_1)^2)^3}. \tag{6.3}$$

**Lemma 6.2** (The regime $\|u\|_2 \gg 1$ is OK). *There exist constants $R_1 = R_1(\beta) > 0$, $d_1 = d_1(\beta) > 0$ such that the following hold:*

*For $m \gtrsim n$, with high probability it holds that*

$$\partial_R f \geq d_1, \quad \forall\, R \geq R_1,\ \forall\, \hat{u} \in \mathbb{S}^{n-1}.$$

*Proof.* We only sketch the proof. Denote $X_k = a_k \cdot e_1$ and $Z_k = a_k \cdot \hat{u}$. Using the inequalities (assume $R \gg 1$ and denote by $C_1 > 0$ a constant depending only on $\beta$)

$$\beta R + X_k^2 \le R(\beta + X_k^2);$$
$$(\beta R + X_k^2)^2 \ge 4\beta R X_k^2;$$
$$\frac{X_k^4}{(\beta R + X_k^2)^2} \le C_1 \cdot (\frac{R}{R^2} + \chi_{|X_k| \ge R^{\frac{1}{4}}}),$$

we have

$$\partial_R f \ge \frac{1}{m}\sum_{k=1}^m \frac{\beta Z_k^4}{(\beta + X_k^2)^2}\phi(\frac{Z_k}{N}) - \frac{1}{m}\sum_{k=1}^m \frac{1}{4R}X_k^2 - \frac{2}{m}\sum_{k=1}^m C_1 \cdot (R^{-1} + \chi_{|X_k| \ge R^{\frac{1}{4}}}) \cdot Z_k^2,$$

where we have chosen $\phi \in C_c^\infty$ such that $0 \le \phi(x) \le 1$ for all $x$, $\phi(x) = 1$ for $|x| \le 1$ and $\phi(x) = 0$ for $|x| \ge 2$. Observe that for $a \sim \mathcal{N}(0, \mathrm{I_n})$, $Z \sim \mathcal{N}(0,1)$,

$$\mathbb{E}(a \cdot \hat{u})^4 \chi_{|a \cdot \hat{u}| \ge N} \le \mathbb{E} Z^4 \chi_{|Z| \ge N} \to 0, \qquad \text{as } N \to \infty.$$

It is also easy to show that

$$\inf_{\hat{u} \in \mathbb{S}^{n-1}} \mathbb{E}\frac{(a \cdot \hat{u})^4}{(\beta + (a \cdot e_1)^2)^2} \gtrsim 1.$$

Thus we can take $N$ large such that

$$\inf_{\hat{u} \in \mathbb{S}^{n-1}} \mathbb{E}\frac{(a \cdot \hat{u})^4}{(\beta + (a \cdot e_1)^2)^2}\phi(\frac{a \cdot \hat{u}}{N}) \gtrsim 1.$$

It is easy to show that by taking $R$ large, for $m \gtrsim n$, it holds with high probability that

$$\frac{1}{m}\sum_{k=1}^m \chi_{|X_k| \ge R^{\frac{1}{4}}} Z_k^2 \le \epsilon.$$

Since all the other terms are OK for union bounds, the desired result then clearly follows by taking $R$ large. $\square$

**Lemma 6.3** (The regime $\|u\|_2 \ll 1$ with $\frac{|u_1|}{\|u\|_2} \le \frac{1}{10}$ is OK)**.** *There exist a constant $R_2 = R_2(\beta) > 0$ such that the following hold:*

*For $m \gtrsim n$, with high probability it holds that*

$$\partial_{u_1 u_1} f \le -2 < 0, \quad \forall\, 0 < R \le R_2,\ \forall\, \hat{u} \in \mathbb{S}^{n-1} \textit{with } |\hat{u} \cdot e_1| \le \frac{1}{10}.$$

*Proof.* We only sketch the proof. Denote $X_k = a_k \cdot e_1$ and $Z_k = a_k \cdot \hat{u}$. A short computation gives

$$\begin{aligned}\partial_{u_1 u_1} f = {} & 4\frac{1}{m}\sum_{k=1}^m \frac{3R X_k^2 Z_k^2 - X_k^4}{\beta R + X_k^2} \\ & - 16\beta(\hat{u} \cdot e_1)R^2 \frac{1}{m}\sum_{k=1}^m \frac{Z_k^3 X_k}{(\beta R + X_k^2)^2} + 16\beta(\hat{u} \cdot e_1)R\frac{1}{m}\sum_{k=1}^m \frac{X_k^3 Z_k}{(\beta R + X_k^2)^2} \\ & + \frac{1}{m}\sum_{k=1}^m (R Z_k^2 - X_k^2)^2 \cdot \frac{6\beta^2 u_1^2 - 2\beta^2 |u'|^2 - 2\beta X_k^2}{(\beta R + X_k^2)^3},\end{aligned}$$

where $u_1 = u \cdot e_1$ and $u' = u - u_1 e_1$. Now observe that

$$\frac{1}{m}\sum_{k=1}^m \frac{Z_k^2}{\beta R + X_k^2} X_k^2 \le \frac{1}{m}\sum_{k=1}^m Z_k^2;$$

$$\frac{1}{m}\sum_{k=1}^m \frac{X_k^4}{\beta R + X_k^2} = \frac{1}{m}\sum_{k=1}^m \frac{(\beta R + X_k^2 - \beta R)^2}{\beta R + X_k^2} \ge \Big(\frac{1}{m}\sum_{k=1}^m (\beta R + X_k^2)\Big) - 2\beta R \ge -\beta R + \frac{1}{m}\sum_{k=1}^m X_k^2;$$

$$\frac{1}{m}\sum_{k=1}^m \frac{R^{\frac{3}{2}}|Z_k|^3 R^{\frac{1}{2}}|X_k|}{(\beta R + X_k^2)^2} \le \epsilon_1 \frac{1}{m}\sum_{k=1}^m \frac{R^2 Z_k^4}{(\beta R + X_k^2)^2} + \frac{1}{\epsilon_1^3}\frac{1}{m}\sum_{k=1}^m \frac{R^2 X_k^4}{(\beta R + X_k^2)^2} \le \frac{R^2}{\epsilon_1^3} + \epsilon_1 \frac{1}{m}\sum_{k=1}^m \frac{R^2 Z_k^4}{(\beta R + X_k^2)^2};$$

$$R\frac{1}{m}\sum_{k=1}^m \frac{|X_k|^3|Z_k|}{(\beta R + X_k^2)^2} \le R\frac{1}{m}\sum_{k=1}^m \frac{|X_k|^3|Z_k|}{(3(\beta R)^{\frac{1}{3}}(\frac{1}{4}X_k^4)^{\frac{1}{3}})^2} \lesssim R^{\frac{1}{3}}\beta^{-\frac{2}{3}}\frac{1}{m}\sum_{k=1}^m |X_k|^{\frac{1}{3}}|Z_k| \lesssim R^{\frac{1}{3}}\beta^{-\frac{2}{3}}\frac{1}{m}\sum_{k=1}^m (X_k^2 + Z_k^2 + 1),$$

where in the above the constant $\epsilon_1 > 0$ will be taken sufficiently small. The needed smallness will become clear momentarily.

Since $|u_1|/\|u\|_2 \le \frac{1}{10}$, it is clear that for some absolute constant $C_1 > 0$,

$$\frac{6\beta^2 u_1^2 - 2\beta^2 |u'|^2 - 2\beta X_k^2}{(\beta R + X_k^2)^3} \le -\beta C_1 \cdot \frac{1}{(\beta R + X_k^2)^2}.$$

Now

$$-\frac{1}{m}\sum_{k=1}^m \frac{(RZ_k^2 - X_k^2)^2}{(\beta R + X_k^2)^2} \le -\frac{1}{m}\sum_{k=1}^m \frac{R^2 Z_k^4}{(\beta R + X_k^2)^2} + 2R\frac{1}{m}\sum_{k=1}^m \frac{Z_k^2}{\beta R + X_k^2}.$$

Now take $\epsilon_1 = \frac{C_1}{1000}$. By Lemma F.1, we can take $R$ sufficiently small such that with high probability

$$2R\frac{1}{m}\sum_{k=1}^m \frac{Z_k^2}{\beta R + X_k^2} < \frac{1}{100}.$$

All the other terms can be treated by taking $R$ sufficiently small, and the desired result follows easily. □

**Lemma 6.4** (The regime $\|u\|_2 \ll 1$ with $\frac{|u_1|}{\|u\|_2} > \frac{1}{10}$ is OK)**.** *There exist a constant $R_3 = R_3(\beta) > 0$ such that the following hold:*

*For $m \gtrsim n$, with high probability it holds that the loss function $f = f(u)$ has no critical points in the regime*

$$\Big\{ u = \sqrt{R}\hat{u} : \quad 0 < R \le R_3,\ \hat{u} \in \mathbb{S}^{n-1} \text{ and } |\hat{u} \cdot e_1| > \frac{1}{10} \Big\}.$$

*Proof of Lemma 6.4.* We assume that for $0 < R \ll 1$ there exists some critical point. The idea is to examine the necessary conditions for a potential critical point and then derive a lower bound on $R$. Denote $X_k = a_k \cdot e_1$ and $Z_k = a_k \cdot \hat{u}$. By (6.2), we have $\partial_R f = 0$ which gives

$$\underbrace{R\frac{1}{m}\sum_{k=1}^m \frac{R\beta Z_k^4 + 2Z_k^4 X_k^2}{(\beta R + X_k^2)^2}}_{=:A_1} = \underbrace{\frac{1}{m}\sum_{k=1}^m \frac{\beta X_k^4 + 2X_k^4 Z_k^2}{(\beta R + X_k^2)^2}}_{=:B_1}.$$

On the other hand, by using $\partial_{u_1} f(u) = 0$, we obtain

$$\beta u_1 R^2 \frac{1}{m}\sum_{k=1}^m \frac{Z_k^4}{(\beta R + X_k^2)^2} - 2R^{\frac{3}{2}}\frac{1}{m}\sum_{k=1}^m \frac{Z_k^3 X_k}{\beta R + X_k^2} = \beta u_1 \frac{1}{m}\sum_{k=1}^m \frac{2RZ_k^2X_k^2 - X_k^4}{(\beta R + X_k^2)^2} - 2R^{\frac{1}{2}}\frac{1}{m}\sum_{k=1}^m \frac{Z_k X_k^3}{\beta R + X_k^2}.$$

Thus

$$-R\underbrace{\widehat{u}_1\frac{1}{m}\sum_{k=1}^m\frac{\beta R Z_k^4}{(\beta R+X_k^2)^2}}_{=:A_2}+R\cdot\underbrace{2\frac{1}{m}\sum_{k=1}^m\frac{Z_k^3X_k}{\beta R+X_k^2}}_{=:A_3}=\underbrace{\beta\widehat{u}_1\frac{1}{m}\sum_{k=1}^m\frac{-2RZ_k^2X_k^2+X_k^4}{(\beta R+X_k^2)^2}+2\frac{1}{m}\sum_{k=1}^m\frac{Z_kX_k^3}{\beta R+X_k^2}}_{=:B_2}.$$

Without loss of generality we assume $\widehat{u}_1>\frac{1}{10}$. Observe that for $0<R\le 1$, we have

$$\frac{1}{m}\sum_{k=1}^m\frac{X_k^4}{(\beta+X_k^2)^2}\lesssim B_1\lesssim 1+\frac{1}{m}\sum_{k=1}^m Z_k^2.$$

Thus with high probability $B_1\sim 1$.

Now by Lemma F.1, for $0<R\ll 1$, we have

$$\frac{1}{m}\sum_{k=1}^m\frac{RZ_k^2X_k^2}{(\beta R+X_k^2)^2}\le\frac{1}{m}\sum_{k=1}^m\frac{RZ_k^2}{\beta R+X_k^2}\ll 1.$$

Also for $0<R\le 1$, we have

$$\frac{1}{m}\sum_{k=1}^m\frac{X_k^4}{(\beta+X_k^2)^2}\le\frac{1}{m}\sum_{k=1}^m\frac{X_k^4}{(\beta R+X_k^2)^2}\le 1.$$

By Lemma F.2, for $0<R\ll 1$, we have

$$c_1\le\frac{1}{m}\sum_{k=1}^m\frac{Z_kX_k^3}{\beta R+X_k^2}\le c_2,$$

where $c_1$, $c_2>0$ are constants depending only on $\beta$.

Thus with high probability we have for $0<R\ll 1$, $B_2\sim 1$.

Now since

$$RA_1=B_1,\qquad -RA_2+RA_3=B_2;$$

we obtain

$$A_1+B_3A_2=B_3A_3,$$

where $B_3=B_1/B_2$. Observe that $A_1>0$, $A_2>0$, and

$$A_1\sim\frac{1}{m}\sum_{k=1}^m\frac{Z_k^4}{\beta R+X_k^2};$$
$$A_3\le\Big(\frac{1}{m}\sum_{k=1}^m\frac{Z_k^4}{\beta R+X_k^2}\Big)^{\frac{3}{4}}\Big(\frac{1}{m}\sum_{k=1}^m\frac{X_k^4}{\beta R+X_k^2}\Big)^{\frac{1}{4}}.$$

It follows easily that with high probability we have

$$A_1\sim 1.$$

But then it follows from the equation $RA_1=B_1$ that we must have $R\sim 1$. Thus the desired result follows. □

**Theorem 6.2** (The regimes $\|u\|_2\ll 1$ and $\|u\|_2\gg 1$ are OK)**.** *For $m\gtrsim n$, with high probability the following hold:*

*(1) We have*

$$\partial_R f \geq d_1, \qquad \forall\, R \geq R_1,\, \forall\, \hat{u} \in \mathbb{S}^{n-1},$$

*where $d_1$, $R_1$ are constants depending only on $\beta$.*

*(2) We have*

$$\partial_{u_1 u_1} f \leq -2 < 0, \quad \forall\, 0 < R \leq R_2,\, \forall\, \hat{u} \in \mathbb{S}^{n-1} \text{with } |\hat{u} \cdot e_1| \leq \frac{1}{10},$$

*where $R_2 > 0$ is a constant depending only on $\beta$.*

*(3) The loss function $f = f(u)$ has no critical points in the regime*

$$\Big\{ u = \sqrt{R}\hat{u} : \quad 0 < R \leq R_3,\, \hat{u} \in \mathbb{S}^{n-1} \text{ and } |\hat{u} \cdot e_1| > \frac{1}{10} \Big\},$$

*where $R_3 > 0$ is a constant depending only on $\beta$.*

*(4) The point $u = 0$ is a local maximum point with strictly negative-definite Hessian,*

$$\sum_{k,l=1}^{n} \xi_k \xi_l (\partial_{kl} f)(0) \leq -d_2 < 0, \qquad \forall\, \xi \in \mathbb{S}^{n-1},$$

*where $d_2 > 0$ is an absolute constant.*

*Proof of Theorem 6.2.* This follows from Lemma 6.1, 6.2, 6.3 and 6.4. □

6.2. **The regime $\|u\|_2 \sim 1$.**

**Lemma 6.5** (The regime $\|u\|_2 \sim 1$ with $\epsilon_0 \leq |\hat{u} \cdot e_1| \leq 1 - \epsilon_0$ is OK)**.** *Let $0 < \epsilon_0 \ll 1$ be given. Assume $0 < c_1 < c_2 < \infty$ are two given constants. Then for $m \gtrsim n$, the following hold with high probability:*

*The loss function $f = f(u)$ has no critical points in the regime:*

$$\Big\{ u = \sqrt{R}\hat{u} : \ c_1 < R < c_2,\ \epsilon_0 \leq |\hat{u} \cdot e_1| \leq 1 - \epsilon_0 \Big\}.$$

*More precisely, introduce the parametrization $\hat{u} = e_1 \cos\theta + e^{\perp} \sin\theta$, where $\theta \in [0, \pi]$ and $e^{\perp} \in \mathbb{S}^{n-1}$ satisfies $e^{\perp} \cdot e_1 = 0$. Then in the aforementioned regime, we have*

$$|\partial_\theta f| \geq \alpha_1 > 0,$$

*where $\alpha_1$ depends only on $(\beta, \epsilon_0, c_1, c_2)$.*

*Proof.* See appendix. □

**Lemma 6.6** (The regime $\|u\|_2 \sim 1$ with $|\hat{u} \cdot e_1| \leq \epsilon_1$ is OK)**.** *Let $0 < \epsilon_1 \ll 1$ be a sufficiently small constant. Assume $0 < c_1 < c_2 < \infty$ are two given constants. Then for $m \gtrsim n$, the following hold with high probability:*

*Consider the regime*

$$\Big\{ u = \sqrt{R}\hat{u} : \ c_1 < R < c_2,\ |\hat{u} \cdot e_1| \leq \epsilon_1 \Big\}.$$

*Introduce the parametrization $\hat{u} = e_1 \cos\theta + e^{\perp} \sin\theta$, where $\theta \in [0, \pi]$ and $e^{\perp} \in \mathbb{S}^{n-1}$ satisfies $e^{\perp} \cdot e_1 = 0$. Then in the aforementioned regime, we have*

$$\partial_{\theta\theta} f \leq -\alpha_2 < 0,$$

*where $\alpha_2 > 0$ depends only on $(\beta, \epsilon_1, c_1, c_2)$.*

*Proof.* See appendix. □

**Theorem 6.3** (The regime $\|u\|_2 \sim 1$, $||\hat{u} \cdot e_1| - 1| \leq \epsilon_0$, $|\|u\|_2 - 1| \geq c(\epsilon_0)$ is OK)**.** *Let $0 < R_1 < 1 < R_2 < \infty$ be given constants. Let $0 < \epsilon_0 \ll 1$ be a given sufficiently small constant and consider the regime $\Big||\hat{u} \cdot e_1| - 1\Big| \leq \epsilon_0$ with $R_1 \leq \|u\|_2^2 \leq R_2$. There exists a constant $c_0 = c_0(\epsilon_0, R_1, R_2, \beta) > 0$ which tends to zero as $\epsilon_0 \to 0$ such that the following hold:*

*For $m \gtrsim n$, with high probability it holds that (below $u = \sqrt{R}\hat{u}$)*

$$\partial_R f < 0, \quad \forall\, R_2 \leq R \leq 1 - c_0, \quad \forall\, \hat{u} \in \mathbb{S}^{n-1} \text{with } ||\hat{u} \cdot e_1| - 1| \leq \epsilon_0;$$
$$\partial_R f > 0, \quad \forall\, 1 + c_0 \leq R \leq R_1, \quad \forall\, \hat{u} \in \mathbb{S}^{n-1} \text{with } ||\hat{u} \cdot e_1| - 1| \leq \epsilon_0.$$

*Proof of Theorem 6.3.* We first consider the regime $R \geq 1 + c$. Let $\phi \in C_c^\infty(\mathbb{R})$ be an even function satisfying $0 \leq \phi(x) \leq 1$ for all $x$, $\phi(x) = 1$ for $|x| \leq 1$ and $\phi(x) = 0$ for $|x| > 2$.

By using (6.2), we have

$$\begin{aligned}
&\partial_R f \\
\geq &\frac{1}{m} \sum_{k=1}^m \frac{R^2 \beta (a_k \cdot \hat{u})^4 \phi(\frac{a_k \cdot \hat{u}}{K}) + 2R (a_k \cdot \hat{u})^4 \phi(\frac{a_k \cdot \hat{u}}{K})(a_k \cdot e_1)^2 - \beta (a_k \cdot e_1)^4 - 2(a_k \cdot e_1)^4 (a_k \cdot \hat{u})^2}{(\beta R + (a_k \cdot e_1)^2)^2}.
\end{aligned} \tag{6.4}$$

By taking $K$ sufficiently large, we can easily obtain

$$\mathbb{E}(1 - \phi(\frac{a \cdot e_1}{K}))(1 + (a \cdot e_1)^2) \ll 1,$$

where $a \sim \mathcal{N}(0, \mathrm{I_n})$. For fixed $K$, it is not difficult to check that the lower bound (6.4) are OK for union bounds and they can be made close to the expectation with high probability, uniformly in $R \sim 1$ and $\hat{u} \in \mathbb{S}^{n-1}$. The perturbation argument (i.e. estimating the error terms coming from replacing $a_k \cdot \hat{u}$ by $a_k \cdot e_1$ and so on) becomes rather easy after taking the expectation. It is then not difficult to show that

$$\partial_R f > 0,$$

for $R \geq 1 + c(\epsilon_0)$.

Next we turn to the regime $R_2 \leq R \leq 1 - c(\epsilon_0)$. Without loss of generality we may assume $|1 - \hat{u} \cdot e_1| \leq \epsilon_0$. The idea is to exploit the decomposition used in the proof of Lemma 6.4.

Namely using $\partial_R f=0$ and $\partial_{u_1} f=0$, we have

$$R\underbrace{\frac{1}{m}\sum_{k=1}^m \frac{R\beta Z_k^4+2Z_k^4X_k^2}{(\beta R+X_k^2)^2}}_{=:A_1}=\underbrace{\frac{1}{m}\sum_{k=1}^m \frac{\beta X_k^4+2X_k^4Z_k^2}{(\beta R+X_k^2)^2}}_{=:B_1};$$

$$-R\,\widehat{u}_1\underbrace{\frac{1}{m}\sum_{k=1}^m \frac{\beta R Z_k^4}{(\beta R+X_k^2)^2}}_{=:A_2}+R\cdot 2\underbrace{\frac{1}{m}\sum_{k=1}^m \frac{Z_k^3X_k}{\beta R+X_k^2}}_{=:A_3}=\underbrace{\beta\widehat{u}_1\frac{1}{m}\sum_{k=1}^m \frac{-2RZ_k^2X_k^2+X_k^4}{(\beta R+X_k^2)^2}+2\frac{1}{m}\sum_{k=1}^m \frac{Z_kX_k^3}{\beta R+X_k^2}}_{=:B_2}.$$

It is not difficult to check that with high probability, we have $B_1\sim 1$, $B_2\sim 1$, and

$$\Big|\frac{B_2}{B_1}-1\Big|\le \eta(\epsilon_0)\ll 1,\qquad \forall\, R_2\le R\le 1,\ \forall\,\hat{u}\in\mathbb{S}^{n-1}\text{ with }|\hat{u}\cdot e_1-1|\le\epsilon_0,$$

where $\eta(\epsilon_0)\to 0$ as $\epsilon_0\to 0$. We then obtain

$$A_1=\Big(1+O(\eta(\epsilon_0))\Big)(-A_2+A_3).$$

From this it is easy (similar to an argument used in the proof of Lemma 6.4) to derive that

$$A_1+A_2+|A_3|\lesssim 1.$$

Now note that the pre-factor of $A_2$ is $\hat{u}_1=1+O(\epsilon_0)$. By using the relation

$$A_1+A_2-A_3=O(\eta(\epsilon_0)),$$

we obtain

$$\frac{1}{m}\sum_{k=1}^m \frac{Z_k^4}{\beta R+X_k^2}-\frac{1}{m}\sum_{k=1}^m \frac{Z_k^3X_k}{\beta R+X_k^2}=O(\eta(\epsilon_0)).$$

By using localization (i.e. decomposing $Z_k^3X_k=Z_k^3\phi(\frac{Z_k}{M})X_k+Z_k^3(1-\phi(\frac{Z_k}{M}))X_k$), Hölder and taking $M$ sufficiently large, one can then derive that (with high probability)

$$\Big|\frac{1}{m}\sum_{k=1}^m \frac{Z_k^4-X_k^4}{\beta R+X_k^2}\Big|+\Big|\frac{1}{m}\sum_{k=1}^m \frac{Z_k^3X_k-X_k^4}{\beta R+X_k^2}\Big|=O(\eta_1(\epsilon_0)),\qquad \forall\, R_2\le R\le 1,\ \forall\,\hat{u}\in\mathbb{S}^{n-1}\text{ with }|\hat{u}\cdot e_1-1|\le\epsilon_0,$$

where $\eta_1(\epsilon_0)\to 0$ as $\epsilon_0\to 0$. It then follows easily that (with high probability)

$$\Big|\frac{1}{m}\sum_{k=1}^m \frac{(Z_k-X_k)^4}{\beta R+X_k^2}\Big|=O(\eta_2(\epsilon_0)),\qquad \forall\, R_2\le R\le 1,\ \forall\,\hat{u}\in\mathbb{S}^{n-1}\text{ with }|\hat{u}\cdot e_1-1|\le\epsilon_0,$$

where $\eta_2(\epsilon_0)\to 0$ as $\epsilon_0\to 0$.

Now observe that for $A_1$, we have

$$\begin{aligned}|Z_k^4-X_k^4|&\le |Z_k-X_k|(O(|Z_k|^3)+O(|X_k|^3))\\&\le C_\epsilon|Z_k-X_k|^4+\epsilon\cdot(O(|Z_k|^4)+O(X_k^4)),\end{aligned}$$

where $C_\epsilon>0$ depends only on $\epsilon$. Clearly by taking $\epsilon>0$ sufficiently small and using the derived quantitative estimates preceding this paragraph, we can guarantee that (with high probability)

$$\Big|A_1-B_1\Big|\ll 1,\qquad \forall\, R_2\le R\le 1,\ \forall\,\hat{u}\in\mathbb{S}^{n-1}\text{ with }|\hat{u}\cdot e_1-1|\le\epsilon_0.$$

It follows that we must have $|R-1|\ll 1$ for a potential critical point. By using (6.2) we have $\partial_R f(R=0)<0$. By using (6.3) we have $\partial_{RR}f>0$. Since we have shown $\partial_R f>0$ for

$R > 1 + c(\epsilon_0)$, it then follows that $\partial_R f = 0$ occurs at a unique point $|R-1| \ll 1$ and $\partial_R f < 0$ for $R < 1 - c(\epsilon_0)$ provided $c(\epsilon_0)$ is suitably re-defined. $\square$

We now show restrictive convexity of the loss function $f(u)$ near the global minimizer $u = \pm e_1$.

**Theorem 6.4** (Restrictive convexity near the global minimizer)**.** *There exists $0 < \epsilon_0 \ll 1$ sufficiently small such that if $m \gtrsim n$, then the following hold with high probability:*

*(1) If $\|u - e_1\|_2 \le \epsilon_0$ and $u \ne e_1$, then for $\xi = \frac{u-e_1}{\|u-e_1\|_2} \in \mathbb{S}^{n-1}$, we have*

$$\sum_{i,j=1}^{n} \xi_i \xi_j (\partial_{ij} f)(u) \ge \gamma > 0,$$

*where $\gamma$ is a constant depending only on $\beta$.*

*(2) If $\|u + e_1\|_2 \le \epsilon_0$, then then for $\xi = \frac{u+e_1}{\|u+e_1\|_2} \in \mathbb{S}^{n-1}$, we have*

$$\sum_{i,j=1}^{n} \xi_i \xi_j (\partial_{ij} f)(u) \ge \gamma > 0,$$

*where $\gamma$ is a constant depending only on $\beta$.*

*(3) Alternatively we can use the parametrization $u = \pm e_1 + t\xi$, where $\xi \in \mathbb{S}^{n-1}$, and $|t| \le \epsilon_0$. Then with this special parametrization, we have*

$$\sum_{i,j=1}^{n} \xi_i \xi_j (\partial_{ij} f)(u) \ge \gamma > 0.$$

*Note that this includes the global minimizers $u = \pm e_1$.*

*In yet other words, $f(u)$ is restrictively convex in a sufficiently small neighborhood of $\pm e_1$.*

*Proof.* See appendix. $\square$

*Proof of Theorem 6.1.* We proceed in several steps as follows.

(1) For the regime $\|u\|_2 \ll 1$ and $\|u\|_2 \gg 1$, we use Theorem 6.2. The point $u = 0$ is a local maximum point with strictly negative-definite Hessian. All other possible critical points must have negative curvature direction.

(2) For the regime $\|u\|_2 \sim 1$, $||\hat{u} \cdot e_1| - 1| \ge \epsilon_0$, we use Lemma 6.5 and 6.6. The loss function either has a nonzero gradient, or it is a strict saddle with a negative curvature direction.

(3) For the regime $\|u\|_2 \sim 1$, $|\hat{u} \cdot e_1| - 1| \le \epsilon_0$, $|\|u\|_2 - 1| \ge c(\epsilon_0)$, we apply Theorem 6.3. The loss function has nonzero gradient in this regime.

(4) Finally for the regime close to the global minimizers $\pm e_1$, we use Theorem 6.4 to show restrictive convexity. This ensures that $\pm e_1$ are the only minimizers.

$\square$

Consider for $\beta_1 > 0$, $\beta_2 > 0$,

$$f(u) = \frac{1}{m}\sum_{k=1}^{m} \frac{((a_k\cdot u)^2 - (a_k\cdot x)^2)^2}{|u|^2 + \beta_1 (a_k\cdot u)^2 + \beta_2 (a_k \cdot x)^2}. \tag{7.1}$$

We shall adopt similar notation as in Section 3.

**Theorem 7.1.** *Let $0 < \beta_1, \beta_2 < \infty$. Assume $\{a_k\}_{k=1}^m$ are i.i.d. standard Gaussian random vectors and $x \neq 0$. There exist positive constants $c$, $C$ depending only on $(\beta_1, \beta_2)$, such that if $m \ge Cn$, then with probability at least $1 - e^{-cm}$ the loss function $f = f(u)$ defined by (7.1) has no spurious local minimizers. The only global minimizer is $\pm x$, and the loss function is strongly convex in a neighborhood of $\pm x$. The point $u = 0$ is a local maximum point with strictly negative-definite Hessian. All other critical points are strict saddles, i.e., each saddle point has a neighborhood where the function has negative directional curvature.*

Without loss of generality we shall assume $x = e_1$ throughout the rest of the proof. Thus we consider

$$f(u) = \frac{1}{m}\sum_{k=1}^{m} \frac{((a_k\cdot u)^2 - (a_k\cdot e_1)^2)^2}{|u|^2 + \beta_1 (a_k\cdot u)^2 + \beta_2 (a_k \cdot e_1)^2}. \tag{7.2}$$

## 7.1. The regimes $\|u\|_2 \ll 1$ and $\|u\|_2 \gg 1$ are fine.

We first investigate the point $u = 0$. It is trivial to verify that $\nabla f(0) = 0$ since $a_k \cdot e_1 \neq 0$ for all $k$ almost surely.

**Lemma 7.1** ($u = 0$ has strictly negative-definite Hessian)**.** *We have $u = 0$ is local maximum point with strictly negative-definite Hessian. More precisely, it holds (almost surely) that*

$$\sum_{k,l=1}^{n} \xi_k \xi_l (\partial_{kl} f)(0) \le -d_1, \quad \forall\, \xi \in \mathbb{S}^{n-1},$$

*where $d_1 > 0$ is a constant depending only on $\beta_2$.*

*Proof of Lemma 7.1.* We begin by noting that since almost surely $a_k \cdot e_1 \neq 0$ for all $k$, the function $f$ is smooth at $u = 0$. It suffices for us to consider (write $u = \sqrt{t}\xi$)

$$G(t) = \frac{1}{m}\sum_{k=1}^{m} \frac{(t(a_k\cdot \xi)^2 - (a_k\cdot e_1)^2)^2}{t + t\beta_1 (a_k\cdot \xi)^2 + \beta_2 (a_k \cdot e_1)^2}.$$

By a simple computation, we have

$$G'(0) = -\frac{1}{\beta_2^2} - \frac{\beta_1 + 2\beta_2}{\beta_2^2}\cdot \frac{1}{m}\sum_{k=1}^{m} (a_k\cdot\xi)^2.$$

The desired conclusion then easily follows. ∎

Write $u = \sqrt{R}\hat{u}$ where $\hat{u} \in S^{n-1}$ and $R > 0$. Denote $X_k = a_k \cdot e_1$. Then

$$f(u) = \frac{1}{m}\sum_{k=1}^{m} \frac{\Big(R(a_k\cdot \hat{u})^2 - X_k^2\Big)^2}{R + \beta_1 R (a_k\cdot \hat{u})^2 + \beta_2 X_k^2}.$$

Clearly

$$\partial_R f = \frac{1}{m}\sum_{k=1}^m \frac{R^2((a_k\cdot\hat{u})^4+\beta_1(a_k\cdot\hat{u})^6)+2R\beta_2(a_k\cdot\hat{u})^4X_k^2-X_k^4-(\beta_1+2\beta_2)(a_k\cdot\hat{u})^2X_k^4}{(R+R\beta_1(a_k\cdot\hat{u})^2+\beta_2X_k^2)^2}; \tag{7.3}$$

$$\partial_{RR} f = 2\frac{1}{m}\sum_{k=1}^m \frac{\Big(1+(a_k\cdot\hat{u})^2(\beta_1+\beta_2)\Big)X_k^4}{(R+R\beta_1(a_k\cdot\hat{u})^2+\beta_2X_k^2)^3}. \tag{7.4}$$

**Lemma 7.2** (The regimes $\|u\|_2 \gg 1$ or $\|u\|_2 \ll 1$ are OK)**.** *There exist constants $R_i = R_i(\beta_1,\beta_2)>0$, $d_i=d_i(\beta_1,\beta_2)>0$, $i=1,2$ such that the following hold:*

*For $m\gtrsim n$, with high probability it holds that*

$$\partial_R f \ge d_1, \quad \forall\, R\ge R_1,\ \forall\, \hat{u}\in\mathbb{S}^{n-1};$$
$$\partial_R f \le -d_2<0, \quad \forall\, 0<R\le R_2,\ \forall\, \hat{u}\in\mathbb{S}^{n-1}.$$

*Proof.* Denote $Z_k = a_k\cdot\hat{u}$. We first consider the regime $R\gg 1$. Observe that

$$\frac{1}{m}\sum_{k=1}^m \frac{R^2Z_k^4}{(R+R\beta_1Z_k^2+\beta_2X_k^2)^2} \gtrsim \frac{1}{m}\sum_{k=1}^m \frac{Z_k^4}{(1+Z_k^2+X_k^2)^2} \gtrsim 1, \quad \forall\, \hat{u}\in\mathbb{S}^{n-1},$$

where the last inequality holds for $m\gtrsim n$ with high probability.

On the other hand we note that

$$\begin{aligned}\frac{1}{m}\sum_{k=1}^m \frac{X_k^4}{(R+R\beta_1Z_k^2+\beta_2X_k^2)^2} &\lesssim \frac{1}{m}\sum_{k=1}^m \frac{X_k^4}{(R+X_k^2)^2}(\chi_{|X_k|\le R^{\frac14}}+\chi_{|X_k|>R^{\frac14}}) \\ &\lesssim R^{-1}+\frac{1}{m}\sum_{k=1}^m \chi_{|X_k|>R^{\frac14}} \ll 1, \qquad \forall\, \hat{u}\in\mathbb{S}^{n-1},\end{aligned}$$

where again the last inequality holds for $R$ sufficiently large, and for $m\gtrsim n$ with high probability.

Similarly we have for $R$ sufficiently large,

$$\frac{1}{m}\sum_{k=1}^m \frac{(\beta_1+2\beta_2)Z_k^2X_k^4}{(R+R\beta_1Z_k^2+\beta_2X_k^2)^2} \lesssim R^{-1}\frac{1}{m}\sum_{k=1}^m Z_k^2 + \frac{1}{m}\sum_{k=1}^m Z_k^2\chi_{|X_k|>R^{\frac14}} \ll 1, \qquad \forall\, \hat{u}\in\mathbb{S}^{n-1}.$$

Thus it follows easily that $\partial_R f\gtrsim 1$ for $R\gg 1$.

Now we turn to the regime $0<R\ll 1$.

First we note that the main negative term is OK. This is due to the fact that for $0<R\le 1$, we have (for $m\gtrsim n$ and with high probability)

$$\frac{1}{m}\sum_{k=1}^m \frac{Z_k^2X_k^4}{(R+RZ_k^2+X_k^2)^2} \ge \frac{1}{m}\sum_{k=1}^m \frac{Z_k^2X_k^4}{(1+Z_k^2+X_k^2)^2} \gtrsim 1, \qquad \forall\, \hat{u}\in\mathbb{S}^{n-1}.$$

On the other hand, we have (for $m \gtrsim n$ and with high probability)

$$\begin{aligned}\frac{1}{m}\sum_{k=1}^m \frac{R^2(Z_k^4+Z_k^6)}{(R+RZ_k^2+X_k^2)^2} &\le \frac{1}{m}\sum_{k=1}^m \frac{RZ_k^4}{R+RZ_k^2+X_k^2}\cdot(\chi_{|X_k|\ge R^{\frac14}|Z_k|}+\chi_{|X_k|<R^{\frac14}|Z_k|})\\ &\le R^{\frac12}\frac{1}{m}\sum_{k=1}^m Z_k^2+\frac{1}{m}\sum_{k=1}^m Z_k^2\chi_{|X_k|<R^{\frac14}|Z_k|}\\ &\le R^{\frac12}\frac{1}{m}\sum_{k=1}^m Z_k^2+\frac{1}{m}\sum_{k=1}^m Z_k^2\chi_{|Z_k|\ge K}+\frac{1}{m}\sum_{k=1}^m K^2\chi_{|X_k|<KR^{\frac14}}\\ &\ll 1, \qquad \forall\, \hat{u}\in\mathbb{S}^{n-1},\end{aligned}$$

if we first take $K$ sufficiently large followed by taking $R$ sufficiently small.

The estimate of the other term $\frac{RZ_k^4X_k^2}{(R+RZ_k^2+X_k^2)^2}$ is similar and we omit further details.

Collecting the estimates, it is then clear that we can obtain the desired estimate for $\partial_R f$ when $0<R\ll 1$. $\square$

7.2. **The regime $\|u\|_2\sim 1$.**

**Lemma 7.3** (The regime $\|u\|_2\sim 1$ with $\epsilon_0\le|\hat{u}\cdot e_1|\le 1-\epsilon_0$ is OK)**.** *Let $0<\epsilon_0\ll 1$ be given. Assume $0<c_1<c_2<\infty$ are two given constants. Then for $m\gtrsim n$, the following hold with high probability:*

*The loss function $f=f(u)$ has no critical points in the regime:*

$$\Big\{u=\sqrt{R}\hat{u}:\ c_1<R<c_2,\ \epsilon_0\le|\hat{u}\cdot e_1|\le 1-\epsilon_0\Big\}.$$

*More precisely, introduce the parametrization $\hat{u}=e_1\cos\theta+e^{\perp}\sin\theta$, where $\theta\in[0,\pi]$ and $e^{\perp}\in\mathbb{S}^{n-1}$ satisfies $e^{\perp}\cdot e_1=0$. Then in the aforementioned regime, we have*

$$|\partial_\theta f|\ge\alpha_1>0,$$

*where $\alpha_1$ depends only on $(\beta,\ \epsilon_0,\ c_1,\ c_2)$.*

*Proof.* We first recall

$$f(u)=\frac{1}{m}\sum_{k=1}^m\frac{\Big(R(a_k\cdot\hat{u})^2-X_k^2\Big)^2}{R+\beta_1R(a_k\cdot\hat{u})^2+\beta_2X_k^2}.$$

Clearly $a_k\cdot\hat{u}=X_k\cos\theta+(a_k\cdot e^{\perp})\sin\theta$, and

$$\begin{aligned}&\partial_\theta(a_k\cdot\hat{u})=X_k(-\sin\theta)+(a_k\cdot e^{\perp})\cos\theta;\\ &\partial_{\theta\theta}(a_k\cdot\hat{u})=-(a_k\cdot\hat{u}).\end{aligned}$$

In particular, if $\theta$ is away from the end-points $0,\pi$, then

$$\partial_\theta(a_k\cdot\hat{u})=(a_k\cdot\hat{u})\cot\theta-X_k\csc\theta.$$

We then obtain (below $Z_k = a_k \cdot \hat{u}$)

$$\begin{aligned}\partial_\theta f =& -\csc\theta \frac{1}{m}\sum_{k=1}^m \frac{2RZ_k(-X_k^2+RZ_k^2)\cdot\Big((\beta_1+2\beta_2)X_k^2+R(2+\beta_1 Z_k^2)\Big)X_k}{(R+\beta_1 RZ_k^2+\beta_2 X_k^2)^2}\\ &+\cot\theta\frac{1}{m}\sum_{k=1}^m \frac{2RZ_k(-X_k^2+RZ_k^2)\cdot\Big((\beta_1+2\beta_2)X_k^2+R(2+\beta_1 Z_k^2)\Big)Z_k}{(R+\beta_1 RZ_k^2+\beta_2 X_k^2)^2}.\end{aligned}$$

Thanks to the strong damping, it is not difficult to check that for any $\epsilon > 0$, if $m \gtrsim n$, then with high probability we have

$$|\partial_\theta f - \mathbb{E}\partial_\theta f| \le \epsilon, \qquad \forall\, c_1 \le R \le c_2,\, \forall\, \hat{u} \in \mathbb{S}^{n-1}.$$

The desired result then follows from Lemma G.1. □

**Lemma 7.4** (The regime $\|u\|_2 \sim 1$ with $|\hat{u}\cdot e_1| \le \epsilon_0$ is OK)**.** *Let $0 < \epsilon_1 \ll 1$ be a sufficiently small constant. Assume $0 < c_1 < c_2 < \infty$ are two given constants. Then for $m \gtrsim n$, the following hold with high probability:*

*Consider the regime*

$$\Big\{u = \sqrt{R}\hat{u}:\; c_1 < R < c_2,\; |\hat{u}\cdot e_1| \le \epsilon_1\Big\}.$$

*Introduce the parametrization $\hat{u} = e_1\cos\theta + e^\perp \sin\theta$, where $\theta \in [0,\pi]$ and $e^\perp \in \mathbb{S}^{n-1}$ satisfies $e^\perp \cdot e_1 = 0$. Then in the aforementioned regime, we have*

$$\partial_{\theta\theta} f \le -\alpha_2 < 0,$$

*where $\alpha_2 > 0$ depends only on ($\beta$, $\epsilon_1$, $c_1$, $c_2$).*

*Proof.* This is similar to the argument in the proof of Lemma 7.3. By a tedious computation, we have

$$\partial_{\theta\theta} f = \frac{1}{m}\sum_{k=1}^m \frac{2RG_k}{\left(R+\beta_2 x^2+\beta_1 RZ_k^2\right)^3},$$

where

$$\begin{aligned}G_k =& -8\beta_1 RZ_k^2\left(-X_k^2+RZ_k^2\right)\left(R+\beta_2X_k^2+\beta_1RZ_k^2\right)(X_k-Z_k\cos\theta)^2\csc^2\theta\\ &-2\left(R+\beta_2X_k^2+\beta_1RZ_k^2\right)^2\Big(X_k^4-3RX_k^2Z_k^2-RZ_k^4-2X_kZ_k\left(X_k^2-3RZ_k^2\right)\cos\theta\\ &\qquad+Z_k^2(X_k^2-2RZ_k^2)\cos 2\theta\Big)\csc^2\theta\\ &+\beta_1\left(X_k^2-RZ_k^2\right)^2\Big(Z_k^2\left(R+\beta_2X_k^2+\beta_1RZ_k^2\right)+4\beta_1RZ_k^2(X_k-Z_k\cos\theta)^2\csc^2\theta\\ &\quad-\left(R+\beta_2X_k^2+\beta_1RZ_k^2\right)(X_k-Z_k\cos\theta)^2\csc^2\theta\Big).\end{aligned}$$

It is then tedious but not difficult to check that that for any $\epsilon > 0$, if $m \gtrsim n$, then with high probability we have

$$|\partial_{\theta\theta} f - \mathbb{E}\partial_{\theta\theta} f| \le \epsilon, \qquad \forall\, c_1 \le R \le c_2,\, \forall\, \hat{u} \in \mathbb{S}^{n-1}.$$

The desired result then follows from Lemma G.1. □

**Theorem 7.2** (The regime $\|u\|_2 \sim 1$, $||\hat{u}\cdot e_1|-1| \le \epsilon_0$, $|\|u\|_2 - 1| \ge c(\epsilon_0)$ is OK)**.** *Let $0 < c_1 < 1 < c_2 < \infty$ be given constants. Let $0 < \epsilon_0 \ll 1$ be a given sufficiently small constant and consider the regime $\Big||\hat{u}\cdot e_1| - 1\Big| \le \epsilon_0$ with $c_1 \le \|u\|_2^2 \le c_2$. There exists a constant $c_0 = c_0(\epsilon_0, c_1, c_2, \beta) > 0$ which tends to zero as $\epsilon_0 \to 0$ such that the following hold:*

*For $m \gtrsim n$, with high probability it holds that (below $u = \sqrt{R}\hat{u}$)*

$$\partial_R f < 0, \quad \forall\, c_2 \le R \le 1 - c_0, \quad \forall\, \hat{u} \in \mathbb{S}^{n-1} \text{with } ||\hat{u}\cdot e_1| - 1| \le \epsilon_0;$$
$$\partial_R f > 0, \quad \forall\, 1 + c_0 \le R \le c_1, \quad \forall\, \hat{u} \in \mathbb{S}^{n-1} \text{with } ||\hat{u}\cdot e_1| - 1| \le \epsilon_0.$$

*Proof.* We rewrite

$$f(u) = \frac{1}{m}\sum_{k=1}^m g(R, (a_k\cdot \hat{u})^2, X_k^2),$$

where

$$g(R,a,b) = \frac{(Ra-b)^2}{R + \beta_1 Ra + \beta_2 b}.$$

It is not difficult to check that for $R \sim 1$, we have

$$|(\partial_R g)(R,a,b) - (\partial_R g)(R,b,b)| \le \|\partial_{Ra} g\|_\infty |b-a| \lesssim |b-a|, \qquad \forall\, a,b \ge 0.$$

On the other hand, note that $(\partial_R g)(1,b,b) = 0$, and for $R \sim 1$,

$$(\partial_{RR} g)(R,b,b) = \frac{2b^2(1+b(\beta_1+\beta_2))^2}{(R+b(\beta_2+\beta_1 R))^3} \sim b.$$

Thus for $R = 1+\eta$, $\eta > 0$ we have

$$\begin{aligned}(\partial_R g)(R,a,b) &\ge \partial_R g(R,b,b) - \gamma_1|b-a| \\ &\ge \gamma_2\cdot\eta\cdot b - \gamma_1|b-a|,\end{aligned}$$

where $\gamma_1 > 0$, $\gamma_2 > 0$ are constants depending only on ($\beta_1$, $\beta_2$, $c_1$, $c_2$). The desired result (for $\partial_R f > 0$ when $R \to 1+$) then follows from this and simple application of Bernstein's inequalities. The estimate for the regime $R \to 1-$ is similar. We omit the details. □

**Theorem 7.3** (Strong convexity near the global minimizer)**.** *There exists $0 < \epsilon_0 \ll 1$ such that if $m \gtrsim n$, then the following hold with high probability:*

*(1) If $\|u - e_1\|_2 \le \epsilon_0$, then*

$$\sum_{i,j=1}^n \xi_i\xi_j(\partial_{ij} f)(u) \ge \gamma > 0, \qquad \forall\, \xi \in \mathbb{S}^{n-1},$$

*where $\gamma$ is a constant.*

*(2) If $\|u + e_1\|_2 \le \epsilon_0$, then*

$$\sum_{i,j=1}^n \xi_i\xi_j(\partial_{ij} f)(u) \ge \gamma > 0, \qquad \forall\, \xi \in \mathbb{S}^{n-1},$$

*where $\gamma$ is a constant.*

*In yet other words, $f(u)$ is strongly convex in a sufficiently small neighborhood of $\pm e_1$.*

*Proof of Theorem 7.3.* See appendix. □

Finally we complete the proof of Theorem 7.1.

*Proof of Theorem 7.1.* We proceed in several steps. All the statements below hold under the assumption that $m \gtrsim n$ and with high probability.

(1) For $u = 0$, we use Lemma 7.1. In particular $u = 0$ is a local maximum point with strictly negative Hessian.

(2) For $\|u\|_2 \ll 1$ or $\|u\|_2 \gg 1$, we use Lemma 7.2. The loss functions has a nonzero gradient ($\partial_R f \neq 0$) in this regime.

(3) For $\|u\|_2 \sim 1$ with $\epsilon_0 \le |\hat{u} \cdot e_1| \le 1 - \epsilon_0$, we use Lemma 7.3 to show that the loss function has a nonzero gradient ($\partial_\theta f \neq 0$) in this regime.

(4) For $\|u\|_2 \sim 1$ with $|\hat{u} \cdot e_1| \le \epsilon_0$, by Lemma 7.4, the loss function has a negative curvature direction (i.e. $\partial_{\theta\theta} f < 0$) in this regime.

(5) For $\|u\|_2 \sim 1$, $||\hat{u} \cdot e_1| - 1| \le \epsilon_0$, $|\|u\|_2 - 1| \ge c(\epsilon_0)$, Theorem 7.2 shows that the gradient of the loss function does not vanish (i.e. $\partial_R f \neq 0$).

(6) For $\|u \pm e_1\| \ll 1$, Theorem 7.3 gives the strong convexity in the full neighborhood.

It is not difficult to check that the above 6 scenarios cover the whole of $\mathbb{R}^n$. We omit further details. □

## 8. Numerical Experiments

In this section, we demonstrate the numerical efficiency of our models by simple gradient descent and compare their performance with other competitive algorithms.

In a concurrent work [5], we considered the following piecewise Smoothed Amplitude Model (SAF):

$$\min_{\boldsymbol{z}\in\mathbb{R}^n} \quad F(\boldsymbol{z}) = \frac{1}{2m}\sum_{i=1}^m \left(\gamma\left(\frac{|\boldsymbol{a}_i^\top \boldsymbol{z}|}{|\boldsymbol{a}_i^\top \boldsymbol{x}|}\right) - 1\right)^2 \cdot |\boldsymbol{a}_i^\top \boldsymbol{x}|^2$$

with the function $\gamma(t)$

$$\gamma(t) := \begin{cases} |t|, & |t| > \beta; \\ \frac{1}{2\beta}t^2 + \frac{\beta}{2}, & |t| \le \beta. \end{cases}$$

In the above the allowed range for $\beta$ is $0 < \beta \le \frac{1}{2}$. We shall showcase some numerical results for this model.

In this work, our first Perturbed Amplitude Flow (PAF1) is model

$$\min_{\boldsymbol{z}\in\mathbb{R}^n} \quad F(\boldsymbol{z}) = \frac{1}{2m}\sum_{i=1}^m \left(\sqrt{\beta\|\boldsymbol{z}\|^2 + |\boldsymbol{a}_i^\top \boldsymbol{z}|^2} - \sqrt{\beta\|\boldsymbol{z}\|^2 + |\boldsymbol{a}_i^\top \boldsymbol{x}|^2}\right)^2.$$

The second Perturbed Amplitude Flow (PAF2) is

$$\min_{\boldsymbol{z}\in\mathbb{R}^n} \quad F(\boldsymbol{z}) = \frac{1}{2m}\sum_{i=1}^m \left(\sqrt{\beta\|\boldsymbol{z}\|^2 + |\boldsymbol{a}_i^\top \boldsymbol{z}|^2 + |\boldsymbol{a}_i^\top \boldsymbol{x}|^2} - \sqrt{\beta\|\boldsymbol{z}\|^2 + 2|\boldsymbol{a}_i^\top \boldsymbol{x}|^2}\right)^2.$$

Our first Quotient Wirtinger Flow (QWF1) model is

$$\min_{z\in\mathbb{R}^n} \quad F(\boldsymbol{z}) = \frac{1}{2m}\sum_{i=1}^{m} \frac{\left(|\boldsymbol{a}_i^\top \boldsymbol{z}|^2 - |\boldsymbol{a}_i^\top \boldsymbol{x}|^2\right)^2}{\beta\|\boldsymbol{z}\|^2 + |\boldsymbol{a}_i^\top \boldsymbol{x}|^2}.$$

The second Quotient Wirtinger Flow (QWF2) model is

$$\min_{z\in\mathbb{R}^n} \quad F(\boldsymbol{z}) = \frac{1}{2m}\sum_{i=1}^{m} \frac{\left(|\boldsymbol{a}_i^\top \boldsymbol{z}|^2 - |\boldsymbol{a}_i^\top \boldsymbol{x}|^2\right)^2}{\beta\|\boldsymbol{z}\|^2 + \beta_1|\boldsymbol{a}_i^\top \boldsymbol{z}|^2 + \beta_2|\boldsymbol{a}_i^\top \boldsymbol{x}|^2}.$$

We have show theoretically that any gradient descent algorithm will not get trapped in a local minimum for the models above. Here we present numerical experiments to show that the models perform very well with randomized initial guess.

We use the following vanilla gradient descent algorithm

$$\boldsymbol{z}_{k+1} = \boldsymbol{z}_k - \mu\nabla F(\boldsymbol{z}_k)$$

with a random initial guess to minimize the loss function $F(\boldsymbol{z})$ given above. The pseudocode for the algorithm is as follows.

**Algorithm 1** Gradient Descend Algorithm Based on Our New Models

**Input:** Measurement vectors: $a_i \in \mathbb{R}^n, i = 1, \ldots, m$; Observations: $y \in \mathbb{R}^m$; Parameters $\beta, \beta_1$ and $\beta_2$; Step size $\mu$; Tolerance $\epsilon > 0$

1: Random initial guess $\boldsymbol{z}_0 \in \mathbb{R}^n$.

2: For $k = 0, 1, 2, \ldots,$ if $\|\nabla F(\boldsymbol{z}_k)\| \geq \epsilon$ do

$$\boldsymbol{z}_{k+1} = \boldsymbol{z}_k - \mu\nabla F(\boldsymbol{z}_k)$$

3: End for

**Output:** The vector $\boldsymbol{z}_T$.

The performance of our SAF, PAF1, PAF2, QWF1 and QWF2 algorithms are conducted via a series of numerical experiments in comparison against Trust Region [27], WF [3], TWF [6] and TAF [31]. Here, it is worth emphasizing that random initialization is used for Trust Region [27] and our SAF, PAF1, PAF2, QWF1, QWF2 algorithms while all other algorithms have adopted a spectral initialization. Our theoretical results are for real Gaussian case, but the algorithms can be easily adapted to the complex Gaussian and CDP cases. All experiments are carried out on a MacBook Pro with a 2.3GHz Intel Core i5 Processor and 8 GB 2133 MHz LPDDR3 memory.

8.1. **Recovery of 1D Signals.** In our numerical experiments, the target vector $\boldsymbol{x} \in \mathbb{R}^n$ is chose randomly from the standard Gaussian distribution and the measurement vectors $\boldsymbol{a}_i$, $i = 1, \ldots, m$ are generated randomly from standard Gaussian distribution or CDP model. For the real Gaussian case, the signal $\boldsymbol{x} \sim \mathcal{N}(0, I_n)$ and measurement vectors $\boldsymbol{a}_i \sim \mathcal{N}(0, I_n)$ for $i = 1, \ldots, m$. For the complex Gaussian case, the signal $\boldsymbol{x} \sim \mathcal{N}(0, I_n) + i\mathcal{N}(0, I_n)$ and measurement vectors $\boldsymbol{a}_i \sim \mathcal{N}(0, I_n/2) + i\mathcal{N}(0, I_n/2)$. For the CDP model, we use masks of octanary patterns as in [3]. For simplicity, our parameters and step size are fixed for all experiments. Specifically, we adopt parameter $\beta = 1/2$ and step size $\mu = 1$ for SAF. We choose the parameter $\beta = 1$, step size $\mu = 0.6$ and $\mu = 2.5$ for PAF1 and PAF2, respectively. For QWF1, we choose parameter $\beta = 1$ and step size $\mu = 0.4$. For QWF2, we adopt parameters $\beta = \beta_2 = 1$, $\beta_1 = 0.1$ and step size $\mu = 0.3$. For Trust Region, WF, TWF and TAF, we use the code provided in the original papers with suggested parameters.

**Example 8.1.** In this example, we test the empirical success rate of SAF, PAF1, PAF2, QWF1 and QWF2 versus the number of measurements. We conduct the experiments for the real Gaussian, complex Gaussian and CDP cases respectively. We choose $n = 128$ and the maximum number of iterations is $T = 2500$. For real and complex Gaussian cases, we vary $m$ within the range $[n, 10n]$. For CDP case, we set the ratio $m/n = L$ from 2 to 10. For each $m$, we run 100 times trials to calculate the success rate. Here, we say a trial to have successfully reconstructed the target signal if the relative error satisfies $\mathrm{dist}(\boldsymbol{z}_T - \boldsymbol{x})/\|\boldsymbol{x}\| \leq 10^{-5}$. The results are plotted in Figure 1. It can be seen that $5n$ Gaussian phaseless measurement or 5 octanary patterns are enough for exactly recovery for SAF. Similarly, $6n$ Gaussian phaseless measurement or 7 octanary patterns are enough for exactly recovery for PAF2, QWF1 and QWF2.

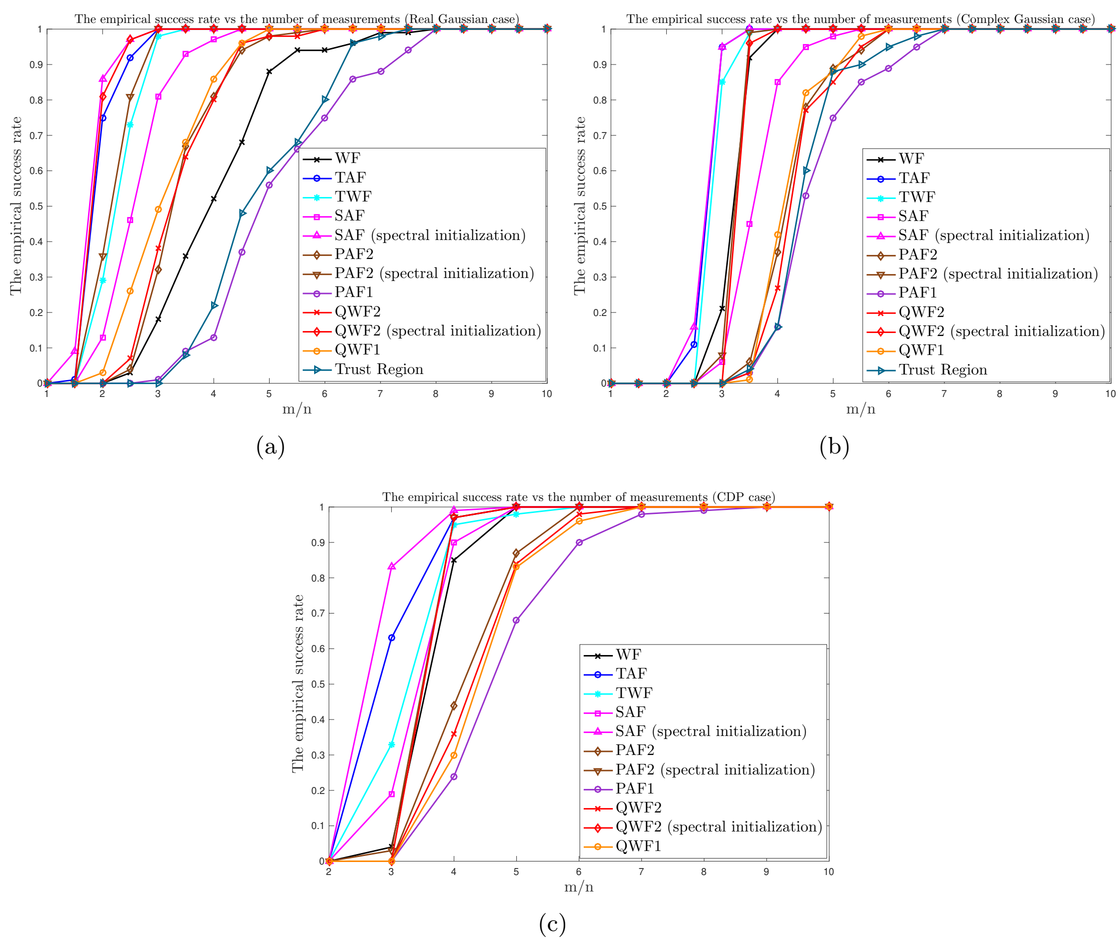
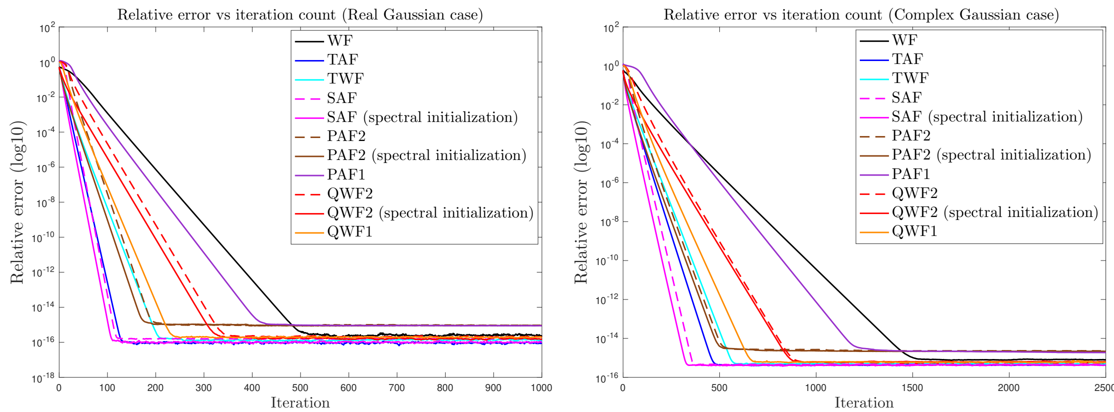



FIGURE 1. The empirical success rate for different $m/n$ based on 100 random trails. (a) Success rate for real Gaussian case, (b) Success rate for complex Gaussian case, (c) Success rate for CDP case.

**Example 8.2.** In this example, we compare the convergence rate of SAF, PAF1, PAF2, QWF1 and QWF2 with those of WF, TWF, TAF for real Gaussian and complex Gaussian cases. We choose $n = 128$ and $m = 6n$. The results are presented in Figure 2. Since our SAF algorithm chooses a random initial guess according to the standard Gaussian distribution instead of adopting a spectral initialization, it sometimes need to escape the saddle points with a small number of iterations. Due to its high efficiency to escape the saddle points, it still performs well comparing with state-of-the-art algorithms with spectral initialization.




(a) (b)

FIGURE 2. Relative error versus number of iterations for SAF, WF, TWF, and TAF method: (a) The noiseless measurements; (b) The noisy measurements.

**Example 8.3.** In this example, we compare the time elapsed and the iteration needed for WF, TWF, TAF and our SAF, PAF1, PAF2, QWF1, QWF2 to achieve the relative error $10^{-5}$ and $10^{-10}$, respectively. We choose $n = 1000$ with $m = 8n$. We adopt the same spectral initialization method for WF, TWF, TAF and the initial guess is obtained by power method with 50 iterations. We run 50 times trials to calculate the average time elapsed and iteration number for those algorithms. The results are shown in Table 1. The numerical results show that SAF takes around 20 and 40 iterations to escape the saddle points for the real and complex Gaussian cases, respectively. Similarly, the iteration numbers of escaping the saddle points are around $15, 42$ for PAF2 and $27, 50$ for QWF2 in the real and complex cases. Since there is no spectral initialization, the high efficiency of escaping saddle points and low computational complexity, the time elapsed of SAF is less than the other methods significantly.

8.2. **Recovery of Natural Image.** We next compare the performance of the above algorithms on recovering a natural image from masked Fourier intensity measurements. The image is the Milky Way Galaxy with resolution $1080 \times 1920$. The colored image has RGB channels. We use $L = 20$ random octanary patterns to obtain the Fourier intensity measurements for each R/G/B channel as in [3]. Table 2 lists the averaged time elapsed and the iteration needed to achieve the relative error $10^{-5}$ and $10^{-10}$ over the three RGB channels. We can see that our algorithms

Table 1. Time Elapsed and Iteration Number among Algorithms on Gaussian Signals with $n = 1000$.

| Algorithm | Real Gaussian | | | | Complex Gaussian | | | |
|---|---|---|---|---|---|---|---|---|
| | $10^{-5}$ | | $10^{-10}$ | | $10^{-5}$ | | $10^{-10}$ | |
| | Iter | Time(s) | Iter | Time(s) | Iter | Time(s) | Iter | Time(s) |
| SAF | 44 | **0.1556** | 68 | **0.2276** | 113 | **1.3092** | 190 | **2.3596** |
| SAF (spectral) | 25 | 0.2631 | 51 | 0.3309 | 67 | 1.4528 | 151 | 2.6122 |
| PAF1 | 108 | 3.3445 | 204 | 5.5768 | 291 | 35.8624 | 591 | 75.3231 |
| PAF2 | 46 | 1.5816 | 84 | 2.1980 | 129 | 15.8295 | 239 | 27.6362 |
| PAF2 (spectral) | 31 | 1.1867 | 66 | 2.0250 | 88 | 9.9784 | 197 | 21.0727 |
| QWF1 | 58 | 2.0589 | 117 | 3.7204 | 155 | 21.6235 | 314 | 37.1972 |
| QWF2 | 88 | 2.4423 | 161 | 4.2229 | 211 | 30.2235 | 422 | 48.1972 |
| QWF 2(spectral) | 61 | 1.8736 | 134 | 3.6696 | 162 | 17.0657 | 370 | 40.0686 |
| WF | 125 | 4.4214 | 229 | 6.3176 | 304 | 34.6266 | 655 | 86.6993 |
| TAF | 29 | 0.2744 | 60 | 0.3515 | 100 | 1.7704 | 211 | 2.7852 |
| TWF | 40 | 0.3181 | 87 | 0.4274 | 112 | 1.9808 | 244 | 3.7432 |
| Trust Region | **21** | 2.9832 | **29** | 4.4683 | **33** | 19.1252 | **42** | 29.0338 |

have good performance comparing with state-of-the-art algorithms with spectral initialization. Furthermore, our algorithms perform well even with $L = 10$ under 300 iterations, while WF fails. Figure 3 shows the image recovered by SAF. The images recovered by other algorithms are not presented because they are similar.

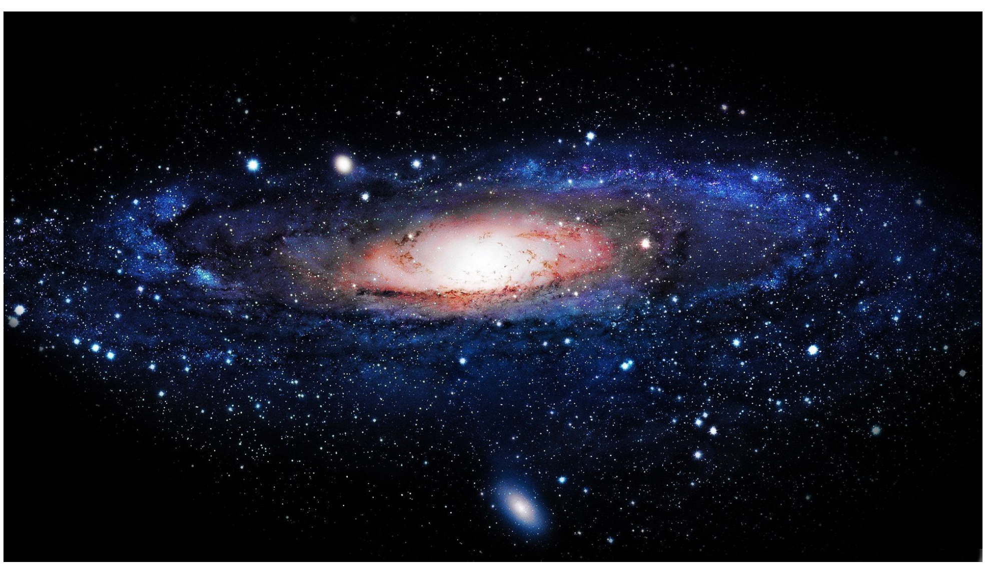


Figure 3. The Milky Way Galaxy image: SAF with $L = 10$ takes 300 iterations, computation time is 455.8 s, relative error is $3.12 \times 10^{-14}$.

Table 2. Time Elapsed and Iteration Number among Algorithms on Recovery of Galaxy Image.

| Algorithm | The Milky Way Galaxy | | | |
|---|---|---|---|---|
| | $10^{-5}$ | | $10^{-10}$ | |
| | Iter | Time(s) | Iter | Time(s) |
| SAF | 92 | **202.47** | 148 | **351.21** |
| PAF1 | 198 | 462.27 | 306 | 710.27 |
| PAF2 | 113 | 260.48 | 187 | 441.55 |
| QWF1 | 168 | 351.32 | 282 | 601.68 |
| QWF2 | 173 | 371.59 | 296 | 709.21 |
| WF | 158 | 381.7 | 277 | 621.63 |
| TAF | **65** | 223.89 | **122** | 368.22 |
| TWF | 68 | 315.14 | 145 | 566.84 |

8.3. **Recovery of signals with noise.** We now demonstrate the robustness of our SAF, PAF1, PAF2, QWF1, QWF2 to noise and compare them with WF, TWF, TAF. We consider the noisy model $y_i = |\langle \boldsymbol{a}_i, \boldsymbol{x}\rangle| + \eta_i$ and add different level of Gaussian noises to explore the relationship between the signal-to-noise rate (SNR) of the measurements and the mean square error (MSE) of the recovered signal. Specifically, SNR and MSE are evaluated by

$$\mathrm{MSE} := 10\log_{10}\frac{\mathrm{dist}^2(\boldsymbol{z},\boldsymbol{x})}{\|\boldsymbol{x}\|^2} \quad \text{and} \quad \mathrm{SNR} = 10\log_{10}\frac{\sum_{i=1}^{m}|\boldsymbol{a}_i^\top \boldsymbol{x}|^2}{\|\eta\|^2},$$

where $\boldsymbol{z}$ is the output of the algorithms given above after 2500 iterations. We choose $n = 128$ and $m = 8n$. The SNR varies from 20db to 60db. The result is shown in Figure 4. We can see that our algorithms perform well for noise phase retrieval.

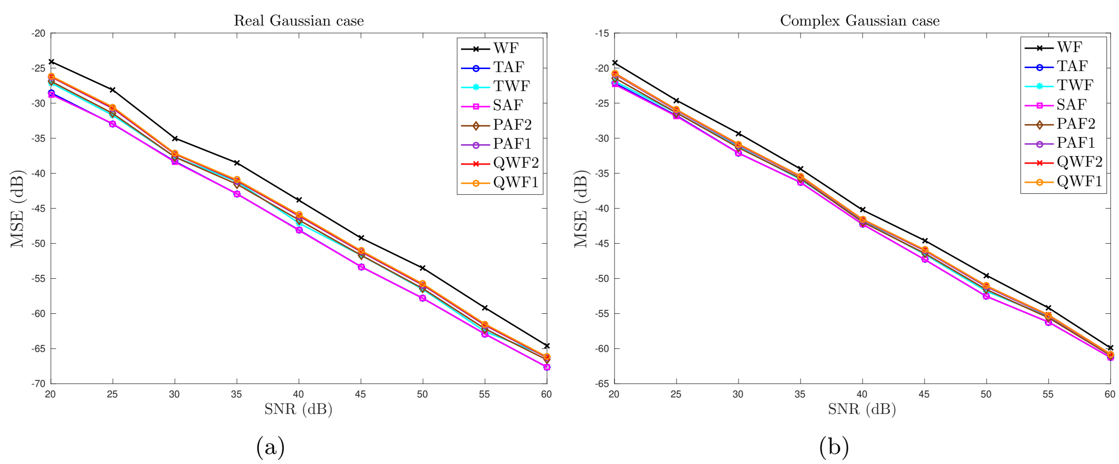



Figure 4. SNR versus relative MSE on a dB-scale under the noisy Gaussian model: (a) Real Gaussian case; (b) Complex Gaussian case.

## Appendix A. Auxiliary estimates

**Notation**. In this section we shall adopt the following convention.

- We shall denote by $c$ and $C$ positive absolute constants which will be needed in the various inequalities below. The values of $c$ and $C$ may vary from line to line. Often $c$ refers to a small absolute constant, and $C$ refers to a large absolute constant.
- For any quantity $X$, we write $X \lesssim 1$ if $X \le C_0$ and $C_0$ is an absolute constant.
- For a random variable $Y$, we shall sometimes use "mean" to denote $\mathbb{E}Y$. For example

  $$\mathbb{P}(|Y - \text{mean}\,| > \epsilon) \le e^{-c\epsilon}$$

  means

  $$\mathbb{P}(|Y - \mathbb{E}Y| > \epsilon) \le e^{-c\epsilon}.$$

  This notation is particularly handy when $Y$ is given by a sum of random variables involving various truncations and modifications.
- For a random variable $Y$, the sub-exponential orlicz norm $\|Y\|_{\psi_1}$ is defined as

  $$\|Y\|_{\psi_1} = \inf\{t > 0 : \mathbb{E}e^{\frac{|X|}{t}} \le 2\}.$$

  In particular $\|Y\|_{\psi_1} \le K \Leftrightarrow \mathbb{E}e^{\frac{|Y|}{K}} \le 2$. Similarly the sub-gaussian orlicz norm $\|Y\|_{\psi_2}$ is

  $$\|Y\|_{\psi_2} = \inf\{t > 0 : \mathbb{E}e^{\frac{|X|^2}{t}} \le 2\}.$$
- We denote $\{a_j\}_{j=1}^m$ as a sequence of i.i.d. random variables which are copies of a standard Gaussian random vector $a: \ \Omega \to \mathbb{R}^n$ satisfying $a \sim \mathcal{N}(0, I_n)$.
- We write $u \in \mathbb{S}^{n-1}$ if $u \in \mathbb{R}^n$ and $\|u\|_2 = \sqrt{\sum_j (u_j)^2} = 1$.

**Lemma A.1** (General Hoeffding's inequality)**.** *Let $X_i$, $1 \le i \le m$ be independent, mean zero, sub-gaussian random variables. Let $b = (b_1, \cdots, b_m) \in \mathbb{R}^m$. Then for every $t \ge 0$, we have*

$$\mathbb{P}(\Big|\sum_{i=1}^m b_i X_i\Big| \ge t) \le 2\exp\Big(-\frac{ct^2}{K^2\|b\|_2^2}\Big),$$

*where* $\max_i \|X_i\|_{\psi_2} = K$.

*Proof.* See Theorem 2.6.3 of [28]. ☐

**Lemma A.2** (General Hoeffding's inequality, practical version)**.** *Let $X_i$, $1 \le i \le m$ be independent, sub-gaussian random variables. Assume*

$$\max_i \|X_i\|_{\psi_2} \lesssim 1.$$

*Let $b = (b_1, \cdots, b_m) \in \mathbb{R}^m$. Then for every $t \ge 0$, we have*

$$\mathbb{P}(\Big|\frac{1}{m}\sum_{i=1}^m b_i(X_i - \mathbb{E}X_i)\Big| \ge t) \le 2\exp\Big(-\frac{cm^2t^2}{\|b\|_2^2}\Big).$$

*Proof.* Apply Lemma A.1 to $X_i - \mathbb{E}X_i$ and observe that $\max_i \|X_i - \mathbb{E}X_i\|_{\psi_2} \lesssim 1$. ☐

**Lemma A.3** (Bernstein's inequality). *Let $X_i$, $1 \le i \le m$ be independent, mean zero, sub-exponential random variables. Let $b = (b_1, \cdots, b_m) \in \mathbb{R}^m$. Then for every $t \ge 0$, we have*

$$\mathbb{P}(\Big|\frac{1}{m}\sum_{i=1}^m b_i X_i\Big| > t) \le 2\exp\left[-c\cdot\min(\frac{m^2t^2}{K^2\|b\|_2^2}, \frac{mt}{K\|b\|_\infty})\right],$$

*where $\max_i \|X_i\|_{\psi_1} = K$. In particular, setting $b_i = 1$, we obtain*

$$\mathbb{P}(\Big|\frac{1}{m}\sum_{i=1}^m X_i\Big| > t) \le 2\exp\left[-c\cdot m\cdot\min(\frac{t^2}{K^2}, \frac{t}{K})\right],$$

*Proof.* See Theorem 2.8.2 of [28]. □

**Lemma A.4** (Bernstein-type inequalities, practical version 1). *Let $X_i$, $1 \le i \le m$ be independent, sub-exponential random variables such that*

$$\max_{1\le i\le m} \|X_i\|_{\psi_1} \le \tau.$$

*Then for any $\epsilon > 0$,*

$$\mathbb{P}(\Big|\frac{1}{m}\sum_{i=1}^m X_i - \mathrm{mean}\Big| > \epsilon) \le 2\exp\left[-c\cdot m\cdot\min(\frac{\epsilon^2}{\tau^2}, \frac{\epsilon}{\tau})\right].$$

*In particular, if $\max_{1\le i\le m}\|X_i\|_{\psi_1} \lesssim 1$ and $0 < \epsilon \le 1$, then*

$$\mathbb{P}(\Big|\frac{1}{m}\sum_{i=1}^m X_i - \mathrm{mean}\Big| > \epsilon) \le 2\exp(-c\cdot m\cdot\epsilon^2).$$

*Proof.* Observe that $\|X_i - \mathbb{E}X_i\|_{\psi_1} \le 2\|X_i\|_{\psi_1}$. By adjusting the absolute constant $c$ slightly in Lemma A.3 one can then obtain the result. □

**Lemma A.5** (Bernstein-type inequalities, practical version 2). *Let $X_i$, $1 \le i \le m$ be independent, bounded random variables such that*

$$\max_{1\le i\le m} \|X_i\|_{\infty} \le \tau.$$

*Then for any $\epsilon > 0$,*

$$\mathbb{P}(\Big|\frac{1}{m}\sum_{i=1}^m X_i - \mathrm{mean}\Big| > \epsilon) \le 2\exp\left[-c\cdot m\cdot\min(\frac{\epsilon^2}{\tau^2}, \frac{\epsilon}{\tau})\right].$$

*In particular, if $\max_{1\le i\le m}\|X_i\|_{\infty} \lesssim 1$ and $0 < \epsilon \le 1$, then*

$$\mathbb{P}(\Big|\frac{1}{m}\sum_{i=1}^m X_i - \mathrm{mean}\Big| > \epsilon) \le 2\exp(-c\cdot m\cdot\epsilon^2).$$

*Proof.* One can give a direct proof from scratch. But it is easier to deduce it from Lemma A.4. Since $\mathbb{E}e^{\frac{|X_i|}{2\tau}} \le e^{\frac{1}{2}} \le 2$, we have $\|X_i\|_{\psi_1} \le 2\|X_i\|_\infty$. The result then follows from Lemma A.4. □

**Lemma A.6.** *For any $\epsilon > 0$, there exists $N_0 = N_0(\epsilon) > 0$, such that for any $N \ge N_0$, we have*

$$\frac{1}{m}\sum_{j=1}^m \mathbb{E}\left((a_j\cdot u)^2(a_j\cdot v)^2\chi_{|a_j\cdot v|\ge N}\right) \le \epsilon,$$

*where $u \in \mathbb{S}^{n-1}$, $v \in \mathbb{S}^{n-1}$.*

*Proof.* Since $a_j$ are i.i.d., it suffices to prove the statement for a single random vector $a \sim \mathcal{N}(0, I_n)$. Noting that $a \cdot u \sim \mathcal{N}(0,1)$ and $a \cdot v \sim \mathcal{N}(0,1)$, we have

$$\begin{aligned}
&\mathbb{E}\left((a\cdot u)^2(a\cdot v)^2\chi_{|a\cdot v|\geq N}\right)\\
\leq& \frac{1}{4}\epsilon\mathbb{E}(a\cdot u)^4 + \frac{1}{\epsilon}\mathbb{E}((a\cdot v)^4\chi_{|a\cdot v|\geq N})\\
\leq& \frac{3}{4}\epsilon + \frac{1}{\epsilon}\cdot\frac{1}{\sqrt{2\pi}}\int_{|y|\geq N} e^{-\frac{y^2}{2}}y^4 dy \leq \epsilon,
\end{aligned}$$

if $N$ is sufficiently large. Note that one can easily quantify $N_0$ in terms of $\epsilon$. However we shall not dwell on this here. $\square$

**Lemma A.7.** *Let $A = \frac{1}{m}\sum_{j=1}^m a_j a_j^T$. Then for any $0 < \epsilon \leq 1$, we have*

$$\mathbb{P}(\|A - \mathrm{I}\|_{\mathrm{op}} > \epsilon) \leq \exp(-cm\epsilon^2), \qquad \text{if } m \geq C\epsilon^{-2}n.$$

*In particular, for $m \geq C\epsilon^{-2}n$, with probability at least $1 - e^{-cm\epsilon^2}$, we have*

$$\Big|\frac{1}{m}\sum_{j=1}^m (a_j\cdot u)(a_j\cdot v) - \text{mean}\Big| \leq \epsilon, \qquad \forall\, u, v \in \mathbb{S}^{n-1}.$$

*Proof.* We briefly sketch the standard proof here for the sake of completeness. By using a $\delta$-net $S_\delta$ on $\mathbb{S}^{n-1}$ with $0 < \delta < \frac{1}{2}$ and $\mathrm{Card}(S_\delta) \leq (1+\frac{2}{\delta})^n$, we have

$$\|A - \mathrm{I}\|_{\mathrm{op}} \leq \frac{1}{1-2\delta}\sup_{x,y\in S_\delta}\langle (A-\mathrm{I})x, y\rangle.$$

Now for a pair of fixed $x_0$, $y_0 \in S_\delta$, since $\|(a_i\cdot x)(a_i\cdot y)\|_{\psi_1} \leq \|a_i\cdot x_0\|_{\psi_2}\|a_i\cdot y_0\|_{\psi_2} \lesssim 1$, by using Lemma A.4, we have for any $0 < \epsilon_1 \leq 1$,

$$\mathbb{P}(\langle (A-\mathrm{I})x_0, y_0\rangle > \epsilon_1) \leq 2\exp(-c\cdot m\cdot\epsilon_1^2).$$

Thus

$$\mathbb{P}(\|A - \mathrm{I}\|_{\mathrm{op}} > \epsilon) \leq 2(1+\frac{2}{\delta})^{2n}\exp\big[-cm(1-2\delta)^2\epsilon^2\big].$$

Taking $\delta = \frac{1}{4}$ and $m \geq O(\epsilon^{-2})n$ then yields the result. $\square$

**Lemma A.8.** *Suppose $h : \mathbb{R} \to \mathbb{R}$ is a locally Lipschitz continuous function such that*

$$|h(z) - h(\tilde{z})| \lesssim (1 + |z| + \tilde{z})|z - \tilde{z}|, \qquad \forall\, z, \tilde{z} \in \mathbb{R}.$$

*Assume that for a standard Gaussian random variable $Z \sim N(0,1)$, we have*

$$\|h(Z)\|_{\psi_1} \lesssim 1.$$

*Then the following hold:*

*For any $0 < \epsilon \leq \frac{1}{2}$, if $m \geq Cn\epsilon^{-2}|\log\epsilon|$, then with probability at least $1 - e^{-cm\epsilon^2}$, we have*

$$\Big|\frac{1}{m}\sum_{j=1}^m h(a_j\cdot u) - \text{mean}\Big| \leq \epsilon, \qquad \forall\, u \in \mathbb{S}^{n-1}.$$

*Proof.* Introduce a $\delta$-net $S_\delta$ on $\mathbb{S}^{n-1}$ with $\mathrm{Card}(S_\delta) \le (1+\frac{2}{\delta})^n$. Observe that by Lemma A.4, for any $0 < \epsilon_1 \le 1$,

$$\mathbb{P}(\sup_{u \in S_\delta} \Big| \frac{1}{m} \sum_{j=1}^m h(a_j \cdot u) - \mathrm{mean} \Big| > \epsilon_1) \le (1+\frac{2}{\delta})^n \cdot 2 \cdot \exp(-c \cdot m \cdot \epsilon_1^2).$$

By Lemma A.7, we have for $m \ge Cn$ with probability at least $1 - e^{-cm}$, it holds that

$$\frac{1}{m} \sum_{j=1}^m |a_j \cdot u|^2 \le 2, \quad \forall\, u \in \mathbb{S}^{n-1}.$$

Thus with probability at least $1 - e^{-cm}$ and uniformly for $u$, $v \in \mathbb{S}^{n-1}$, we have

$$\begin{aligned}
&\Big| \frac{1}{m} \sum_{j=1}^m h(a_j \cdot u) - \frac{1}{m} \sum_{j=1}^m h(a_j \cdot v) \Big| \\
\le & C_1 \frac{1}{m} \sum_{j=1}^m |a_j \cdot (u-v)| + C_1 \frac{1}{m} \sum_{j=1}^m |a_j \cdot (u-v)| (|a_j \cdot u| + |a_j \cdot v|) \\
\le & C_1 \Big( \frac{1}{m} \sum_{j=1}^m 10 \cdot C_1 \cdot \epsilon^{-1} |a_j \cdot (u-v)|^2 + \frac{\epsilon}{40 C_1} \\
& \qquad + \frac{1}{m} \cdot \frac{1000 C_1}{\epsilon} \sum_{j=1}^m |a_j \cdot (u-v)|^2 + \frac{1}{m} \cdot \frac{\epsilon}{100 C_1} \sum_{j=1}^m (|a_j \cdot u|^2 + |a_j \cdot v|^2) \Big) \\
\le & \frac{\epsilon}{4} + C_2 \cdot \epsilon^{-1} \|u - v\|_2,
\end{aligned}$$

where $C_1 > 0$, $C_2 > 0$ are absolute constants.

Now we take $\epsilon_1 = \frac{\epsilon}{4}$ and $\delta = \frac{\epsilon^2}{4C_2}$. It follows that for $m \ge Cn$ with probability at least

$$1 - e^{-cm} - (1 + \frac{8C_2}{\epsilon^2})^n \cdot 2 \cdot \exp(-c \cdot m \cdot \epsilon^2),$$

the desired inequality holds uniformly for all $u \in \mathbb{S}^{n-1}$. Clearly we need to take $m$ larger, i.e. $m \ge Cn\epsilon^{-2} |\log \epsilon|$ to get the probability bound $1 - \exp(-c \cdot m \cdot \epsilon^2)$. □

*Corollary* A.1. Let $0 < \epsilon \le \frac{1}{2}$. There exists $N_0(\epsilon) = C\sqrt{|\log \epsilon|} > 0$ such that the following hold for any $N \ge N_0$:

If $m \ge C\epsilon^{-2} |\log \epsilon| n$, then with probability at least $1 - e^{-cm\epsilon^2}$, we have

$$\frac{1}{m} \sum_{j=1}^m \chi_{|a_j \cdot u| > N} \le \epsilon, \qquad \forall\, u \in \mathbb{S}^{n-1}.$$

*Proof.* We choose $\phi \in C_c^\infty(\mathbb{R})$ such that $\phi(z) \equiv 1$ for $|z| \le \frac{1}{2}$, $\phi(z) = 0$ for $|z| \ge 1$, and $0 \le \phi(z) \le 1$ for all $z \in \mathbb{R}$. Clearly then

$$\frac{1}{m} \sum_{j=1}^m \chi_{|a_j \cdot u| > N} \le \frac{1}{m} \sum_{j=1}^m (1 - \phi(\frac{a_j \cdot u}{N})).$$

We can then apply Lemma A.8 with $h(z) = 1 - \phi(\frac{z}{N})$. Note that (below $Z \sim \mathcal{N}(0,1)$ is a standard normal random variable)

$$\mathrm{mean} = \mathbb{E}(1 - \phi(\frac{Z}{N})) \le \mathbb{E}\chi_{|Z| \ge \frac{1}{2}N} \le O(e^{-N^2}) \le \frac{1}{2}\epsilon,$$

if $N \ge N_0(\epsilon)$. □

**Lemma A.9.** *Let $X_i$: $1 \le i \le m$ be independent random variables with*

$$\max_{1\le i\le m} \mathbb{E}|X_i|^4 \lesssim 1.$$

*Then for any $t>0$,*

$$\mathbb{P}(\Big|\frac{1}{m}\sum_{j=1}^m X_i - \mathrm{mean}\Big| > t) \lesssim \frac{1}{m^2 t^4}.$$

*Proof.* Without loss of generality we can assume $X_i$ has zero mean. The result then follows from the observation that

$$\mathbb{E}(\sum_{j=1}^m X_j)^4 \lesssim \sum_{i<j} \mathbb{E}X_i^2 X_j^2 + \sum_i \mathbb{E}X_i^4 \lesssim m^2.$$

□

**Lemma A.10.** *Let $0 < \epsilon \le \frac{1}{2}$. There exists $N_0(\epsilon) > 0$ such that the following hold for any $N \ge N_0$:*

*If $m \ge C\epsilon^{-4}|\log\epsilon| n$, then with probability at least $1-\frac{c}{m^2}$, we have*

$$\frac{1}{m}\sum_{j=1}^m (a_j\cdot e_1)^4 \chi_{|a_j\cdot u|\ge N} \le \epsilon, \qquad \forall\, u \in \mathbb{S}^{n-1}.$$

*Proof.* By using Cauchy-Schwartz, we have

$$\begin{aligned}&\frac{1}{m}\sum_{j=1}^m (a_j\cdot e_1)^4 \chi_{|a_j\cdot u|\ge N}\\ &\le \frac{\eta}{4}\frac{1}{m}\sum_{j=1}^m (a_j\cdot e_1)^8 + \frac{1}{\eta}\frac{1}{m}\sum_{j=1}^m \chi_{|a_j\cdot u|\ge N}.\end{aligned}$$

By using Lemma A.9, we have for any $t_1>0$,

$$\mathbb{P}\Big(\Big|\frac{1}{m}\sum_{j=1}^m (a_j\cdot e_1)^8 - \mathrm{mean}\Big| > t_1\Big) \lesssim \frac{1}{m^2 t_1^4}.$$

Now choose $t_1$ to be a sufficiently large absolute constant and $\eta = \frac{\epsilon}{C_1}$ for some sufficiently large absolute constant $C_1$. The desired result then follows from Corollary A.1. Note that due to the factor $\frac{1}{\eta}$ we have to work with $\epsilon^2$ instead of $\epsilon$. □

**Lemma A.11.** *For any $0<\epsilon\le\frac{1}{2}$, there exist $C_1 = C_1(\epsilon) > 0$, $c_1 = c_1(\epsilon) > 0$, such that if $m \ge C_1 \cdot n$, then the following holds with probability at least $1-e^{-c_1 m}$:*

$$\frac{1}{m}\sum_{j=1}^m (a_j\cdot u)^4 \ge 3-\epsilon, \qquad \forall\, u\in\mathbb{S}^{n-1}.$$

*Proof.* Let $\phi \in C_c^\infty(\mathbb{R})$ be such that $0 \le \phi(x) \le 1$ for all $x\in\mathbb{R}$, $\phi(x) = 1$ for $|x|\le 1$ and $\phi(x) = 0$ for $|x|\ge 2$. Then for any $N \ge 1$,

$$\frac{1}{m}\sum_{j=1}^m (a_j\cdot u)^4 \ge \frac{1}{m}\sum_{j=1}^m (a_j\cdot u)^4 \phi(\frac{a_j\cdot u}{N}).$$

We first take $N$ sufficiently large (depending only on $\epsilon$) such that

$$\frac{1}{m}\sum_{j=1}^{m}\mathbb{E}(a_j\cdot u)^4\phi(\frac{a_j\cdot u}{N}) = \mathbb{E}Z^4\phi(\frac{Z}{N}) \geq 3-\frac{\epsilon}{2},$$

where $Z\sim\mathcal{N}(0,1)$.

Then by using Lemma A.8, we have for $m\geq C_1\cdot n$,

$$\mathbb{P}(\Big|\frac{1}{m}\sum_{j=1}^{m}(a_j\cdot u)^4\phi(\frac{a_j\cdot u}{N})-\text{mean}\Big| > \frac{\epsilon}{2}) \leq 2\exp(-2c_1 m) \leq \exp(-c_1\cdot m).$$

The desired result then easily follows. □

**Lemma A.12.** *Suppose $f_1, f_2 : \mathbb{R}\to\mathbb{R}$ are Lipschitz continuous functions such that*

$$\begin{aligned}
&|f_1(z)|+|f_2(z)| \leq L\cdot(1+|z|), \qquad \forall\, z\in\mathbb{R};\\
&|f_k(z)-f_k(\tilde{z})| \leq L\cdot|z-\tilde{z}|, \qquad \forall\, z,\tilde{z}\in\mathbb{R},\ k=1,2,
\end{aligned}$$

*where $L>0$ is a constant. Then the following hold:*

*For any $0<\epsilon\leq\frac{1}{2}$, there are constants $C_1=C_1(L,\epsilon)>0$, $c_1=c_1(L,\epsilon)>0$, such that if $m\geq C_1\cdot n$, then with probability at least $1-e^{-c_1 m}$, we have*

$$\Big|\frac{1}{m}\sum_{j=1}^{m}f_1(a_j\cdot u)f_2(a_j\cdot v)-\text{mean}\Big| \leq \epsilon, \qquad \forall\, u,\, v\in\mathbb{S}^{n-1}.$$

*Proof.* Introduce a $\delta$-net $S_\delta$ on $\mathbb{S}^{n-1}$ with $\mathrm{Card}(S_\delta)\leq(1+\frac{2}{\delta})^n$. By Lemma A.4, for any $0<\epsilon_1\leq 1$, we have

$$\mathbb{P}(\sup_{u,v\in S_\delta}\Big|\frac{1}{m}\sum_{j=1}^{m}f_1(a_j\cdot u)f_2(a_j\cdot v)-\text{mean}\Big| > \epsilon_1) \leq (1+\frac{2}{\delta})^{2n}\cdot 2\cdot\exp(-\tilde{c}_1\cdot m),$$

where $\tilde{c}_1>0$ depends only on $(L,\epsilon_1)$.

Next by Lemma A.7 and A.8, with probability at least $1-e^{-cm}$, it holds that

$$\begin{aligned}
&\Big|\frac{1}{m}\sum_{k=1}^{m}|a_j\cdot w|-\text{mean}\Big| \leq 0.01, \qquad \forall\, w\in\mathbb{S}^{n-1};\\
&\Big|\frac{1}{m}\sum_{k=1}^{m}|a_j\cdot w|^2-\text{mean}\Big| \leq 0.01, \qquad \forall\, w\in\mathbb{S}^{n-1}.
\end{aligned}$$

Consequently, on the same set (i.e. same $a_j$ satisfying the above two inequalities), the following hold: for any $u, v \in \mathbb{S}^{n-1}$, take $\tilde{u}, \tilde{v} \in S_\delta$ such that $\|u-\tilde{u}\|_2 \le \delta$, $\|v-\tilde{v}\|_2 \le \delta$, then

$$
\begin{aligned}
&\frac{1}{m}\sum_{j=1}^m |f_1(a_j\cdot u) f_2(a_j\cdot v) - f_1(a_j\cdot\tilde{u}) f_2(a_j\cdot\tilde{v})| \\
\le\, & \frac{1}{m}\sum_{j=1}^m |f_1(a_j\cdot u) - f_1(a_j\cdot\tilde{u})||f_2(a_j\cdot v)| + \frac{1}{m}\sum_{j=1}^m |f_1(a_j\cdot\tilde{u})||f_2(a_j\cdot v) - f_2(a_j\cdot\tilde{v})| \\
\le\, & \frac{1}{m}\sum_{j=1}^m L^2|a_j\cdot(u-\tilde{u})|(1+|a_j\cdot v|) + \frac{1}{m}\sum_{j=1}^m L^2(1+|a_j\cdot\tilde{u}|)|a_j\cdot(v-\tilde{v})| \\
\le\, & \frac{1}{m}L^2\sum_{j=1}^m(|a_j\cdot(u-\tilde{u})| + |a_j\cdot(v-\tilde{v})|) + \frac{1}{m}\sum_{j=1}^m L^2\cdot\frac{1}{4\delta}|a_j\cdot(u-\tilde{u})|^2 + \frac{1}{m}\sum_{j=1}^m L^2\delta|a_j\cdot v|^2 \\
& \quad + \frac{1}{m}\sum_{j=1}^m L^2\cdot\frac{1}{4\delta}|a_j\cdot(v-\tilde{v})|^2 + \frac{1}{m}L^2\sum_{j=1}^m \delta|a_j\cdot\tilde{u}|^2 \\
\le\, & 20L^2(\|u-\tilde{u}\|_2 + \|v-\tilde{v}\|_2) + \frac{1}{\delta}L^2(\|u-\tilde{u}\|_2^2 + \|v-\tilde{v}\|_2^2) + 20\delta L^2 \le 100\delta L^2.
\end{aligned}
$$

Now set $\delta = \frac{\epsilon}{200L^2}$ and $\epsilon_1 = \frac{\epsilon}{3}$. The desired conclusion then follows with probability at least

$$1 - e^{-cm} - (1+\frac{2}{\delta})^{2n}\cdot 2\cdot\exp(-\tilde{c}_1\cdot m) \ge 1 - e^{-c_1 m},$$

if $m \ge C_1 n$. □

*Corollary* A.2. If $m \ge Cn$, then with probability at least $1-e^{-cm}$, we have

$$\frac{1}{m}\sum_{j=1}^m (a_j\cdot u)^2(a_j\cdot v)^2 \ge c_1 > 0, \qquad \forall\, u, v \in \mathbb{S}^{n-1},$$

where $c_1 > 0$ is an absolute constant.

*Proof.* Step 1. Write $v = su + \sqrt{1-s^2}u^\perp$, where $|s| \le 1$, and $u^\perp \in \mathbb{S}^{n-1}$ is such that $u^\perp\cdot u = 0$. Let $a \sim \mathcal{N}(0, I_n)$ and denote $X = a\cdot u$, $Y = a\cdot u^\perp$. Clearly $X \sim \mathcal{N}(0,1)$, $Y \sim \mathcal{N}(0,1)$ and $X$, $Y$ are independent. Now let $N \ge 4$. We have

$$
\begin{aligned}
& \mathbb{E}(a\cdot u)^2(a\cdot v)^2\chi_{|a\cdot u|\le N}\chi_{|a\cdot v|\le N} \\
=\, & \mathbb{E}X_1^2(sX_1 + \sqrt{1-s^2}Y_1)^2\chi_{|X_1|\le N}\chi_{|sX_1+\sqrt{1-s^2}Y_1|\le N} \\
\ge\, & \mathbb{E}X_1^2(sX_1 + \sqrt{1-s^2}Y_1)^2\chi_{|X_1|\le\frac{N}{4}}\chi_{|Y_1|\le\frac{N}{4}} \\
=\, & \mathbb{E}(s^2X_1^4 + (1-s^2)X_1^2Y_1^2)\chi_{|X_1|\le\frac{N}{4}}\chi_{|Y_1|\le\frac{N}{4}} \ge 2c_1 > 0,
\end{aligned}
$$

where $c_1 > 0$ is an absolute constant, and $N$ is taken to be a sufficiently large absolute constant.

Step 2. Let $\phi \in C_c^\infty(\mathbb{R})$ be such that $\phi(x) = x^2$ for $|x| \le N$ and $\phi(x) = 0$ for $|x| \ge N+1$. Clearly if $m \ge Cn$, then with probability at least $1-e^{-cm}$, we have

$$\Big|\frac{1}{m}\sum_{j=1}^m \phi(a_j\cdot u)\phi(a_j\cdot v) - \text{mean}\Big| \le \frac{1}{2}c_1$$

and thus

$$\frac{1}{m}\sum_{j=1}^{m}(a_j\cdot u)^2(a_j\cdot v)^2 \geq \frac{1}{m}\sum_{j=1}^{m}\phi(a_j\cdot u)\phi(a_j\cdot v) > c_1, \quad \forall\, u,v\in\mathbb{S}^{n-1}.$$

□

**Lemma A.13.** *Let $h:\mathbb{R}\to\mathbb{R}$ be a Lipschitz continuous function such that*

$$\sup_{z\in\mathbb{R}}\frac{|h(z)|}{1+|z|}\lesssim 1;$$
$$\sup_{z\neq\tilde z}\frac{|h(z)-h(\tilde z)|}{|z-\tilde z|}\lesssim 1.$$

*Define*

$$F=\{u\in\mathbb{R}^n:\ \|u\|_2\leq 1,\quad u\cdot e_1=0\}.$$

*Suppose $b=(b_1,\cdots,b_m)\in\mathbb{R}^m$ satisfies*

$$\frac{1}{m}\sum_{j=1}^{m}b_j^2\lesssim 1.$$

*For any $0<\epsilon\leq\frac{1}{2}$, if $m\geq Cn\epsilon^{-2}|\log\epsilon|$, then with probability at least $1-e^{-cm\epsilon^2}$, we have*

$$\Big|\frac{1}{m}\sum_{j=1}^{m}b_jh(a_j\cdot u)-\mathrm{mean}\Big|\leq\epsilon, \qquad \forall\, u\in F.$$

*Proof.* First it is easy to check that $\max_j\|h(a_j\cdot u)\|_{\psi_2}\lesssim 1$. By Lemma A.2, for each $u\in F$, we have

$$\mathbb{P}(\Big|\frac{1}{m}\sum_{j=1}^{m}b_jh(a_j\cdot u)-\mathrm{mean}\Big|>\frac{\epsilon}{4})\leq 2\exp(-cm\epsilon^2).$$

Now let $\delta>0$ and introduce a $\delta$-net $S_\delta$ on the set $F$. Note that the set $F$ can be identified as a unit ball in $\mathbb{R}^{n-1}$. We have $\mathrm{Card}(S_\delta)\leq(1+\frac{2}{\delta})^n$. Thus

$$\mathbb{P}(\sup_{u\in S_\delta}\Big|\frac{1}{m}\sum_{j=1}^{m}b_jh(a_j\cdot u)-\mathrm{mean}\Big|>\frac{\epsilon}{4})\leq 2(1+\frac{2}{\delta})^n\exp(-cm\epsilon^2).$$

By Lemma A.7, if $m\geq Cn$, then with probability at least $1-e^{-cm}$, we have

$$\frac{1}{m}\sum_{j=1}^{m}(a_j\cdot u)^2\leq 2\|u\|_2^2, \qquad \forall\, u\in\mathbb{R}^n.$$

Now if $u\in S_\delta$, $v\in F$ with $\|v-u\|_2\leq\delta$, then with probability at least $1-e^{-cm}$, we have

$$\begin{aligned}
&\Big|\frac{1}{m}\sum_{j=1}^{m}b_j\cdot h(a_j\cdot u)-\frac{1}{m}\sum_{j=1}^{m}b_j\cdot h(a_j\cdot v)\Big|\\
\leq&\ \frac{1}{m}\sum_{j=1}^{m}|b_j|\cdot K_0\cdot|a_j\cdot(u-v)|\\
\leq& K_0\frac{1}{m}(\sum_{j=1}^{m}b_j^2)^{\frac{1}{2}}(\sum_{j=1}^{m}|a_j\cdot(u-v)|^2)^{\frac{1}{2}}\\
\leq& K_1\|u-v\|_2,
\end{aligned}$$

where $K_0 > 0$, $K_1 > 0$ are absolute constants. On the other hand

$$\frac{1}{m}\sum_{j=1}^{m}|b_j|\mathbb{E}|h(a_j\cdot u)-h(a_j\cdot v)|$$
$$\lesssim \frac{1}{m}\sum_{j=1}^{m}|b_j|\cdot\|u-v\|_2 \lesssim \|u-v\|_2.$$

It follows that for some absolute constant $K_3 > 0$,

$$\Big|\frac{1}{m}\sum_{j=1}^{m} b_j\cdot\big(h(a_j\cdot u)-\mathbb{E}(h(a_j\cdot u))\big) - \frac{1}{m}\sum_{j=1}^{m} b_j\cdot\big(h(a_j\cdot v)-\mathbb{E}(h(a_j\cdot v))\big)\Big|$$
$$\le K_3\|u-v\|_2.$$

Now set $\delta = \frac{\epsilon}{10K_3+10}$. The desired conclusion then follows with probability at least

$$1-2(1+\frac{2}{\delta})^n\exp(-cm\epsilon^2)-e^{-cm} \ge 1-e^{-c_1 m\epsilon^2},$$

where $c_1 > 0$ is an absolute constant, and we need to take $m \ge Cn\epsilon^{-2}|\log\epsilon|$. □

**Lemma A.14.** *Let $h:\mathbb{R}\to\mathbb{R}$ be a Lipschitz continuous function such that*

$$\sup_{z\in\mathbb{R}}\frac{|h(z)|}{1+|z|}\lesssim 1;$$
$$\sup_{z\ne\tilde{z}}\frac{|h(z)-h(\tilde{z})|}{|z-\tilde{z}|}\lesssim 1.$$

*Let $f_3:\mathbb{R}\to\mathbb{R}$ be such that*

$$\sup_{z\in\mathbb{R}}\frac{f_3(z)}{1+|z|^3}\lesssim 1.$$

*Define*

$$F=\{u\in\mathbb{R}^n:\ \|u\|_2\le 1,\quad u\cdot e_1=0\}.$$

*Then for any $0<\epsilon\le\frac{1}{2}$, there exist $C_1=C_1(\epsilon)>0$, $C_2=C_2(\epsilon)>0$, such that if $m\ge C_1\cdot n$, then the following holds with probability at least $1-\frac{C_2}{m^2}$:*

$$\Big|\frac{1}{m}\sum_{j=1}^{m} f_3(a_j\cdot e_1)h(a_j\cdot u)-\mathrm{mean}\Big|\le\epsilon,\qquad \forall\, u\in F.$$

*Proof.* Step 1. Set $b_j=f_3(a_j\cdot e_1)$. Denote

$$K_0=\mathbb{E}b_1^2\lesssim 1.$$

By Lemma A.9, we have

$$\mathbb{P}(\Big|\frac{1}{m}\sum_{j=1}^{m} b_j^2-K_0\Big|>t)\lesssim\frac{1}{m^2t^4}.$$

Then with probability at least $1-\frac{1}{m^2}$, we have

$$\frac{1}{m}\sum_{j=1}^{m} b_j^2\le B_0,$$

where $B_0>0$ is some absolute constant.

Step 2. Denote $\tilde{a}_j = a_j - (a_j \cdot e_1)e_1$. An important observation is that $(a_j \cdot e_1)_{1\le j\le m}$ and $(\tilde{a}_j)_{1\le j\le m}$ are independent. Note that for $u \in F$ we have $a_j \cdot u = \tilde{a}_j \cdot u$. Thus for every $\tilde{b}_j$ with the property $\frac{1}{m}\sum_{j=1}^m (\tilde{b}_j)^2 \le B_0$, we have the following as a consequence of Lemma A.13: For any $0 < \epsilon \le \frac{1}{2}$, if $m \ge Cn\epsilon^{-2}|\log \epsilon|$, then with probability at least $1 - e^{-cm\epsilon^2}$, we have

$$\Big|\frac{1}{m}\sum_{j=1}^m \tilde{b}_j \cdot \big(h(\tilde{a}_j \cdot u) - \mathbb{E}(h(\tilde{a}_j \cdot u))\big)\Big| \le \frac{1}{3}\epsilon, \qquad \forall\, u \in F.$$

Step 3. By using the results from Step 1 and Step 2, with probability at least $1 - \frac{2}{m^2}$, we have

$$\Big|\frac{1}{m}\sum_{j=1}^m b_j \cdot \big(h(\tilde{a}_j \cdot u) - \mathbb{E}(h(\tilde{a}_j \cdot u))\big)\Big| \le \frac{1}{3}\epsilon, \qquad \forall\, u \in F.$$

Now note that $|\mathbb{E}(h(\tilde{a}_j \cdot u))| = |\mathbb{E}(h(\tilde{a}_1 \cdot u))| \lesssim 1$. By Lemma A.9, we have

$$\mathbb{P}(\Big|\frac{1}{m}\sum_{j=1}^m (b_j - \mathbb{E}b_j)\Big| > t) \lesssim \frac{1}{m^2 t^4}.$$

Choosing $t = \frac{\epsilon}{K_1}$ where $K_1 > 0$ is a sufficiently large absolute constant such that

$$\frac{1}{K_1} \cdot |\mathbb{E}b_1| \le \frac{\epsilon}{10}$$

then yields the result. $\square$

*Corollary* A.3. For any $0 < \epsilon \le \frac{1}{2}$, there exist $C_1 = C_1(\epsilon) > 0$, $C_2 = C_2(\epsilon) > 0$, such that if $m \ge C_1 \cdot n$, then the following holds with probability at least $1 - \frac{C_2}{m^2}$:

$$\Big|\frac{1}{m}\sum_{j=1}^m (a_j \cdot e_1)^3 (a_j \cdot u) - \text{mean}\Big| \le \epsilon, \qquad \forall\, u \in \mathbb{S}^{n-1}.$$

*Proof.* We decompose $u = u_1 e_1 + \tilde{u}$, where $\tilde{u} \cdot e_1 = 0$. The result then easily follows from Lemma A.14. $\square$

**Lemma A.15.** *For any $0 < \epsilon \le \frac{1}{2}$, there exists $N_0 = N_0(\epsilon) > 0$, $C_1 = C_1(\epsilon) > 0$, $C_2 = C_2(\epsilon) > 0$, such that if $m \ge C_1 n$, then the following hold with probability at least $1 - \frac{C_2}{m^2}$: For any $N \ge N_0$, we have*

$$\frac{1}{m}\sum_{j=1}^m |a_j \cdot e_1|^3 |a_j \cdot u| \chi_{|a_j \cdot e_1| \ge N} \le \epsilon, \qquad \forall\, u \in \mathbb{S}^{n-1};$$

$$\frac{1}{m}\sum_{j=1}^m |a_j \cdot e_1|^3 |a_j \cdot u| \chi_{|a_j \cdot u| \ge N} \le \epsilon, \qquad \forall\, u \in \mathbb{S}^{n-1}.$$

*Proof.* We only sketch the proof. Write $u = (u \cdot e_1)e_1 + \tilde{u}$, where $\tilde{u} \in F = \{\tilde{u} \in \mathbb{R}^n : \ \|\tilde{u}\|_2 \le 1, \tilde{u} \cdot e_1 = 0\}$. For the first inequality, note that (observe $|u \cdot e_1| \le 1$)

$$\begin{aligned}&\frac{1}{m}\sum_{j=1}^m |a_j \cdot e_1|^3 |a_j \cdot u| \chi_{|a_j \cdot e_1| \ge N}\\ \le&\frac{1}{m}\sum_{j=1}^m |a_j \cdot e_1|^4 \chi_{|a_j \cdot e_1| \ge N} + \frac{1}{m}\sum_{j=1}^m |a_j \cdot e_1|^3 |a_j \cdot \tilde{u}| \chi_{|a_j \cdot e_1| \ge N}.\end{aligned}$$

For the first term one can use Lemma A.9. For the second term one can use Lemma A.14.

Now for the second inequality, we write

$$\begin{aligned}&\frac{1}{m}\sum_{j=1}^{m}|a_j\cdot e_1|^3|a_j\cdot u|\chi_{|a_j\cdot u|\geq N}\\ \leq\ &\underbrace{\frac{1}{m}\sum_{j=1}^{m}|a_j\cdot e_1|^3\chi_{|a_j\cdot e_1|\leq M}|a_j\cdot u|\chi_{|a_j\cdot u|\geq N}}_{=:H_1}+\underbrace{\frac{1}{m}\sum_{j=1}^{m}|a_j\cdot e_1|^3\chi_{|a_j\cdot e_1|>M}|a_j\cdot u|}_{=:H_2}.\end{aligned}$$

For $H_2$, by using the estimates already obtained in the beginning part of this proof, it is clear that we can take $M$ sufficiently large such that $H_2\leq\epsilon/2$.

After $M$ is fixed, we return to the estimate of $H_1$. The result then follows from Lemma A.12 by taking $N$ sufficiently large. Note that we can work with a smoothed cut-off function instead of the strict cut-off. $\square$

**Lemma A.16.** *Define*

$$F=\{u\in\mathbb{R}^n:\ \|u\|_2\leq 1,\quad u\cdot e_1=0\}.$$

*Suppose* $b=(b_1,\cdots,b_m)\in\mathbb{R}^m$ *satisfies*

$$\frac{1}{m}\sum_{j=1}^{m}b_j^2\lesssim 1,\qquad \frac{1}{\ln m}\max_{1\leq j\leq m}|b_j|\lesssim 1.$$

*For any* $0<\epsilon\leq\frac{1}{2}$, *if* $m\geq C\epsilon^{-1}n\ln n$, *then with probability at least* $1-\exp(-c\frac{m}{\ln m}\epsilon)$, *it holds that*

$$\Big|\frac{1}{m}\sum_{j=1}^{m}b_j(a_j\cdot u)(a_j\cdot v)-\mathrm{mean}\Big|\leq\epsilon,\qquad\forall\,u,v\in F.$$

*Proof.* Let $0<\delta<\frac{1}{2}$ and introduce a $\delta$-net $S_\delta$ on the set $F$. Note that the set $F$ can be identified as a unit ball in $\mathbb{R}^{n-1}$. We have $\mathrm{Card}(S_\delta)\leq(1+\frac{2}{\delta})^n$. Introduce the operator

$$A=\frac{1}{m}\sum_{j=1}^{m}b_j(a_ja_j^T-\mathrm{I}).$$

We have

$$\|A\|_{\mathrm{op}}=\sup_{x,y\in F}\langle Ax,y\rangle\leq\frac{1}{1-2\delta}\sup_{x,y\in S_\delta}\langle Ax,y\rangle.$$

By Lemma A.3, for each $u,v\in F$, we have

$$\mathbb{P}\Big(\Big|\frac{1}{m}\sum_{j=1}^{m}b_j(a_j\cdot u)(a_j\cdot v)-\mathrm{mean}\Big|>\frac{\epsilon}{4}\Big)\leq 2\exp(-c\min\{m\epsilon^2,\frac{m\epsilon}{\ln m}\})\leq 2\exp(-c\frac{m}{\ln m}\epsilon).$$

Thus

$$\mathbb{P}\Big(\sup_{u,v\in S_\delta}\Big|\frac{1}{m}\sum_{j=1}^{m}b_j(a_j\cdot u)(a_j\cdot v)-\mathrm{mean}\Big|>\frac{\epsilon}{4}\Big)\leq 2(1+\frac{2}{\delta})^n\exp(-c\frac{m}{\ln m}\epsilon).$$

Taking $\delta=\frac{1}{4}$ and $m\gtrsim n\ln n$ then yields the result. $\square$

*Corollary* A.4. Suppose $h_1$, $h_2 : \mathbb{R} \to \mathbb{R}$ are locally Lipschitz continuous functions such that

$$\sup_{z\in\mathbb{R}} \frac{|h_1(z)| + |h_2(z)|}{1+|z|} \lesssim 1,$$
$$\sup_{z\neq\tilde{z}} \frac{|h_i(z) - h_i(\tilde{z})|}{|z-\tilde{z}|} \lesssim 1, \quad i = 1, 2.$$

Define

$$F = \{u \in \mathbb{R}^n : \ \|u\|_2 \le 1, \quad u \cdot e_1 = 0\}.$$

Suppose $b = (b_1, \cdots, b_m) \in \mathbb{R}^m$ satisfies

$$\frac{1}{m}\sum_{j=1}^m b_j^2 \lesssim 1, \qquad \frac{1}{\ln m} \max_{1\le j\le m} |b_j| \lesssim 1.$$

For any $0 < \epsilon \le \frac{1}{2}$, there exist a constant $C_1 = C_1(\epsilon) > 0$, such that if $m \ge C_1 n \ln n$, then with probability at least $1 - \exp(-c\frac{m}{\ln m}\epsilon)$, it holds that

$$\Big|\frac{1}{m}\sum_{j=1}^m b_j h_1(a_j\cdot u) h_2(a_j \cdot v) - \text{mean}\Big| \le \epsilon, \qquad \forall\, u, v \in F.$$

*Proof.* Let $0 < \delta < \frac{1}{2}$ and introduce a $\delta$-net $S_\delta$ with $\mathrm{Card}(S_\delta) \le (1+\frac{2}{\delta})^n$ on the set $F$. As we shall see momentarily, we will need to take $\delta = O(\epsilon)$. By Lemma A.3, we have

$$\mathbb{P}(\sup_{u,v\in S_\delta}\Big|\frac{1}{m}\sum_{j=1}^m b_j h_1(a_j\cdot u)h_2(a_j\cdot v) - \text{mean}\Big| > \frac{\epsilon}{4}) \le 2(1+\frac{2}{\delta})^n \exp(-c\frac{m}{\ln m}\epsilon).$$

Now for any $u, v \in F$, $\tilde{u}, \tilde{v} \in S_\delta$ with $\|u - \tilde{u}\|_2 \le \delta$, $\|v - \tilde{v}\|_2 \le \delta$, we have

$$\begin{aligned}
&\Big|h_1(a_j\cdot u)h_2(a_j\cdot v) - h_1(a_j\cdot\tilde{u})h_2(a_j\cdot\tilde{v})\Big| \\
&\le |h_1(a_j\cdot u) - h_1(a_j\cdot\tilde{u})|\cdot|h_2(a_j\cdot v)| + |h_1(a_j\cdot\tilde{u})|\cdot|h_2(a_j\cdot v) - h_2(a_j\cdot\tilde{v})| \\
&\le K_0|a_j\cdot(u-\tilde{u})|(1+|a_j\cdot v|) + K_0(1+|a_j\cdot\tilde{u}|)|a_j\cdot(v-\tilde{v})| \\
&\le K_0\Big(\frac{1}{\delta}(a_j\cdot(u-\tilde{u}))^2 + \delta(a_j\cdot v)^2 + 2\delta + \frac{1}{\delta}(a_j\cdot(v-\tilde{v}))^2 + \delta(a_j\cdot\tilde{u})^2\Big),
\end{aligned}$$

where $K_0$ is an absolute constant. Now introduce the operator

$$A = \frac{1}{m}\sum_{j=1}^m |b_j|(a_j a_j^T - \mathrm{I}).$$

By Lemma A.16, with probability at least $1 - \exp(-c\frac{m}{\ln m})$, we have

$$\langle Ax, y\rangle \le 1, \qquad \forall\, x, y \in F.$$

Thus with the same probability, we have

$$\begin{aligned}
&\frac{1}{m}\sum_{j=1}^{m}|b_j|\cdot\Big|h_1(a_j\cdot u)h_2(a_j\cdot v)-h_1(a_j\cdot\tilde{u})h_2(a_j\cdot\tilde{v})\Big| \\
&\le K_0\Big(\frac{1}{\delta}\frac{1}{m}\sum_{j=1}^{m}|b_j|((a_j\cdot(u-\tilde{u}))^2+(a_j\cdot(v-\tilde{v}))^2) \\
&\qquad+\delta\frac{1}{m}\sum_{j=1}^{m}|b_j|((a_j\cdot v)^2+(a_j\cdot\tilde{u})^2)+2\delta\frac{1}{m}\sum_{j=1}^{m}|b_j|\Big) \\
&\le K_1\delta,
\end{aligned}$$

where $K_1>0$ is another absolute constant. It is also not difficult to control the differences in expectation, i.e. for some absolute constant $K_2>0$,

$$\begin{aligned}
&\frac{1}{m}\sum_{j=1}^{m}|b_j|\cdot\Big|\mathbb{E}h_1(a_j\cdot u)h_2(a_j\cdot v)-\mathbb{E}h_1(a_j\cdot\tilde{u})h_2(a_j\cdot\tilde{v})\Big| \\
&\le K_2\delta,
\end{aligned}$$

Now take $\delta=\frac{\epsilon}{4(K_1+K_2)}$ and the desired result clearly follows by taking $\frac{m}{\ln m}\gg n$. □

**Lemma A.17.** *For any $0<\epsilon\le\frac{1}{2}$, there are constants $C_1=C_1(\epsilon)>0$, $C_2=C_2(\epsilon)>0$, such that if $m\ge C_1 n\ln n$, then with probability at least $1-\frac{C_2}{m^2}$, the following hold:*

$$\Big|\frac{1}{m}\sum_{j=1}^{m}(a_j\cdot\xi)^2(a_j\cdot e_1)^2-\text{mean}\Big|\le\epsilon,\qquad\forall\,\xi\in\mathbb{S}^{n-1}.$$

*Proof.* Write $\xi=\xi_1e_1+u$, where $u\cdot e_1=0$. Then

$$\begin{aligned}
&\frac{1}{m}\sum_{j=1}^{m}(a_j\cdot\xi)^2(a_j\cdot e_1)^2 \\
&=\frac{1}{m}\sum_{j=1}^{m}(a_j\cdot e_1)^4(\xi_1)^2+2\frac{1}{m}\sum_{j=1}^{m}(a_j\cdot e_1)^3\cdot\xi_1\cdot(a_j\cdot u) \\
&\qquad+\frac{1}{m}\sum_{j=1}^{m}(a_j\cdot e_1)^2(a_j\cdot u)^2.
\end{aligned}$$

Clearly the first two terms can be easily handled by Lemma A.9 and Lemma A.14 respectively. For these terms we actually only need $m\ge Cn$. To handle the last term we need $m\gtrsim n\ln n$. The main observation is that $(a_j\cdot e_1)$ and $(a_j\cdot u)$ are independent.

Write $b_j=(a_j\cdot e_1)^2$ and observe that with probability at least $1-O(m^{-2})$, we have

$$\sum_{j=1}^{m}b_j^2\le 100m,\qquad\max_{1\le j\le m}b_j\le 100\ln m.$$

For $m\gtrsim n\ln n$, by using Lemma A.16, it holds with probability at least $1-O(m^{-2})-e^{-c\frac{m}{\ln m}\epsilon}=1-O(m^{-2})$ that

$$\Big|\frac{1}{m}\sum_{j=1}^{m}b_j((a_j\cdot u)^2-\|u\|_2^2)\Big|\le\frac{\epsilon}{2},\qquad\forall\,u\in F=\{v\in\mathbb{R}^n:\,\|v\|_2\le 1,\,v\cdot e_1=0\}.$$

By Lemma A.9, we have with probability $1-O(m^{-2})$,

$$\|u\|_2\Big|\frac{1}{m}\sum_{j=1}^m b_j-\text{mean}\Big|\le\Big|\frac{1}{m}\sum_{j=1}^m b_j-\text{mean}\Big|\le\frac{\epsilon}{2}.$$

The desired result then easily follows. □

## Appendix B. Alternative analysis for the regime $|s|\ll 1$

In this appendix, we outline an alternative analysis for the regime $|s|\ll 1$. The main (modest) goal is to show that when $m\gtrsim n\ln n$, there cannot appear any spurious critical points other than $u=\pm e_1$ in this regime. In the argument we pin-point exactly where the assumption $m\gtrsim n\ln n$ is needed.

**Theorem B.1.** *If $||\hat{u}\cdot e_1|-1|\ll 1$, then the only critical points of the loss function $f=f(u)$ in this regime are $\hat{u}=\pm e_1$ and $R=1$.*

The rest of this section is devoted to the proof of Theorem B.1. We shall focus on the angular part (i.e. $\hat{u}$). By using the characterization $R=B/A$ (see (2.2)), it is clear that if $\hat{u}=\pm e_1$, it must follow that $R=1$. Thus it suffices to localize $\hat{u}$.

To alleviate the notation we shall drop the hat and write $\hat{u}$ as $u$.

We use the parametrization $u=se_1^{\perp}+\sqrt{1-s^2}e_1$ (the regime $u=se_1^{\perp}-\sqrt{1-s^2}e_1$ can be similarly treated) and consider

$$F(s)=(\frac{1}{m}\sum_{j=1}^m(a_j\cdot u)^2(a_j\cdot e_1)^2)(\frac{1}{m}\sum_{j=1}^m(a_j\cdot e_1)(a_j\cdot u)^3)-(\frac{1}{m}\sum_{j=1}^m(a_j\cdot u)^4)(\frac{1}{m}\sum_{j=1}^m(a_j\cdot u)(a_j\cdot e_1)^3).$$

By (2.3), any critical point must satisfy $F(s)=0$.

Observe that $F(0)=0$.

**Lemma B.1.** *We have $F'(0)=0$.*

*Proof.* First observe that

$$\frac{1}{m}\sum_{j=1}^m(a_j\cdot u)^2(a_j\cdot e_1)^2=\underbrace{\frac{1}{m}\sum_{j=1}^m(a_j\cdot e_1)^4}_{=:B_1}+2s\cdot\underbrace{\frac{1}{m}\sum_{j=1}^m(a_j\cdot e_1)^3(a_j\cdot e_1^{\perp})}_{=:B_2}+O(s^2),$$

where $O(s^2)$ denotes terms of order $s^2$ and higher. Similarly

$$\frac{1}{m}\sum_{j=1}^m(a_j\cdot e_1)(a_j\cdot u)^3=B_1+3sB_2+O(s^2),$$

$$\frac{1}{m}\sum_{j=1}^m(a_j\cdot u)^4=B_1+4sB_2+O(s^2),$$

$$\frac{1}{m}\sum_{j=1}^m(a_j\cdot u)(a_j\cdot e_1)^3=B_1+sB_2+O(s^2).$$

Thus $F(s)=O(s^2)$ and $F'(0)=0$. □

*Proof of Theorem B.1.* We shall argue by contradiction. Suppose $\tilde{u} = \sqrt{\tilde{R}}w$, $w \in \mathbb{S}^{n-1}$ is another critical point in the small neighborhood of $e_1$ and for some $0 < s_1 \ll 1$,

$$w = s_1 e_1^{\perp} + \sqrt{1 - s_1^2} e_1.$$

Now we consider only for $|s| \le s_1$ with $u = s e_1^{\perp} + \sqrt{1 - s^2} e_1$. We shall show that $F''(s) \le -c_1 < 0$ and this will lead to a contradiction. ( Since $F(0) = F'(0) = 0$, we obtain $F(s) \le -\frac{c_1}{2} s^2$. But by (2.3), we must have $F(s_1) = 0$. Hence the contradiction.)

Denote $u' = \frac{d}{ds} u$ and $u'' = \frac{d^2}{ds^2} u$. Then

$$F''(s) = (\frac{2}{m}\sum_{j=1}^m (a_j \cdot u)(a_j \cdot u'')(a_j \cdot e_1)^2)(\frac{1}{m}\sum_{j=1}^m (a_j \cdot u)^3 (a_j \cdot e_1)) \tag{B.1}$$

$$+ (\frac{2}{m}\sum_{j=1}^m (a_j \cdot u')^2 (a_j \cdot e_1)^2)(\frac{1}{m}\sum_{j=1}^m (a_j \cdot u)^3 (a_j \cdot e_1)) \tag{B.2}$$

$$+ (\frac{4}{m}\sum_{j=1}^m (a_j \cdot u)(a_j \cdot u')(a_j \cdot e_1)^2)(\frac{3}{m}\sum_{j=1}^m (a_j \cdot u)^2 (a_j \cdot u')(a_j \cdot e_1)) \tag{B.3}$$

$$+ (\frac{1}{m}\sum_{j=1}^m (a_j \cdot u)^2 (a_j \cdot e_1)^2)(\frac{3}{m}\sum_{j=1}^m (a_j \cdot u)^2 (a_j \cdot u'')(a_j \cdot e_1)) \tag{B.4}$$

$$+ (\frac{1}{m}\sum_{j=1}^m (a_j \cdot u)^2 (a_j \cdot e_1)^2)(\frac{6}{m}\sum_{j=1}^m (a_j \cdot u)(a_j \cdot u')^2 (a_j \cdot e_1)) \tag{B.5}$$

$$- (\frac{4}{m}\sum_{j=1}^m (a_j \cdot u)^3 (a_j \cdot u''))(\frac{1}{m}\sum_{j=1}^m (a_j \cdot u)(a_j \cdot e_1)^3) \tag{B.6}$$

$$- (\frac{12}{m}\sum_{j=1}^m (a_j \cdot u)^2 (a_j \cdot u')^2)(\frac{1}{m}\sum_{j=1}^m (a_j \cdot u)(a_j \cdot e_1)^3) \tag{B.7}$$

$$- (\frac{8}{m}\sum_{j=1}^m (a_j \cdot u)^3 (a_j \cdot u'))(\frac{1}{m}\sum_{j=1}^m (a_j \cdot u')(a_j \cdot e_1)^3) \tag{B.8}$$

$$- (\frac{1}{m}\sum_{j=1}^m (a_j \cdot u)^4)(\frac{1}{m}\sum_{j=1}^m (a_j \cdot u''(a_j \cdot e_1)^3). \tag{B.9}$$

Observe that $u' = e_1^{\perp} - \frac{s}{\sqrt{1-s^2}} e_1$ and $u'' = -(1-s^2)^{-\frac{3}{2}} e_1$. The sign in (B.9) is NOT favorable for us and this is where we need the contradiction argument.

Clearly

$$u = \frac{s}{s_1}(w - \sqrt{1 - s_1^2} e_1) + \sqrt{1 - s^2} e_1 = \frac{s}{s_1} w + (\sqrt{1 - s^2} - \frac{s}{s_1}\sqrt{1 - s_1^2}) e_1 =: \alpha w + \beta e_1.$$

Note that $\alpha$ and $\beta$ are $O(1)$ constants.

To deal with the sum $S = \frac{1}{m}\sum_{j=1}^m (a_j \cdot u)^4$, we can rewrite it as

$$\begin{aligned} S &= \frac{1}{m}\sum_{j=1}^m (\alpha a_j \cdot w + \beta a_j \cdot e_1)^4 \\ &= \frac{1}{m}\sum_{j=1}^m (\alpha^4 (a_j \cdot w)^4 + 4\alpha^3 \beta (a_j \cdot w)^3 (a_j \cdot e_1) + 6\alpha^2 \beta^2 (a_j \cdot w)^2 (a_j \cdot e_1)^2 + 4\alpha \beta^3 (a_j \cdot w)(a_j \cdot e_1)^3 + \beta^4 (a_j \cdot e_1)^4). \end{aligned}$$

By using the strong estimates of $w$ established in Theorem 2.3, we can then estimate all these terms up to some controllable error! Thus (B.9) is under control. Other terms are similarly controlled as long as $u'$ does not appear.

The remaining task is to control the other terms which contain $u' = e_1^\perp - \frac{s}{\sqrt{1-s^2}} e_1$. Our main good term is (B.7) which has a favorable sign.

`Estimate of` (B.8). We take $\eta > 0$ small and bound it as:

$$\eta (a_j \cdot u)^2 (a_j \cdot u')^2 + C_\eta (a_j \cdot u)^4.$$

Note that the factor $(\frac{1}{m}\sum_{j=1}^m (a_j \cdot u')(a_j \cdot e_1)^3)$ is well under control and very small. The first term can be controlled using (B.7) whereas the second term is bounded (since we already estimated the sum $S$ earlier). So (B.8) is controllable. Here we only need $m \gtrsim n$.

`Estimate of (B.2).` This is the main term where we need the assumption $m \gtrsim n \log n$. Here we have to control

$$\frac{1}{m}\sum_{j=1}^m (a_j \cdot e_1^\perp)^2 (a_j \cdot e_1)^2$$

uniformly in $e_1^\perp$. For this we need $m \gtrsim n \log n$. Note that we cannot use (B.7): $(a_j \cdot u)^2 (a_j \cdot u')^2 \approx (a_j \cdot u)^2 (a_j \cdot e_1^\perp)^2$ (the error in this approximation is $s^2 (a_j \cdot u)^2 (a_j \cdot e_1)^2$ which is controllable in both ways) to kill this term.

*Remark.* A very tempting idea is to adopt an $s$-dependent cut-off. This appears to work for each individual $s$. However one need to show union bounds also in $s$ which is not trivial under the mere assumption $m \gtrsim n$.

`Estimate of (B.3).` Note that

$$\begin{aligned}(a_j \cdot u)(a_j \cdot u')(a_j \cdot e_1)^2 &= (a_j \cdot u)(a_j \cdot e^\perp)(a_j \cdot e_1)^2 - (a_j \cdot u)(a_j \cdot e_1)^3 \frac{s}{\sqrt{1-s^2}} \\ &= s(a_j \cdot e^\perp)^2 (a_j \cdot e_1)^2 + \sqrt{1-s^2}(a_j \cdot e_1)^3 (a_j \cdot e_1^\perp) - (a_j \cdot u)(a_j \cdot e_1)^3 \frac{s}{\sqrt{1-s^2}}.\end{aligned}$$

Clearly here again we need $m \gtrsim n \log n$ to deal with the term $(a_j \cdot e_1^\perp)^2 (a_j \cdot e_1)^2$. The other terms are easily controlled under the assumption $m \gtrsim n$.

Note that $(a_j \cdot u)(a_j \cdot u')(a_j \cdot e_1)^2$ is small (can be made to be as small as possible by shrinking $s$ and so on). For the other factor $(a_j \cdot u)^2 (a_j \cdot u')(a_j \cdot e_1)$ we simply bound it as

$$(a_j \cdot u)^2 (a_j \cdot u')(a_j \cdot e_1) \le \frac{1}{2}(a_j \cdot u)^2 (a_j \cdot u')^2 + \frac{1}{2}(a_j \cdot u)^2 (a_j \cdot e_1)^2.$$

The smallness of the pre-factor $(a_j \cdot u)(a_j \cdot u')(a_j \cdot e_1)^2$ helps to bound the overall term. Thus (B.3) is OK.

`Estimate of (B.5).` Here we estimate it as

$$(a_j \cdot u)(a_j \cdot u')^2 (a_j \cdot e_1) \le \frac{1}{2}(a_j \cdot u)^2 (a_j \cdot u')^2 + \frac{1}{2}(a_j \cdot e_1)^2 (a_j \cdot u')^2.$$

This is nearly a sharp estimate. The first term can be dealt with by using (B.7). To control the second term we need the crucial assumption $m \gtrsim n \log n$.

All the terms can be treated and we have $F''(s) \leq -c_1 < 0$, arriving at a contradiction. $\square$

## Appendix C. Auxiliary estimates for Section 3

*Proof of Lemma 3.4.* Recall that

$$\mathbb{E}\sqrt{\beta + X_t^2}\sqrt{\beta\rho^2 + X_1^2} =: h_\infty(\rho, t).$$

Since $\rho \sim 1$ we shall slightly abuse notation and write $h_\infty(\rho, t)$ simply as $h(t)$ in this proof.

Denote $g(x) = \sqrt{\beta + x^2}$. Clearly

$$g' = \frac{x}{\sqrt{\beta + x^2}}, \quad g'' = \beta(\beta + x^2)^{-\frac{3}{2}}.$$

Since $X_t = tX_1 + \sqrt{1-t^2}Y_1$, where $X_1$ and $Y_1$ are independent standard 1D Gaussian random variables, we clearly have

$$h(t) = \int g(tx + \sqrt{1-t^2}y)k(x)\rho_1(x,y)dxdy,$$

where

$$k(x) = \sqrt{\beta\rho^2 + x^2}, \qquad \rho_1(x,y) = \frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}.$$

Observe that

$$\begin{aligned}
&\partial_x(g(tx + \sqrt{1-t^2}y)) = g' \cdot t;\\
&\partial_y(g(tx + \sqrt{1-t^2}y)) = g' \cdot \sqrt{1-t^2};\\
&\partial_t(g(tx + \sqrt{1-t^2}y)) = g' \cdot (x - \frac{t}{\sqrt{1-t^2}}y) = \frac{1}{\sqrt{1-t^2}}(x\partial_y g - y\partial_x g).
\end{aligned}$$

The third identity is the key to obtaining cancellation when calculating $h'(t)$ and $h''(t)$.

Observe that

$$(x\partial_y - y\partial_x)\rho_1 \equiv 0, \qquad \forall\, x, y \in \mathbb{R}.$$

Then clearly

$$\begin{aligned}
h'(t) &= \frac{1}{\sqrt{1-t^2}}\int \Big((x\partial_y - y\partial_x)(g(tx + \sqrt{1-t^2}y))\Big)k(x)\rho_1(x,y)dxdy\\
&= \frac{1}{\sqrt{1-t^2}}\int g(tx + \sqrt{1-t^2}y) \cdot yk'(x)\rho_1 dxdy\\
&= \frac{1}{\sqrt{1-t^2}}2\int_{x>0,y>0}(g(tx + \sqrt{1-t^2}y) - g(tx - \sqrt{1-t^2}y)) \cdot yk'(x)\rho_1 dxdy.
\end{aligned}$$

Clearly then

$$\begin{aligned}
&h'(t) > 0, \quad 0 < t < 1;\\
&h'(t) < 0, \quad -1 < t < 0.
\end{aligned}$$

Moreover,

$$|h'(t)| \leq 4\int \|g'\|_\infty y^2 \cdot \rho_1 dxdy \lesssim 1, \quad \forall\, |t| < 1.$$

Note that we can actually obtain $|h'(t)| \lesssim 1$ for all $|t| \leq 1$.

On the other hand, for $0 < t < 1$,

$$
\begin{aligned}
h'(t) &\gtrsim \int_{x>0,y>0} \frac{tyx}{\sqrt{\beta+(tx+\sqrt{1-t^2}y)^2}+\sqrt{\beta+(tx-\sqrt{1-t^2}y)^2}} \cdot y \cdot \frac{x}{\sqrt{\beta\rho^2+x}} \cdot e^{-\frac{x^2+y^2}{2}} dxdy \\
&\gtrsim \int_{x\sim 10,y\sim 10} \frac{tyx}{\sqrt{\beta+(tx+\sqrt{1-t^2}y)^2}+\sqrt{\beta+(tx-\sqrt{1-t^2}y)^2}} \cdot y \cdot \frac{x}{\sqrt{\beta\rho^2+x}} \cdot e^{-\frac{x^2+y^2}{2}} dxdy \\
&\gtrsim t.
\end{aligned}
$$

Note that the implied constants here are allowed to depend on $\beta$. Similarly one can show $-h'(t) \gtrsim |t|$ for $-1 < t < 0$.

Next we treat $h''(t)$ in the regime $|t| \ll 1$. Observe that

$$
\frac{d}{dt}\Big(\frac{1}{\sqrt{1-t^2}}\Big) = O(t), \qquad \frac{1}{\sqrt{1-t^2}} = 1 + O(t^2).
$$

Thus

$$
\begin{aligned}
h''(t) &= O(t) + \frac{1}{\sqrt{1-t^2}} \int g'(tx+\sqrt{1-t^2}y)(x - \frac{t}{\sqrt{1-t^2}}y)yk'(x)\rho_1 dxdy \\
&= O(t) + \int g'(tx+\sqrt{1-t^2}y)(x - \frac{t}{\sqrt{1-t^2}}y)yk'(x)\rho_1 dxdy.
\end{aligned}
$$

Observe that the contribution of $\frac{t}{\sqrt{1-t^2}}y$ is bounded by $O(t)$. Then

$$
h''(t) = O(t) + \int g'(tx+\sqrt{1-t^2}y)y\frac{x^2}{\sqrt{\beta\rho^2+x^2}}\rho_1 dxdy.
$$

Denote $x_t = tx + \sqrt{1-t^2}y$. Then $y = \frac{x_t - tx}{\sqrt{1-t^2}}$. The contribution due to $\frac{tx}{\sqrt{1-t^2}}$ is also $O(t)$. Thus

$$
h''(t) = O(t) + \int g'(x_t)\frac{x_t}{\sqrt{1-t^2}}\frac{x^2}{\sqrt{\beta\rho^2+x^2}}\rho_1 dxdy.
$$

Note that $g'(z)z = \frac{z^2}{\sqrt{\beta+z^2}}$. Thus for $|t| \ll 1$, if $9 \le y \le 11$, $\frac{1}{4} \le x \le \frac{1}{2}$, then $x_t \sim 1$, and the main term is $O(1)$. Thus

$$
h''(t) \gtrsim 1
$$

for all $|t| \ll 1$. $\square$

If we take the limit $\beta \to 0$ in $h(\rho,t) = \mathbb{E}\sqrt{\beta+X_t^2}\sqrt{\beta\rho^2+X_1^2}$. Then we obtain the expression

$$
\mathbb{E}|X_t||X_1|.
$$

Understanding this limiting case is of some importance for the case $\beta > 0$. The following proposition gives a very explicit characterization.

**Proposition C.1.** *We have for $|t| < 1$:*

$$
\begin{aligned}
&\frac{\pi}{2}\mathbb{E}|X_t||X_1| = (\frac{\pi}{2} - \arccos t)t + \sqrt{1-t^2}; \\
&(\frac{\pi}{2}\mathbb{E}|X_t||X_1|)' = \frac{\pi}{2} - \arccos t; \\
&(\frac{\pi}{2}\mathbb{E}|X_t||X_1|)'' = \frac{1}{\sqrt{1-t^2}}.
\end{aligned}
$$

*Proof.* We recall $X_t = tX + \sqrt{1-t^2}Y$ where $X = X_1 \sim \mathcal{N}(0,1)$, $Y \sim \mathcal{N}(0,1)$ are independent. Without loss of generality we can assume $0 \le t < 1$. Denote $t = \sin\theta_0$ where $0 \le \theta_0 < \frac{\pi}{2}$. Then by using polar coordinates, we have

$$\begin{aligned}
\frac{\pi}{2}\mathbb{E}|X_tX_1| &= \frac{1}{4}\int_{\mathbb{R}^2} |tx + \sqrt{1-t^2}y| \cdot |x| \cdot e^{-\frac{x^2+y^2}{2}} dxdy \\
&= \frac{1}{4}\int_0^{2\pi} |\sin(\theta+\theta_0)| \cdot |\cos\theta| d\theta \cdot \int_0^\infty r^3 e^{-\frac{r^2}{2}} dr \\
&= \frac{1}{2}\int_0^{2\pi} |\sin(\theta+\theta_0)| \cdot |\cos\theta| d\theta \\
&= \int_0^{\pi} |\sin(\theta+\theta_0)| \cdot |\cos\theta| d\theta.
\end{aligned}$$

Now observe that

$$\begin{aligned}
&\int_0^{\pi} |\sin(\theta+\theta_0)| \cdot |\cos\theta| d\theta \\
=& \int_0^{\frac{\pi}{2}} \sin(\theta+\theta_0) \cdot \cos\theta d\theta - \int_{\frac{\pi}{2}}^{\pi-\theta_0} \sin(\theta+\theta_0) \cdot \cos\theta d\theta \\
&+ \int_{\pi-\theta_0}^{\pi} \sin(\theta+\theta_0) \cdot \cos\theta d\theta.
\end{aligned}$$

The desired result then easily follows by an explicit computation. $\square$

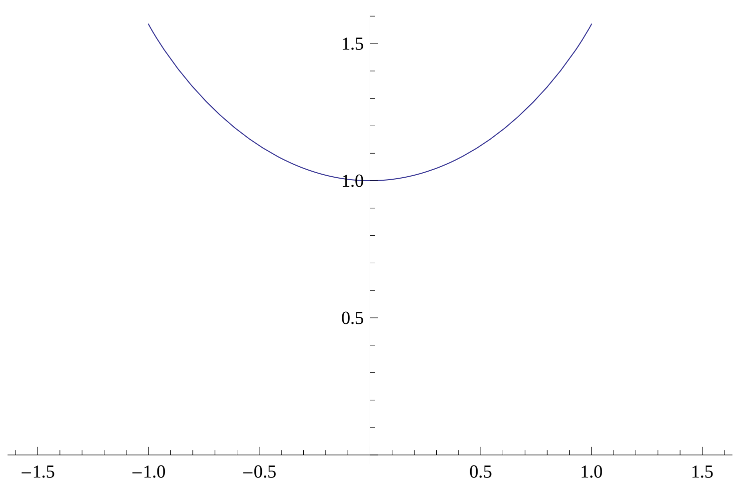



FIGURE 5. The function $(\frac{\pi}{2} - \arccos t)t + \sqrt{1-t^2}$

*Remark.* One may wonder why at $t = 1$, the derivative is formally given by $\frac{\pi}{2}$ instead of being zero since $u = x$ should be a critical point. The reason is due to the artificial singularity introduced by our representation. To see this, one can consider the regular variable $t = \cos\theta$ with $\theta \in [0,\pi]$, then

$$f(\theta) = (\frac{\pi}{2} - \theta)\cos\theta + \sin\theta;$$

Then clearly $f'(0) = 0$ and $f''(0) = -\frac{\pi}{2} < 0$. On the other hand,

$$f(\theta) = \tilde{f}(t) = \tilde{f}(\cos\theta).$$

Then

$$f'(\theta) = \tilde{f}'(\cos\theta)(-\sin\theta).$$

Thus

$$\tilde{f}'(1) = \lim_{\theta\to 0} \frac{f'(\theta)}{-\sin\theta} = -f''(0) > 0.$$

**Lemma C.1.** *Let $\psi : \mathbb{R} \to \mathbb{R}$ be continuously differentiable such that*

$$\sup_{z\in\mathbb{R}} |\psi(z)| + \sup_{z\in\mathbb{R}} \sqrt{1+z^2}|\psi'(z)| \lesssim 1.$$

*Let*

$$\psi_0(R, z) = z\sqrt{R+z^2}, \quad z \in \mathbb{R},\, c_1 \le R \le c_2,$$

*where $0 < c_1 < c_2 < \infty$ are two fixed constants. For any $0 < \epsilon \le 1$, if $m \gtrsim n$, then the following hold with high probability:*

$$\Big| \frac{1}{m} \sum_{j=1}^m \psi(a_j \cdot u)\psi_0(R, a_j \cdot e_1) - \text{mean} \Big| \le \epsilon, \qquad \forall\, u \in \mathbb{S}^{n-1}, \quad \forall\, c_1 \le R \le c_2.$$

*Proof.* Step 1. Let $0 < \eta < \frac{1}{2}$ be a constant whose value will be chosen sufficiently small. Let $\phi \in C_c^\infty(\mathbb{R})$ be such that $0 \le \phi(x) \le 1$ for all $x$, $\phi(x) = 1$ for $|x| \le 1$, and $\phi(x) = 0$ for $|x| \ge 2$. Denote

$$\langle x \rangle = \sqrt{1+x^2}, \qquad x \in \mathbb{R}.$$

Consider first the piece

$$\begin{aligned}
I_1 &= \frac{1}{m} \sum_{j=1}^m \psi(a_j \cdot u)\phi(\frac{a_j \cdot u}{\eta\langle a_j \cdot e_1\rangle})\psi_0(R, a_j \cdot e_1) \\
&= \frac{1}{m} \sum_{j=1}^m \psi(a_j \cdot u)\phi(\frac{a_j \cdot u}{\eta\langle a_j \cdot e_1\rangle})\psi_0(R, a_j \cdot e_1)\phi(\frac{a_j \cdot e_1}{K}) \\
&\quad + \frac{1}{m} \sum_{j=1}^m \psi(a_j \cdot u)\phi(\frac{a_j \cdot u}{\eta\langle a_j \cdot e_1\rangle})\psi_0(R, a_j \cdot e_1)(1 - \phi(\frac{a_j \cdot e_1}{K})) \\
&=: \; I_{1,a} + I_{1,b}.
\end{aligned}$$

where $K = \eta^{-\frac{1}{6}}$.

Thanks to the cut-off $\phi(\frac{a_j \cdot e_1}{K})$, we have $|a_j \cdot e_1| \le 2K$ on its support. Thus

$$|\psi_0(R, a_j \cdot e_1)|\phi(\frac{a_j \cdot e_1}{K}) \lesssim K^2,$$

and

$$|I_{1,a}| \le \alpha_0 K^2 \frac{1}{m} \sum_{j=1}^m \phi(\frac{a_j \cdot u}{\eta\langle 2K\rangle}),$$

where $\alpha_0 > 0$ is an absolute constant. Clearly

$$K^2 \mathbb{E}\phi(\frac{a_j \cdot u}{\eta\langle 2K\rangle}) \le K^2 \frac{1}{\sqrt{2\pi}} \int e^{-\frac{x^2}{2}} \phi(\frac{x}{\eta\langle 2K\rangle}) dx$$
$$\lesssim \eta^{-\frac{1}{3}} \cdot \eta^{\frac{5}{6}} = \eta^{\frac{1}{2}}.$$

By Bernstein's inequality, we have with high probability,

$$\Big| \frac{1}{m} \sum_{j=1}^m \phi(\frac{a_j \cdot u}{\eta\langle 2K\rangle}) - \text{mean} \Big| \le \eta.$$

Thus for $\eta > 0$ sufficiently small,

$$|I_{1,a}| \lesssim \eta^{\frac{1}{2}} \le \frac{\epsilon}{10}.$$

For $I_{1,b}$, we have

$$|I_{1,b}| \le \alpha_1 \cdot \frac{1}{m} \sum_{j=1}^m (a_j \cdot e_1)^2 (1 - \phi(\frac{a_j \cdot e_1}{K})),$$

where $\alpha_1 > 0$ is a constant. Similar to the estimate in $I_{1,a}$, we have with high probability,

$$\alpha_1 \cdot \frac{1}{m} \sum_{j=1}^m (a_j \cdot e_1)^2 (1 - \phi(\frac{a_j \cdot e_1}{K})) \lesssim \eta \le \frac{\epsilon}{10}.$$

Thus with high probability, it holds that for sufficiently small $\eta$,

$$|I_1| \le \frac{\epsilon}{5},$$

By a simple estimate we have $|\mathbb{E} I_1| \le \frac{\epsilon}{5}$ for sufficiently small $\eta$. Thus

$$|I_1 - \mathbb{E} I_1| \le \frac{2\epsilon}{5}.$$

Step 2. We now consider the main piece

$$I_2(u, R) = \frac{1}{m} \sum_{j=1}^m \psi(a_j \cdot u) \cdot \Big(1 - \phi(\frac{a_j \cdot u}{\eta\langle a_j \cdot e_1\rangle})\Big) \cdot \psi_0(R, a_j \cdot e_1).$$

Note that $\eta > 0$ is fixed in step 1. For simplicity we denote

$$X_j = a_j \cdot e_1, \quad h_1(z) = \psi(z) \cdot \Big(1 - \phi(\frac{z}{\eta\langle X_j\rangle})\Big).$$

Thanks to the cut-off $1 - \phi(\frac{z}{\eta\langle X_j\rangle})$ and the fact that $|\psi'(z)| \lesssim \langle z\rangle^{-1}$, we have

$$|h_1'(z)| \lesssim \eta^{-1} \langle X_j\rangle^{-1} \lesssim \langle X_j\rangle^{-1},$$

where in the last inequality we have included $\eta^{-1}$ into the implied constant. Since in this step $\eta > 0$ is a fixed constant this will not cause any problem. Clearly then

$$|h_1(z) - h_1(\tilde{z})| \lesssim \langle X_j\rangle^{-1} |z - \tilde{z}|, \qquad \forall\, z, \tilde{z} \in \mathbb{R}. \tag{C.1}$$

Also

$$|\psi_0(R, X_j) - \psi_0(\tilde{R}, X_j)| \le |X_j| \cdot |R - \tilde{R}|^{\frac{1}{2}}, \qquad \forall\, c_1 \le R,\ \tilde{R} \le c_2. \tag{C.2}$$

We shall need these important estimates below.

Let $\delta > 0$ be a small constant whose smallness will be specified later. We choose a $\delta$-net $F_\delta$ covering the set $\mathbb{S}^{n-1} \times \{R:\, c_1 \le R \le c_2\}$. We endow the set $S^{n-1} \times \{R:\, c_1 \le R \le c_2\}$ with the simple metric:

$$d((u,R),(\tilde{u},\tilde{R})) = \|u-\tilde{u}\|_2 + |R-\tilde{R}|.$$

Note that

$$\mathrm{Card}(F_\delta) \le \exp(C_\delta n),$$

where $C_\delta > 0$ depends only on $\delta$. By Bernstein's inequality, we have for any $0 < \eta_1 \le \frac{1}{2}$,

$$\mathbb{P}\Big(\sup_{(u,R)\in F_\delta} \Big| I_2(u,R) - \mathrm{mean}\Big| \ge \eta_1\Big) \le 2e^{C_\delta n} \cdot e^{-c\eta_1^2 m}.$$

Thus with high probability and taking $\eta_1 = \frac{\epsilon}{10}$, we have

$$|I_2(u,R) - \mathbb{E}I_2(u,R)| \le \frac{\epsilon}{10}, \qquad \forall\, (u,R) \in F_\delta.$$

Now let $(u,R) \in F_\delta$, and consider any $(\tilde{u},\tilde{R})$ such that

$$\|u-\tilde{u}\|_2 + |R-\tilde{R}| \le \delta.$$

By using the estimates (C.1), (C.2), we have

$$\begin{aligned}
&|I_2(u,R) - I_2(\tilde{u},\tilde{R})| \\
&= \Big| \frac{1}{m}\sum_{j=1}^m (h_1(a_j\cdot u)\psi_0(R,X_j) - h_1(a_j\cdot\tilde{u})\psi_0(\tilde{R},X_j))\Big| \\
&\le \frac{1}{m}\sum_{j=1}^m |h_1(a_j\cdot u) - h_1(a_j\cdot\tilde{u})||\psi_0(R,X_j)| + \frac{1}{m}\sum_{j=1}^m |h_1(a_j\cdot\tilde{u})||\psi_0(R,X_j) - \psi_0(\tilde{R},X_j)| \\
&\le \alpha_2 \frac{1}{m}\sum_{j=1}^m \langle X_j\rangle^{-1} |a_j\cdot(u-\tilde{u})||X_j|\cdot\langle X_j\rangle + \alpha_2\frac{1}{m}\sum_{j=1}^m |R-\tilde{R}|^{\frac{1}{2}}|X_j|\langle X_j\rangle \\
&\le \alpha_2\frac{1}{m}\sum_{j=1}^m \Big(\frac{1}{2}|a_j\cdot(u-\tilde{u})|^2\cdot\delta^{-1} + \frac{1}{2}\delta|X_j|^2\Big) + \alpha_2|R-\tilde{R}|^{\frac{1}{2}}\frac{1}{m}\sum_{j=1}^m (|X_j|^2+1),
\end{aligned}$$

where $\alpha_2 > 0$ is an absolute constant. By Bernstein's inequality, it holds with high probability that

$$\frac{1}{m}\sum_{j=1}^m |a_j\cdot v|^2 \le 2, \qquad \forall\, v \in \mathbb{S}^{n-1}.$$

Thus

$$|I_2(u,R) - I_2(\tilde{u},\tilde{R})| \le 10\alpha_2(\delta + \delta^{\frac{1}{2}}).$$

Also it is easy to check that

$$|\mathbb{E}(I_2(u,R) - I_2(\tilde{u},\tilde{R}))| \le 10\alpha_2(\delta + \delta^{\frac{1}{2}}).$$

Therefore

$$|I_2(u,R) - \mathbb{E}I_2(u,R) - (I_2(\tilde{u},\tilde{R}), -\mathbb{E}I_2(\tilde{u},\tilde{R})| \le 20\alpha_2(\delta + \delta^{\frac{1}{2}}).$$

Now take $\delta$ such that

$$20\alpha_2(\delta + \delta^{\frac{1}{2}}) \le \frac{\epsilon}{10}.$$

We then obtain (with high probability)

$$|I_2(u,R) - \mathbb{E}I_2(u,R)| \le \frac{\epsilon}{5}, \qquad \forall\, u \in S^{n-1}, \forall\, c_1 \le R \le c_2.$$

Together with the estimate of $I_1$ in step 1, we obtain the desired conclusion. $\square$

**Lemma C.2.** *Let $\psi : \mathbb{R} \to \mathbb{R}$ be Lipschitz continuous such that*

$$\sup_{z\in\mathbb{R}} \frac{|\psi(z)|}{1+|z|} + \sup_{z\ne\tilde{z}\in\mathbb{R}} \frac{|\psi(z)-\psi(\tilde{z})|}{|z-\tilde{z}|} \lesssim 1.$$

*Let $0 < c_1 < c_2 < \infty$ be two fixed constants. For any $0 < \epsilon \le 1$, if $m \gtrsim n$, then the following hold with high probability:*

$$\Big|\frac{1}{m}\sum_{j=1}^m \psi(a_j\cdot u)\sqrt{R+(a_j\cdot e_1)^2} - \text{mean}\Big| \le \epsilon, \qquad \forall\, u \in \mathbb{S}^{n-1}, \quad \forall c_1 \le R \le c_2.$$

*Proof.* The main point is use a $\delta$-covering of the set $\mathbb{S}^{n-1} \times \{R: \, c_1 \le R \le c_2\}$. Note that

$$\begin{aligned}
&\Big|\psi(a_j\cdot u)\sqrt{R+(a_j\cdot e_1)^2} - \psi(a_j\cdot \tilde{u})\sqrt{\tilde{R}+(a_j\cdot e_1)^2}\Big| \\
&\le |\psi(a_j\cdot u) - \psi(a_j\cdot\tilde{u})|\sqrt{R+(a_j\cdot e_1)^2} + |\psi(a_j\cdot\tilde{u})|\cdot|\sqrt{R+(a_j\cdot e_1)^2} - \sqrt{\tilde{R}+(a_j\cdot e_1)^2}| \\
&\lesssim |a_j\cdot(u-\tilde{u})|(1+|a_j\cdot e_1|) + (1+|a_j\cdot\tilde{u}|)\cdot|R-\tilde{R}|^{\frac{1}{2}}.
\end{aligned}$$

The argument is then similar to that in Lemma C.1. We omit details. $\square$

Consider

$$h = \frac{1}{m}\sum_{j=1}^m \sqrt{\beta+(a_j\cdot\hat{u})^2}\cdot\sqrt{R+X_j^2},$$

where

$$\begin{aligned}
&\hat{u} = te_1 + \sqrt{1-t^2}e^{\perp}, \quad |t|<1, \, e^{\perp}\cdot e_1 = 0, e^{\perp} \in \mathbb{S}^{n-1}; \\
&X_j = a_j \cdot e_1, \quad 0 < c_1 \le R \le c_2 < \infty.
\end{aligned}$$

In the above we take $c_1 > 0$, $c_2 > 0$ as two fixed constants. In our original model, $R = \beta\rho^2$ and $\rho \sim 1$, and therefore this assumption is quite natural. In the lemma below we shall study $h$ in the regime

$$|t| \le 1 - \epsilon_0,$$

where $0 < \epsilon_0 \ll 1$. The smallness of $\epsilon_0$ will be needed later when we study the regime $||\hat{u}\cdot e_1| - 1| \ll 1$. Here we we shall show that away from $|t| = 1$ we have good control of $h$.

**Lemma C.3.** *Let $0 < \epsilon_0 \ll 1$ be fixed. For any $0 < \epsilon \le 1$, if $m \gtrsim n$, then with high probability it holds that*

$$|\partial_t h - \mathbb{E}\partial_t h| \le \epsilon, \qquad \forall\, |t| \le 1-\epsilon_0, \, e^{\perp}\cdot e_1 = 0, e^{\perp}\in\mathbb{S}^{n-1}, c_1 \le R \le c_2.$$

*Proof of Lemma C.3.* Denote $g(x) = \sqrt{\beta + x^2}$ and

$$Z_j = a_j \cdot \hat{u} = tX_j + \sqrt{1-t^2}Y_j, \qquad Y_j = a_j \cdot e^{\perp}.$$

Clearly

$$\begin{aligned}\frac{d}{dt}Z_j &= X_j - \frac{t}{\sqrt{1-t^2}}Y_j \\ &= \frac{1}{1-t^2}X_j - \frac{t}{1-t^2}Z_j.\end{aligned}$$

Therefore

$$\begin{aligned}\partial_t h &= \frac{1}{1-t^2} \cdot \frac{1}{m}\sum_{j=1}^m g'(Z_j)X_j\sqrt{R+X_j^2} - \frac{t}{1-t^2} \cdot \frac{1}{m}\sum_{j=1}^m g'(Z_j)Z_j\sqrt{R+X_j^2} \\ &=: \frac{1}{1-t^2}H_1 - \frac{t}{1-t^2}H_2.\end{aligned}$$

By Lemma C.1, it holds with high probability that

$$|H_1 - \mathbb{E}H_1| \le (1-\epsilon_0^2) \cdot \frac{\epsilon}{3}, \qquad \forall\, \hat{u} \in \mathbb{S}^{n-1}, c_1 \le R \le c_2.$$

For $H_2$, we observe that

$$g'(x)x = \frac{x^2}{\sqrt{\beta + x^2}}.$$

By Lemma C.2, it then holds with high probability that

$$|H_2 - \mathbb{E}H_2| \le (1-\epsilon_0^2) \cdot \frac{\epsilon}{3}, \qquad \forall\, \hat{u} \in \mathbb{S}^{n-1}, c_1 \le R \le c_2.$$

The desired result then easily follows. □

**Lemma C.4.** *Let $0 < \epsilon_0 \ll 1$ be fixed. For any $0 < \epsilon \le \frac{1}{2}$, if $m \gtrsim n$, then with high probability it holds that*

$$\partial_{tt}h \ge \mathbb{E}\partial_{tt}h - \epsilon, \qquad \forall\, |t| \le 1-\epsilon_0,\, e^{\perp} \cdot e_1 = 0, e^{\perp} \in \mathbb{S}^{n-1}, c_1 \le R \le c_2.$$

*Furthermore, it holds with probability at least $1 - O(m^{-2})$ that*

$$|\partial_{tt}h - \mathbb{E}\partial_{tt}h| \le \epsilon, \qquad \forall\, |t| \le 1-\epsilon_0,\, e^{\perp} \cdot e_1 = 0, e^{\perp} \in \mathbb{S}^{n-1}, c_1 \le R \le c_2.$$

*Proof of Lemma C.4.* We adopt the same notation as in Lemma C.3.

Observe that

$$\begin{aligned}\partial_{tt}h &= \frac{1}{m}\sum_{j=1}^m g'(Z_j)\frac{d^2}{dt^2}Z_j\sqrt{R+X_j^2} + \frac{1}{m}\sum_{j=1}^m g''(Z_j)\Big(\frac{d}{dt}Z_j\Big)^2\sqrt{R+X_j^2} \\ &=: H_1 + H_2.\end{aligned}$$

We first deal with $H_1$. Note that

$$\frac{d^2}{dt^2}Z_j = -(1-t^2)^{-\frac{3}{2}}Y_j.$$

Since

$$Y_j = \frac{1}{\sqrt{1-t^2}}(Z_j - tX_j),$$

we obtain

$$H_1 = -(1-t^2)^{-2}\frac{1}{m}\sum_{j=1}^m g'(Z_j)Z_j\sqrt{R+X_j^2} + \frac{t}{(1-t^2)^2}\frac{1}{m}\sum_{j=1}^m g'(Z_j)X_j\sqrt{R+X_j^2}.$$

By similar estimates as in Lemma C.3, we have with high probability,

$$\begin{aligned}
&\Big|\frac{1}{m}\sum_{j=1}^m g'(Z_j)Z_j\sqrt{R+X_j^2} - \text{mean}\Big| \le \frac{\epsilon}{20}\cdot(1-\epsilon_0^2)^2, \qquad \forall\, u\in\mathbb{S}^{n-1},\, c_1\le R\le c_2;\\
&\Big|\frac{1}{m}\sum_{j=1}^m g'(Z_j)X_j\sqrt{R+X_j^2} - \text{mean}\Big| \le \frac{\epsilon}{20}\cdot(1-\epsilon_0^2)^2, \qquad \forall\, u\in\mathbb{S}^{n-1},\, c_1\le R\le c_2.
\end{aligned}$$

Thus

$$|H_1 - \mathbb{E}H_1| \le \frac{\epsilon}{10}, \qquad \forall\, |t|\le 1-\epsilon_0,\, e^\perp\cdot e_1 = 0, e^\perp\in\mathbb{S}^{n-1}, c_1\le R\le c_2.$$

Next we deal with $H_2$. Observe that

$$g''(x) = \beta(\beta+x^2)^{-\frac{3}{2}} > 0.$$

Let $\phi\in C_c^\infty(\mathbb{R})$ be such that $0\le\phi(x)\le 1$ for all $x$, $\phi(x)=1$ for $|x|\le 1$ and $\phi(x)=0$ for $|x|\ge 2$. Then

$$H_2 = H_3 + \frac{1}{m}\sum_{j=1}^m g''(Z_j)\cdot\Big(\frac{d}{dt}Z_j\Big)^2\cdot\Big(1-\phi(\frac{Z_j}{\eta\langle X_j\rangle})\Big)\sqrt{R+X_j^2},$$

where $H_3\ge 0$ is given by

$$H_3 = \frac{1}{m}\sum_{j=1}^m g''(Z_j)\cdot\Big(\frac{d}{dt}Z_j\Big)^2\cdot\phi(\frac{Z_j}{\eta\langle X_j\rangle})\sqrt{R+X_j^2}.$$

We first show that if $0<\eta\le\frac{1}{2}$ is taken sufficiently small, then

$$\mathbb{E}H_3 \le \frac{\epsilon}{20}; \tag{C.3}$$

and with probability at least $1-O(m^{-2})$ ,

$$H_3 \le \frac{\epsilon}{20}, \qquad \forall\, |t|\le 1-\epsilon_0, \quad e^\perp\cdot e_1 = 0, e^\perp\in\mathbb{S}^{n-1}, c_1\le R\le c_2. \tag{C.4}$$

Here we stress that since $H_3\ge 0$, if we only care about the lower bound, we can just discard it in order to obtain a high-in-probability statement. On the other hand, to get a two-way bound of $H_3$, we need to work with weaker statements due to the high-moment terms (i.e. more than quadratic) of $X_j$ in $H_3$.

Recall that

$$\begin{aligned}
\Big|\frac{d}{dt}Z_j\Big| &= \Big|X_j - \frac{t}{\sqrt{1-t^2}}Y_j\Big| = \Big|\frac{1}{1-t^2}X_j - \frac{t}{1-t^2}Z_j\Big|\\
&\lesssim |X_j| + |Z_j|.
\end{aligned}$$

We have

$$
\begin{aligned}
H_3 &\lesssim \frac{1}{m}\sum_{j=1}^m \phi(\frac{Z_j}{\eta\langle X_j\rangle})\langle Z_j\rangle^{-3}\cdot(X_j^2+Z_j^2)\cdot(1+|X_j|)\\
&\lesssim \frac{1}{m}\sum_{j=1}^m \phi(\frac{Z_j}{\eta\langle X_j\rangle})(1+|X_j|^3).
\end{aligned}
$$

Let $K=\eta^{-\frac{1}{8}}$. Then

$$
H_3 \lesssim K^3\frac{1}{m}\sum_{j=1}^m \phi(\frac{Z_j}{\eta\langle 2K\rangle})+\frac{1}{m}\sum_{j=1}^m(1+|X_j|^3)(1-\phi(\frac{X_j}{K})).
$$

Clearly then for $\eta>0$ sufficiently small,

$$
\begin{aligned}
\mathbb{E}H_3 &\lesssim K^3\mathbb{E}\phi(\frac{Z_1}{\eta\langle 2K\rangle})+\mathbb{E}(1+|X_1|^3)(1-\phi(\frac{X_1}{K}))\\
&\le \frac{\epsilon}{20}.
\end{aligned}
$$

For $m\gtrsim n$, it holds with high probability that

$$
\Big|\frac{1}{m}\sum_{j=1}^m \phi(\frac{Z_j}{\eta\langle 2K\rangle})-\mathrm{mean}\Big|\le\eta.
$$

On the other hand, by Lemma A.9, it holds with probability at least $1-O(m^{-2})$ that

$$
\Big|\frac{1}{m}\sum_{j=1}^m(1+|X_j|^3)(1-\phi(\frac{X_j}{K}))-\mathrm{mean}\Big|\le\eta,\qquad \forall\, u\in\mathbb{S}^{n-1}.
$$

Thus for $\eta>0$ sufficiently small, (C.3) and (C.4) hold.

Now we consider the main piece

$$
H_4=\frac{1}{m}\sum_{j=1}^m g''(Z_j)\cdot\Big(\frac{d}{dt}Z_j\Big)^2\cdot\Big(1-\phi(\frac{Z_j}{\eta\langle X_j\rangle})\Big)\sqrt{R+X_j^2}.
$$

By using

$$
\begin{aligned}
\Big(\frac{d}{dt}Z_j\Big)^2 &= \Big(\frac{1}{1-t^2}X_j-\frac{t}{1-t^2}Z_j\Big)^2\\
&=\frac{1}{(1-t^2)^2}(X_j^2+t^2Z_j^2-2tX_jZ_j).
\end{aligned}
$$

Then

$$
\begin{aligned}
H_4 &= (1-t^2)^{-2}\frac{1}{m}\sum_{j=1}^m g''(Z_j)\Big(1-\phi(\frac{Z_j}{\eta\langle X_j\rangle})\Big)X_j^2\sqrt{R+X_j^2}\\
&\quad+(1-t^2)^{-2}t^2\frac{1}{m}\sum_{j=1}^m g''(Z_j)Z_j^2\Big(1-\phi(\frac{Z_j}{\eta\langle X_j\rangle})\Big)\sqrt{R+X_j^2}\\
&\quad-2t(1-t^2)^{-2}\frac{1}{m}\sum_{j=1}^m g''(Z_j)Z_j\Big(1-\phi(\frac{Z_j}{\eta\langle X_j\rangle})\Big)X_j\sqrt{R+X_j^2}.
\end{aligned}
$$

Define

$$
h_j(x)=g''(x)\cdot(1-\phi(\frac{x}{\eta\langle X_j\rangle})).
$$

Clearly, thanks to the cut-off $1-\phi$, we have

$$\|h_j\|_\infty \lesssim \langle X_j \rangle^{-3};$$
$$\|h_j'\|_\infty \lesssim \langle X_j \rangle^{-4}.$$

It is then easy to check that the summands in $H_4$ are bounded. Moreover

$$|h_j(a_j \cdot u) - h_j(a_j \cdot \tilde{u})| \lesssim |a_j \cdot (u-\tilde{u})| \cdot \langle X_j \rangle^{-4}.$$

Similar bounds also hold for the other summands in $H_4$. Thus by a similar union bound argument as in Lemma C.1 (and taking care of the covering in the $R$-variable), we have with high probability that

$$|H_4 - \mathbb{E}H_4| \le \frac{\epsilon}{20}, \qquad \forall\, |t| \le 1-\epsilon_0, e^\perp \cdot e_1 = 0, e^\perp \in \mathbb{S}^{n-1}, c_1 \le R \le c_2.$$

Collecting all the estimates, we then obtain the desired estimate for $\partial_{tt} h$. □

**Lemma C.5.** *Let $X \sim \mathcal{N}(0,1)$, $Y \sim \mathcal{N}(0,1)$ be independent. Define*

$$H(\rho, s) = \mathbb{E}\sqrt{\beta + (\sqrt{1-s^2}X + sY)^2}\sqrt{\beta\rho^2 + X^2};$$
$$h(\rho, s) = \frac{1}{2}(1+2\beta)\rho^2 - \rho H(\rho, s).$$

*Then it holds that*

$$\sup_{|\rho-1|\ll 1, |s| \ll 1} \sum_{j=1}^{3} (|\partial^j H| + |\partial^j h|) \lesssim 1.$$

*where $\partial = \partial_\rho$ or $\partial_s$.*

*Proof.* Clearly it suffices for us to prove the estimate for $H$ since the estimate for $h$ will follow from it.

We first deal with $\partial_{sss} H$ which appears to be the most difficult case and simultaneously $\partial_s H$, $\partial_{ss} H$. In some terms we shall even exhibit $(\beta,\, \rho)$-independent bounds which will be of interest for future investigations.

Denote $A = \sqrt{1-s^2}x + sy$. Then

$$\partial_s A = -\frac{s}{\sqrt{1-s^2}} x + y;$$
$$\partial_y A = s, \qquad \partial_x A = \sqrt{1-s^2};$$
$$\partial_s A = \frac{1}{\sqrt{1-s^2}}(-x\partial_y + y\partial_x) A.$$

Now we have

$$2\pi H = \int \sqrt{\beta + A^2}\sqrt{\beta\rho^2 + x^2} e^{-\frac{x^2+y^2}{2}} dx dy.$$

Since

$$\partial_s(\sqrt{\beta + A^2}) = \frac{1}{\sqrt{1-s^2}}(-x\partial_y + y\partial_x)(\sqrt{\beta + A^2}),$$
$$(-x\partial_y + y\partial_x)(e^{-\frac{x^2+y^2}{2}}) = 0,$$

we obtain (by using integration by parts) that

$$2\pi\partial_s H = \frac{1}{\sqrt{1-s^2}}\int \sqrt{\beta+A^2}(-y\partial_x)(\sqrt{\beta\rho^2+x^2})e^{-\frac{x^2+y^2}{2}}dxdy$$
$$= -\frac{1}{\sqrt{1-s^2}}\int \sqrt{\beta+A^2}\frac{xy}{\sqrt{\beta\rho^2+x^2}}e^{-\frac{x^2+y^2}{2}}dxdy.$$

Note that the pre-factor $\frac{1}{\sqrt{1-s^2}}$ is smooth in the regime $|s|\ll 1$, therefore to compute the higher order $\partial_s$-derivatives of $H$, it suffices for us to treat

$$H_1 = \int \sqrt{\beta+A^2}\frac{xy}{\sqrt{\beta\rho^2+x^2}}e^{-\frac{x^2+y^2}{2}}dxdy.$$

Then

$$\partial_s H_1 = \frac{1}{\sqrt{1-s^2}}\int \sqrt{\beta+A^2}(x\partial_y - y\partial_x)(\frac{xy}{\sqrt{\beta\rho^2+x^2}})e^{-\frac{x^2+y^2}{2}}dxdy.$$

The most difficult term is the piece corresponding to $y\partial_x$. Thus we consider

$$H_2 = \int \sqrt{\beta+A^2}(y\partial_x)(\frac{xy}{\sqrt{\beta\rho^2+x^2}})e^{-\frac{x^2+y^2}{2}}dxdy$$
$$= \int \sqrt{\beta+A^2}\cdot y^2\cdot \frac{\beta\rho^2}{(\beta\rho^2+x^2)^{\frac{3}{2}}}e^{-\frac{x^2+y^2}{2}}dxdy.$$

Thus

$$\partial_s H_2 = \int \frac{A}{\sqrt{\beta+A^2}}(-\frac{s}{\sqrt{1-s^2}}x+y)y^2\cdot \frac{\beta\rho^2}{(\beta\rho^2+x^2)^{\frac{3}{2}}}e^{-\frac{x^2+y^2}{2}}dxdy.$$

The piece corresponding to $-\frac{s}{\sqrt{1-s^2}}x$ is clearly fine. So we only need to treat

$$H_3 = \int \frac{A}{\sqrt{\beta+A^2}}y^3\cdot \frac{\beta\rho^2}{(\beta\rho^2+x^2)^{\frac{3}{2}}}e^{-\frac{x^2+y^2}{2}}dxdy.$$

Observe that for $\eta>0$,

$$\int_{\mathbb{R}} \frac{\eta^2}{(\eta^2+x^2)^{\frac{3}{2}}}dx = \int_{\mathbb{R}} \frac{1}{(1+x^2)^{\frac{3}{2}}}dx.$$

Thus $H_3$ is bounded by an absolute constant. Collecting the estimates, we have

$$|\partial_{sss}H| \lesssim 1.$$

Now we deal with $\partial_\rho H$, $\partial_{\rho\rho}H$, and $\partial_{\rho\rho\rho}H$. This case is easy. Denote $B=\beta\rho^2+x^2$. Then

$$\partial_\rho(\sqrt{B}) = B^{-\frac{1}{2}}\beta\rho;$$
$$\partial_{\rho\rho}(\sqrt{B}) = B^{-\frac{1}{2}}\beta - B^{-\frac{3}{2}}\beta^2\rho^2;$$
$$\partial_{\rho\rho\rho}(\sqrt{B}) = -B^{-\frac{3}{2}}\beta^2\rho + 3B^{-\frac{5}{2}}(\beta\rho)^3 - B^{-\frac{3}{2}}\beta^2 2\rho.$$

Clearly all terms are bounded and we have

$$|\partial_\rho H| + |\partial_{\rho\rho}H| + |\partial_{\rho\rho\rho}H| \lesssim 1.$$

Next clearly $\partial_{\rho s}H$ and $\partial_{\rho\rho s}H$ are OK.

We only need to treat $\partial_{\rho ss}H$. The main term of $\partial_{ss}H$ is

$$H_4 = \int \frac{A}{\sqrt{\beta + A^2}}(-\frac{s}{\sqrt{1-s^2}}x + y)\frac{xy}{\sqrt{\beta\rho^2 + x^2}}e^{-\frac{x^2+y^2}{2}}dxdy.$$

Now

$$\partial_\rho H_4 = -\int \frac{A}{\sqrt{\beta + A^2}}(-\frac{s}{\sqrt{1-s^2}}x + y)\frac{xy\beta\rho}{(\beta\rho^2 + x^2)^{\frac{3}{2}}}e^{-\frac{x^2+y^2}{2}}dxdy.$$

Clearly for any $0 < \eta \lesssim 1$,

$$\int \frac{\eta^2|x|}{(\eta^2 + x^2)^{\frac{3}{2}}}dx = \eta \int \frac{|x|}{(1 + x^2)^{\frac{3}{2}}}dx < \infty.$$

Thus $\partial_{\rho ss}H$ is also OK for us. □

**Lemma C.6** (Calculation of $\partial^2 h$ at $(\rho = 1,\, s = 0)$)**.** *Let*

$$H(\rho, s) = \mathbb{E}\sqrt{\beta + (\sqrt{1-s^2}X + sY)^2}\sqrt{\beta\rho^2 + X^2};$$
$$h(\rho, s) = \frac{1}{2}(1 + 2\beta)\rho^2 - \rho H(\rho, s).$$

*Then at $\rho = 1$, $s = 0$, we have*

$$(\partial_{\rho\rho}H)(1, 0) = -\gamma_1 < 0, \quad (\partial_{\rho s}H)(1, 0) = 0;$$
$$(\partial_{ss}H)(1, 0) = -\gamma_2 < 0;$$
$$(\partial_s h)(\rho, 0) = 0, \forall\, \rho > 0, \quad (\partial_\rho h)(1, 0) = 0;$$
$$(\partial_{\rho\rho}h)(1, 0) = \gamma_3 > 0, \quad (\partial_{\rho s}h)(1, 0) = 0;$$
$$(\partial_{ss}h)(1, 0) = \gamma_4 > 0,$$

*where $\gamma_i > 0$, $i = 1, \cdots, 4$ are constants depending on $\beta$.*

*Proof.* `Calculation of` $\partial_{ss}H$.

Denote $A = \sqrt{1-s^2}x + sy$. Then

$$\partial_s A = -\frac{s}{\sqrt{1-s^2}}x + y, \quad \partial_y A = s, \qquad \partial_x A = \sqrt{1-s^2};$$
$$\partial_s A = \frac{1}{\sqrt{1-s^2}}(-x\partial_y + y\partial_x)A.$$

Now we have

$$2\pi H = \int \sqrt{\beta + A^2}\sqrt{\beta\rho^2 + x^2}e^{-\frac{x^2+y^2}{2}}dxdy.$$

Then

$$2\pi\partial_s H = \frac{1}{\sqrt{1-s^2}}\int \sqrt{\beta + A^2}(-y\partial_x)(\sqrt{\beta\rho^2 + x^2})e^{-\frac{x^2+y^2}{2}}dxdy$$
$$= -\frac{1}{\sqrt{1-s^2}}\int \sqrt{\beta + A^2}\frac{xy}{\sqrt{\beta\rho^2 + x^2}}e^{-\frac{x^2+y^2}{2}}dxdy.$$

One should observe that $\partial_s H\Big|_{\rho>0,s=0} = 0$.

Then

$$2\pi\partial_{ss}H\Big|_{\rho=1,s=0} = -\int \frac{A}{\sqrt{\beta+A^2}}\Big|_{s=0}\cdot\frac{xy^2}{\sqrt{\beta+x^2}}e^{-\frac{x^2+y^2}{2}}dxdy$$
$$= -\int \frac{x^2y^2}{\beta+x^2}e^{-\frac{x^2+y^2}{2}}dxdy.$$

<u>`Calculation of` $\partial_{\rho\rho}H$.</u> Clearly

$$2\pi\partial_\rho H = \int \sqrt{\beta+A^2}\frac{\beta\rho}{\sqrt{\beta\rho^2+x^2}}e^{-\frac{x^2+y^2}{2}}dxdy.$$

Observe that

$$2\pi\partial_\rho H\Big|_{\rho=1,s=0} = \int \beta e^{-\frac{x^2+y^2}{2}}dxdy.$$

Then

$$2\pi\partial_{\rho\rho}H\Big|_{\rho=1,s=0} = \int \sqrt{\beta+A^2}\Big|_{\rho=1,s=0}\frac{\beta}{\sqrt{\beta\rho^2+x^2}}e^{-\frac{x^2+y^2}{2}}dxdy$$
$$-\int \sqrt{\beta+A^2}\Big|_{\rho=1,s=0}\frac{\beta^2\rho^2}{(\beta\rho^2+x^2)^{\frac{3}{2}}}e^{-\frac{x^2+y^2}{2}}dxdy$$
$$= \int \frac{\beta x^2}{\beta+x^2}e^{-\frac{x^2+y^2}{2}}dxdy.$$

<u>`Calculation of` $\partial_{\rho s}H$.</u> We have

$$2\pi\partial_{\rho s}H\Big|_{\rho=1,s=0} = \int \frac{A}{\sqrt{\beta+A^2}}\Big|_{s=0}y\frac{\beta\rho}{\sqrt{\beta\rho^2+x^2}}e^{-\frac{x^2+y^2}{2}}dxdy$$
$$= \int \frac{\beta xy}{\beta+x^2}e^{-\frac{x^2+y^2}{2}}dxdy = 0.$$

Now we calculate the corresponding Hessian for $h = \frac{1}{2}(1+2\beta)\rho^2 - \rho H$. Clearly

$$\partial_{\rho\rho}h\Big|_{\rho=1,s=0} = 1+2\beta-(\partial_{\rho\rho}H+2\partial_\rho H) = 1+2\beta-\frac{1}{2\pi}\int(\frac{\beta x^2}{\beta+x^2}+2\beta)e^{-\frac{x^2+y^2}{2}}dxdy$$
$$= \frac{1}{2\pi}\int(1-\frac{\beta x^2}{\beta+x^2})e^{-\frac{x^2+y^2}{2}}dxdy.$$

By Lemma C.7 this is clearly positive and has a lower bound depending only in terms of $\beta$.

On the other hand,

$$\partial_{\rho s}h\Big|_{\rho=1,s=0} = -\partial_s H - \partial_{\rho s}H = 0.$$

Finally

$$\partial_{ss}h = -\partial_{ss}H = \frac{1}{2\pi}\int\frac{x^2y^2}{\beta+x^2}e^{-\frac{x^2+y^2}{2}}dxdy > 0.$$

□

**Lemma C.7.** *For any $0<\beta<\infty$, we have*

$$\int(1-\frac{\beta x^2}{\beta+x^2})e^{-\frac{x^2}{2}}dx > 0.$$

*Proof.* For $0 < \beta \le 1$, this is obvious. For $\beta > 1$, denote $\epsilon = \frac{1}{\beta}$. Then

$$\tilde{h}(\epsilon) = \int (1 - \frac{x^2}{1+\epsilon x^2}) e^{-\frac{x^2}{2}} dx;$$
$$\tilde{h}'(\epsilon) = \int \frac{\epsilon x^4}{(1+\epsilon x^2)^2} e^{-\frac{x^2}{2}} dx > 0.$$

Clearly $\tilde{h}(0) = 0$. Thus $\tilde{h}(\epsilon) > 0$ for all $0 < \epsilon < \infty$. □

**Lemma C.8.** *Let $0 < c_1 < c_2 < \infty$ be fixed. Consider for $\xi \in \mathbb{S}^{n-1}$, $u \in \mathbb{R}^n$ with $c_1 \le \|u\|_2 \le c_2$, the following:*

$$I_1 = I_1(\xi, u) = \frac{1}{m}\sum_{k=1}^m (\beta|u|^2 + (a_k \cdot u)^2)^{-\frac{3}{2}} (\beta|u|^2 + (a_k \cdot e_1))^{\frac{1}{2}} \cdot (a_k \cdot \xi)^2,$$
$$I_2 = I_2(\xi, u) = \frac{1}{m}\sum_{k=1}^m (\beta|u|^2 + (a_k \cdot u)^2)^{-\frac{3}{2}} (\beta|u|^2 + (a_k \cdot e_1))^{\frac{1}{2}} \cdot (a_k \cdot u) \cdot (a_k \cdot \xi).$$

*For any $0 < \epsilon \le 1$, if $m \gtrsim n$, then it holds with probability at least $1 - O(m^{-2})$ that*

$$|I_1 - \mathbb{E}I_1| \le \epsilon, \qquad \forall \xi \in \mathbb{S}^{n-1}, \quad \forall c_1 \le \|u\|_2 \le c_2;$$
$$|I_2 - \mathbb{E}I_2| \le \epsilon, \qquad \forall \xi \in \mathbb{S}^{n-1}, \quad \forall c_1 \le \|u\|_2 \le c_2.$$

*Proof of Lemma C.8.* We first note that, in order to prove the statement for $I_2$, it suffices for us to prove the statement for $I_1$ under a more general condition (instead of $\xi \in \mathbb{S}^{n-1}$):

$$\|\xi\|_2 \le c_3 := 2 + c_2.$$

The reason is as follows. By using the simple identity

$$(a_k \cdot (\xi + u))^2 = (a_k \cdot \xi)^2 + (a_k \cdot u)^2 + 2(a_k \cdot \xi)(a_k \cdot u),$$

we have

$$I_2(\xi, u) = \frac{1}{2} I_1(\xi + u, u) - \frac{1}{2} I_1(\xi, u) - I_3,$$

where

$$I_3 = \frac{1}{m}\sum_{k=1}^m (\beta|u|^2 + (a_k \cdot u)^2)^{-\frac{3}{2}} (a_k \cdot u)^2 (\beta|u|^2 + (a_k \cdot e_1)^2)^{\frac{1}{2}}.$$

Clearly $I_3$ is OK for union bounds and we have with high probability

$$|I_3 - \mathbb{I}_3| \le \epsilon, \qquad \forall c_1 \le \|u\|_2 \le c_2.$$

Thus to prove the statement for $I_2$ it suffices for us to prove it for $I_1$ uniformly in $\xi$ with $\|\xi\|_2 \le c_3$.

Next we observe that for $\xi \ne 0$ with $\|\xi\|_2 \le c_3$, we have

$$\begin{aligned} |I_1(\xi, u) - \mathbb{E}I_1(\xi, u)| \le & \ \|\xi\|_2 |I_1(\frac{\xi}{\|\xi\|_2}, u) - \mathbb{E}I_1(\frac{\xi}{\|\xi\|_2}, u)| \\ \le & \ c_3 |I_1(\frac{\xi}{\|\xi\|_2}, u) - \mathbb{E}I_1(\frac{\xi}{\|\xi\|_2}, u)|. \end{aligned}$$

Thus it suffices for us to prove the statement for $I_1$ under the original assumption $\xi \in \mathbb{S}^{n-1}$.

Now let $\phi \in C_c^\infty(\mathbb{R})$ be such that $0 \le \phi(x) \le 1$ for all $x$, $\phi(x) = 1$ for $|x| \le 1$ and $\phi(x) = 0$ for $|x| \ge 2$. Let $\delta > 0$ be a sufficiently small constant. The needed smallness will be specified later. We write (below $\langle x \rangle = (1+x^2)^{\frac{1}{2}}$)

$$\begin{aligned} I_1 &= \frac{1}{m}\sum_{k=1}^m (\beta|u|^2 + (a_k \cdot u)^2)^{-\frac{3}{2}} (\beta|u|^2 + (a_k \cdot e_1)^2)^{\frac{1}{2}} \cdot (a_k \cdot \xi)^2 \cdot \phi(\frac{a_k \cdot u}{\delta \langle a_k \cdot \xi \rangle}) \\ &\quad + \frac{1}{m}\sum_{k=1}^m (\beta|u|^2 + (a_k \cdot u)^2)^{-\frac{3}{2}} (\beta|u|^2 + (a_k \cdot e_1)^2)^{\frac{1}{2}} \cdot (a_k \cdot \xi)^2 \cdot \Bigl(1 - \phi(\frac{a_k \cdot u}{\delta \langle a_k \cdot \xi \rangle})\Bigr) \\ &=: I_{1,a} + I_{1,b}. \end{aligned}$$

<u>Estimate of $I_{1,a}$</u>. Let $K = \delta^{-\frac{1}{9}}$. Then

$$\begin{aligned} |I_{1,a}| &\le \frac{1}{m}\sum_{k=1}^m (\beta|u|^2 + (a_k \cdot u)^2)^{-\frac{3}{2}} (\beta|u|^2 + (a_k \cdot e_1)^2)^{\frac{1}{2}} \cdot (a_k \cdot \xi)^2 \cdot \phi(\frac{a_k \cdot u}{\delta \langle a_k \cdot \xi \rangle}) \phi(\frac{a_k \cdot \xi}{K}) \\ &\quad + \frac{1}{m}\sum_{k=1}^m (\beta|u|^2 + (a_k \cdot u)^2)^{-\frac{3}{2}} (\beta|u|^2 + (a_k \cdot e_1)^2)^{\frac{1}{2}} \cdot (a_k \cdot \xi)^2 \cdot \phi(\frac{a_k \cdot u}{\delta \langle a_k \cdot \xi \rangle}) \cdot \Bigl(1 - \phi(\frac{a_k \cdot \xi}{K})\Bigr) \\ &\lesssim K^2 \frac{1}{m}\sum_{k=1}^m (1 + (a_k \cdot e_1)^2)^{\frac{1}{2}} \phi(\frac{a_k \cdot u}{\delta \langle 2K \rangle}) \\ &\quad + \frac{1}{m}\sum_{k=1}^m (1 + (a_k \cdot e_1)^2)^{\frac{1}{2}} \cdot (a_k \cdot \xi)^2 \cdot \Bigl(1 - \phi(\frac{a_k \cdot \xi}{K})\Bigr) \\ &\lesssim K^5 \frac{1}{m}\sum_{k=1}^m \phi(\frac{a_k \cdot u}{\delta \langle 2K \rangle}) + \frac{1}{m}\sum_{k=1}^m (1 + (a_k \cdot e_1)^2) \cdot K^{-1} \\ &\quad + K \cdot \frac{1}{m}\sum_{k=1}^m (a_k \cdot \xi)^4 \cdot \Bigl(1 - \phi(\frac{a_k \cdot \xi}{K})\Bigr). \end{aligned}$$

Clearly for sufficiently small $\delta$, we have

$$\mathbb{E}|I_{1,a}| \le \frac{\epsilon}{10}.$$

Furthermore, with probability at least $1 - O(m^{-2})$, we have

$$|I_{1,a}| \le \frac{\epsilon}{10}, \qquad \forall\, c_1 \le \|u\|_2 \le c_2,\ \forall\, \xi \in \mathbb{S}^{n-1}.$$

<u>Estimate of $I_{1,b}$</u>. Thanks to the cut-off $1 - \phi(\frac{a_k \cdot u}{\delta \langle a_k \cdot \xi \rangle})$, we have $|a_k \cdot u| \gtrsim \langle a_k \cdot \xi \rangle$ on its support. It is then easy to check that the summands in $I_{1,b}$ are sub-exponential random variables. It remains for us to check the union bound.

To this end, take $u$, $\tilde{u}$ with $c_1 \le \|u\|_2, \|\tilde{u}\|_2 \le c_2$, and $\xi$, $\tilde{\xi} \in \mathbb{S}^{n-1}$. Then clearly

$$\begin{aligned} &\Bigl| (\beta|u|^2 + (a_k \cdot u)^2)^{-\frac{3}{2}} (\beta|u|^2 + (a_k \cdot e_1)^2)^{\frac{1}{2}} \cdot (a_k \cdot \xi)^2 \cdot \Bigl(1 - \phi(\frac{a_k \cdot u}{\delta \langle a_k \cdot \xi \rangle})\Bigr) \\ &\qquad - (\beta|\tilde{u}|^2 + (a_k \cdot \tilde{u})^2)^{-\frac{3}{2}} (\beta|\tilde{u}|^2 + (a_k \cdot e_1)^2)^{\frac{1}{2}} \cdot (a_k \cdot \tilde{\xi})^2 \cdot \Bigl(1 - \phi(\frac{a_k \cdot \tilde{u}}{\delta \langle a_k \cdot \tilde{\xi} \rangle})\Bigr) \Bigr| \\ &\lesssim (1 + |a_k \cdot e_1|) \cdot \Bigl( \|u - \tilde{u}\|_2 + |a_k \cdot (u - \tilde{u})| + |a_k \cdot (\xi - \tilde{\xi})| \Bigr). \end{aligned}$$

Here in the above derivation we have used the fact that the function (it differs from the actual one by some minor change of parameters)

$$G(t,s) = \langle t\rangle^{-3} s^2 \Big(1 - \phi(\frac{t}{\langle s\rangle})\Big)$$

satisfies

$$|G(t,s) - G(\tilde{t},\tilde{s})| \lesssim |t-\tilde{t}| + |s-\tilde{s}|.$$

It is then clear that $I_{1,b}$ is OK for union bounds and we have with high probability

$$|I_{1,b} - \mathbb{E}I_{1,b}| \le \frac{\epsilon}{10}, \qquad \forall\, c_1 \le \|u\|_2 \le c_2,\, \forall\, \xi \in S^{n-1}.$$

The desired estimate for $I_1$ then easily follows. $\square$

**Lemma C.9.** *Let $0 < c_1 < c_2 < \infty$ be fixed. Consider*

$$f_0(u) = -\frac{1}{m}\sum_{k=1}^m \sqrt{\beta|u|^2 + (a_k\cdot u)^2}\sqrt{\beta|u|^2 + (a_k\cdot e_1)^2}.$$

*For any $0 < \epsilon \le 1$, if $m \gtrsim n$, then it holds with probability at least $1 - O(m^{-2})$ that*

$$\Big|\sum_{i,j=1}^n \xi_i\xi_j(\partial_{ij}f_0)(u) - \sum_{i,j=1}^n \xi_i\xi_j\mathbb{E}(\partial_{ij}f_0)(u)\Big| \le \epsilon, \qquad \forall\, \xi \in \mathbb{S}^{n-1}, \quad \forall\, c_1 \le \|u\|_2 \le c_2.$$

*Proof of Lemma C.9.* To simplify the notation, write $a_k$ as $a$, and denote

$$\begin{aligned}
&A = \beta|u|^2 + (a\cdot u)^2, \qquad B = \beta|u|^2 + (a\cdot e_1)^2;\\
&\partial_i A = 2\beta u_i + 2(a\cdot u)a_i, \qquad \partial_{ij}A = 2\beta\delta_{ij} + 2a_ia_j;\\
&\partial_i B = 2\beta u_i, \qquad \partial_{ij}B = 2\beta\delta_{ij}.
\end{aligned}$$

We need to compute $\partial_{ij}\tilde{F}$ for

$$\tilde{F} = A^{\frac12}B^{\frac12}.$$

Clearly

$$\begin{aligned}
\partial_i\tilde{F} &= \frac12 A^{-\frac12}\partial_i A B^{\frac12} + \frac12 B^{-\frac12}\partial_i B A^{\frac12};\\
\partial_{ij}\tilde{F} &= -\frac14 A^{-\frac32}\partial_i A\partial_j A B^{\frac12}\\
&\quad + \frac12 A^{-\frac12}\partial_{ij}AB^{\frac12}\\
&\quad + \frac12 A^{-\frac12}\partial_i A\frac12 B^{-\frac12}\partial_j B\\
&\quad - \frac14 B^{-\frac32}\partial_j B\partial_i B A^{\frac12}\\
&\quad + \frac12 B^{-\frac12}\partial_{ij}BA^{\frac12}\\
&\quad + \frac14 B^{-\frac12}A^{-\frac12}\partial_i B\partial_j A.
\end{aligned}$$

We then have

$$\sum_{i,j=1}^{n} \xi_i \xi_j (\partial_{ij} f_0)(u)$$

$$= \frac{1}{4} \cdot \frac{1}{m} \sum_{k=1}^{m} A_k^{-\frac{3}{2}} |\xi \cdot \nabla A_k|^2 B_k^{\frac{1}{2}} \tag{C.5}$$

$$+ \frac{1}{4} \cdot \frac{1}{m} \sum_{k=1}^{m} B_k^{-\frac{3}{2}} |\xi \cdot \nabla B_k|^2 A_k^{\frac{1}{2}} \tag{C.6}$$

$$- \frac{1}{2} \cdot \frac{1}{m} \sum_{k=1}^{m} A_k^{-\frac{1}{2}} \langle \xi, (\nabla^2 A_k)\xi \rangle B_k^{\frac{1}{2}} \tag{C.7}$$

$$- \frac{1}{2} \cdot \frac{1}{m} \sum_{k=1}^{m} A_k^{-\frac{1}{2}} B_k^{-\frac{1}{2}} (\nabla A_k \cdot \xi)(\nabla B_k \cdot \xi) \tag{C.8}$$

$$- \frac{1}{2} \cdot \frac{1}{m} \sum_{k=1}^{m} B_k^{-\frac{1}{2}} A_k^{\frac{1}{2}} \langle \xi, (\nabla^2 B_k)\xi \rangle, \tag{C.9}$$

where $A_k = \beta |u|^2 + (a_k \cdot u)^2$, $B_k = \beta |u|^2 + (a_k \cdot e_1)^2$, and we have denoted

$$\langle \xi, (\nabla^2 A_k)\xi \rangle = \sum_{i,j=1}^{n} \xi_i \xi_j \partial_{ij} A_k.$$

Estimate of (C.9). We have

$$\frac{1}{m} \sum_{k=1}^{m} B_k^{-\frac{1}{2}} A_k^{\frac{1}{2}} \langle \xi, (\nabla^2 B_k)\xi \rangle$$
$$= 2\beta |\xi|^2 \Big( \frac{1}{m} \sum_{k=1}^{m} B_k^{-\frac{1}{2}} A_k^{\frac{1}{2}} \Big).$$

The summand consists of sub-exponential random variables and are clearly OK for union bounds. Thus with high probability, it holds that

$$\Big| \frac{1}{m} \sum_{k=1}^{m} B_k^{-\frac{1}{2}} A_k^{\frac{1}{2}} - \text{mean} \Big| \le \frac{\epsilon}{100(1+2\beta)}, \qquad \forall\, c_1 \le \|u\|_2 \le c_2.$$

Thus the contribution of (C.9) is OK for us.

Estimate of (C.6). We have

$$\frac{1}{m} \sum_{k=1}^{m} B_k^{-\frac{3}{2}} A_k^{\frac{1}{2}} |\xi \cdot \nabla B_k|^2$$
$$= 4\beta^2 (\xi \cdot u)^2 \Big( \frac{1}{m} \sum_{k=1}^{m} B_k^{-\frac{3}{2}} A_k^{\frac{1}{2}} \Big).$$

Again the summand consists of sub-exponential random variables and are clearly OK for union bounds. Thus with high probability, it holds that

$$\Big| \frac{1}{m} \sum_{k=1}^{m} B_k^{-\frac{3}{2}} A_k^{\frac{1}{2}} - \text{mean} \Big| \le \frac{\epsilon}{100(1+4\beta^2 c_2^2)}, \qquad \forall\, c_1 \le \|u\|_2 \le c_2.$$

Thus the contribution of (C.6) is OK for us.

Estimate of (C.8). We have

$$\begin{aligned}
&\frac{1}{m}\sum_{k=1}^m A_k^{-\frac12} B_k^{-\frac12}(\nabla A_k\cdot\xi)(\nabla B_k\cdot\xi)\\
=&\frac{1}{m}\sum_{k=1}^m A_k^{-\frac12} B_k^{-\frac12}\cdot\Big(2\beta(u\cdot\xi)+2(a_k\cdot u)(a_k\cdot\xi)\Big)2\beta(\xi\cdot u)\\
=&4\beta^2(\xi\cdot u)^2\frac{1}{m}\sum_{k=1}^m A_k^{-\frac12} B_k^{-\frac12}\\
&\quad+4\beta(\xi\cdot u)\frac{1}{m}\sum_{k=1}^m A_k^{-\frac12} B_k^{-\frac12}(a_k\cdot\xi)(a_k\cdot u). \qquad (C.10)
\end{aligned}$$

The first term is clearly under control and therefore we focus only on (C.10). For this observe that for any $u$, $\tilde u$ with $c_1\le\|u\|_2,\|\tilde u\|_2\le c_2$, $\xi$, $\tilde\xi\in\mathbb{S}^{n-1}$, it holds that

$$\begin{aligned}
&\Big|\frac{a_k\cdot u}{\sqrt{\beta|u|^2+|a_k\cdot u|^2}}-\frac{a_k\cdot\tilde u}{\sqrt{\beta|u|^2+|a_k\cdot\tilde u|^2}}\Big|\lesssim|a_k\cdot(u-\tilde u)|,\\
&\Big|\frac{a_k\cdot\tilde u}{\sqrt{\beta|u|^2+|a_k\cdot\tilde u|^2}}-\frac{a_k\cdot\tilde u}{\sqrt{\beta|\tilde u|^2+|a_k\cdot\tilde u|^2}}\Big|\lesssim\|u-\tilde u\|_2,\\
&\Big|\frac{a_k\cdot u}{\sqrt{\beta|u|^2+|a_k\cdot u|^2}}\cdot\frac{a_k\cdot\xi}{\sqrt{\beta|u|^2+|a_k\cdot e_1|^2}}-\frac{a_k\cdot\tilde u}{\sqrt{\beta|\tilde u|^2+|a_k\cdot\tilde u|^2}}\cdot\frac{a_k\cdot\tilde\xi}{\sqrt{\beta|\tilde u|^2+|a_k\cdot e_1|^2}}\Big|\\
&\qquad\lesssim(|a_k\cdot(u-\tilde u)|+\|u-\tilde u\|_2)|a_k\cdot\xi|+|a_k\cdot(\xi-\tilde\xi)|.
\end{aligned}$$

Thus (C.10) is OK for union bounds and we have with high probability,

$$\Big|\frac{1}{m}\sum_{k=1}^m A_k^{-\frac12} B_k^{-\frac12}(a_k\cdot\xi)(a_k\cdot u)-\mathrm{mean}\Big|\le\frac{\epsilon}{200(1+4\beta c_2)},\qquad\forall\, c_1\le\|u\|_2\le c_2,\,\forall\,\xi\in\mathbb{S}^{n-1}.$$

Thus (C.8) is under control.

Estimate of (C.7). We have

$$\begin{aligned}
&-\frac12\cdot\frac{1}{m}\sum_{k=1}^m A_k^{-\frac12}\langle\xi,(\nabla^2A_k)\xi\rangle B_k^{\frac12}\\
=&-\beta|\xi|^2\frac{1}{m}\sum_{k=1}^m A_k^{-\frac12}B_k^{\frac12}\\
&\quad-\frac{1}{m}\sum_{k=1}^m A_k^{-\frac12}B_k^{\frac12}(a_k\cdot\xi)^2. \qquad (C.11)
\end{aligned}$$

The first term is clearly under control. Therefore we only need to treat (C.11). We shall treat it together with (C.12) below.

Estimate of (C.5). We have

$$\begin{aligned}
&\frac{1}{4}\cdot\frac{1}{m}\sum_{k=1}^{m}A_k^{-\frac{3}{2}}|\xi\cdot\nabla A_k|^2B_k^{\frac{1}{2}}\\
&=\frac{1}{m}\sum_{k=1}^{m}A_k^{-\frac{3}{2}}B_k^{\frac{1}{2}}(a_k\cdot u)^2(a_k\cdot\xi)^2 &\text{(C.12)}\\
&\quad+2\beta(\xi\cdot u)\frac{1}{m}\sum_{k=1}^{m}A_k^{-\frac{3}{2}}B_k^{\frac{1}{2}}(a_k\cdot u)(a_k\cdot\xi) &\text{(C.13)}\\
&\quad+\beta^2(\xi\cdot u)^2\frac{1}{m}\sum_{k=1}^{m}A_k^{-\frac{3}{2}}B_k^{\frac{1}{2}}. &\text{(C.14)}
\end{aligned}$$

Clearly (C.14) is perfectly under control. Now observe

$$\text{(C.11)}+\text{(C.12)}=-\beta|u|^2\frac{1}{m}\sum_{k=1}^{m}A_k^{-\frac{3}{2}}B_k^{\frac{1}{2}}(a_k\cdot\xi)^2.$$

One can then apply Lemma C.8 to get the desired estimate for this term as well as (C.13). $\square$

## Appendix D. Technical estimates for Section 4

**Lemma D.1.** *Denote $X_j=a_j\cdot e_1$ and $Z_j=a_j\cdot\hat{u}$, where $\hat{u}\in\mathbb{S}^{n-1}$. For any $\epsilon>0$, there exists $R=R(\epsilon,\beta)>0$, such that if $m\gtrsim n$, then the following hold with high probability:*

$$\frac{1}{m}\sum_{j=1}^{m}(\beta+Z_j^2)\sqrt{\frac{\rho^2(\beta+2Z_j^2)}{\rho^2(\beta+Z_j^2)+X_j^2}}\le\epsilon,\qquad\forall\,0<\rho\le R,\quad\forall\,\hat{u}\in\mathbb{S}^{n-1}.$$

*Proof of Lemma D.1.* We shall only sketch the proof. Choose $\phi\in C_c^\infty(\mathbb{R})$ such that $0\le\phi(x)\le1$ for all $x$, $\phi(x)=1$ for $|x|\le1$ and $\phi(x)=0$ for $|x|\ge2$. Then

$$\begin{aligned}
&(\beta+Z_j^2)\sqrt{\frac{\rho^2(\beta+2Z_j^2)}{\rho^2(\beta+Z_j^2)+X_j^2}}\\
&\le(\beta+Z_j^2)\sqrt{\frac{\rho^2(\beta+2Z_j^2)}{\rho^2(\beta+Z_j^2)+X_j^2}}\phi(\frac{Z_j}{K}) &\text{(D.1)}\\
&\quad+(\beta+Z_j^2)\cdot\sqrt{2}\cdot\Big(1-\phi(\frac{Z_j}{K})\Big),
\end{aligned}$$

where $K>0$ is a constant to be specified momentarily. Clearly by taking $K$ sufficiently large, we have with high probability that

$$\frac{1}{m}\sum_{j=1}^{m}(\beta+Z_j^2)\cdot\sqrt{2}\cdot\Big(1-\phi(\frac{Z_j}{K})\Big)\le\frac{\epsilon}{10},\qquad\forall\,\hat{u}\in\mathbb{S}^{n-1}.$$

It then remains for us to deal with (D.1). Thanks to the smooth cut-off, we have

$$\begin{aligned}
\text{(D.1)}&\le\rho C_{K,\beta}\cdot\frac{1}{\sqrt{\rho^2\beta+X_j^2}}\\
&\le\rho C_{K,\beta}\cdot\frac{1}{\eta}+E_{K,\beta}\cdot\phi(\frac{X_j}{\eta}),
\end{aligned}$$

where $C_{K,\beta} > 0$, $E_{K,\beta} > 0$ are constants depending only on $K$ and $\beta$. We first choose $\eta > 0$ sufficiently small such that with high probability,

$$E_{K,\beta}\Big|\frac{1}{m}\sum_{j=1}^{m}\phi(\frac{X_j}{\eta})\Big| \le \frac{\epsilon}{10}.$$

Then the desired result follows by taking $\rho$ sufficiently small. $\square$

**Lemma D.2.** *Let* $\gamma_1 > 0$, $\gamma_2 > 0$ *and* $\gamma_3 \ge 0$. *Consider*

$$g(\theta_0) = \int_0^{\pi} \sqrt{\gamma_1 + \gamma_2\cos^2(\theta-\theta_0) + \gamma_3\sin^2\theta}\sqrt{\gamma_1 + 2\gamma_3\sin^2\theta}d\theta.$$

*Then*

$$g'(\theta_0) \ge 0, \qquad \forall\, \theta_0 \in [0, \frac{\pi}{2}).$$

*Furthermore, if* $\gamma_1 \sim 1$, $\gamma_2 \sim 1$, $\gamma_3 \ge 0$, *then*

$$g'(\theta_0) \gtrsim \frac{1}{1+\gamma_3}\sin 2\theta_0.$$

*In particular we have*

$$g''(0) \gtrsim \frac{1}{1+\gamma_3}.$$

*Remark* D.1. There exists a subtle balance of coefficients in the expression of $g(\theta_0)$ without which we cannot have the positivity of $g'$. As a counter-example, consider

$$f(s,b) = \int_0^{\pi} (1 + b\cos^2(\theta - s) + 2\sin^2\theta)^{\frac{1}{2}}(1+\sin^2\theta)^{\frac{1}{2}}d\theta.$$

One can check that $\partial_s f(s,b) < 0$ for $b < 1.99$ and $\partial_s f(s,b) > 0$ for some $b \ge 2$ and $s$.

*Proof of Lemma D.2.* Clearly

$$\begin{aligned} g'(\theta_0) &= \gamma_2 \int_0^{\pi} \sqrt{\frac{\gamma_1 + 2\gamma_3\sin^2\theta}{\gamma_1 + \gamma_2\cos^2(\theta-\theta_0) + \gamma_3\sin^2\theta}}\sin 2(\theta-\theta_0)d\theta \\ &= \gamma_2 \int_0^{\pi} \sqrt{\frac{\gamma_1 + 2\gamma_3\sin^2(\theta+\theta_0)}{\gamma_1 + \gamma_2\cos^2\theta + \gamma_3\sin^2(\theta+\theta_0)}}\sin 2\theta d\theta \\ &= \gamma_2 \int_0^{\frac{\pi}{2}} \Big(\sqrt{\frac{\gamma_1 + 2\gamma_3\sin^2(\theta+\theta_0)}{\gamma_1 + \gamma_2\cos^2\theta + \gamma_3\sin^2(\theta+\theta_0)}} - \sqrt{\frac{\gamma_1 + 2\gamma_3\sin^2(\theta-\theta_0)}{\gamma_1 + \gamma_2\cos^2\theta + \gamma_3\sin^2(\theta-\theta_0)}}\Big)\sin 2\theta d\theta. \end{aligned}$$

Clearly for $\theta, \theta_0 \in [0, \frac{\pi}{2})$, we have

$$\sin(\theta+\theta_0) \ge |\sin(\theta-\theta_0)|.$$

The non-negativity of $g'$ then follows from the monotonicity of the function (below $a \ge 1$ is a constant)

$$\tilde{g}(z) = \frac{\gamma_1 + 2\gamma_3 z}{a\gamma_1 + \gamma_3 z} = 2 - \frac{(2a-1)\gamma_1}{\gamma_3 z + a\gamma_1}, \quad z \ge 0.$$

Next if $\gamma_1, \gamma_2 \sim 1, \gamma_3 \ge 0$, then clearly (note that $a = 1 + \frac{\gamma_2}{\gamma_1}\cos^2\theta \ge 1$, $a \sim 1$)

$$\tilde{g}'(z) \gtrsim \frac{1}{1+\gamma_3}, \qquad \forall\, z \in [0,1].$$

Thus

$$\begin{aligned}
g'(\theta_0) &\gtrsim \frac{1}{1+\gamma_3}\int_0^{\frac{\pi}{2}}(\sin^2(\theta+\theta_0)-\sin^2(\theta-\theta_0))\sin 2\theta d\theta \\
&\gtrsim \frac{1}{1+\gamma_3}\int_0^{\frac{\pi}{2}}\sin^2(2\theta)d\theta\sin 2\theta_0 \\
&\gtrsim \frac{1}{1+\gamma_3}\sin 2\theta_0.
\end{aligned}$$

Since $g''(0)=\lim_{\theta_0\to 0+}\frac{g'(\theta_0)}{\theta_0}$, the estimate for $g''(0)$ easily follows. □

*Proof of Lemma 4.3.* Clearly

$$\begin{aligned}
h_\infty(\rho,t) &= \mathbb{E}\sqrt{\beta\rho^2+\rho^2X_t^2+X_1^2}\sqrt{\beta\rho^2+2X_1^2} \\
&= \frac{1}{2\pi}\int_{\mathbb{R}^2}\sqrt{\beta\rho^2+\rho^2(tx+\sqrt{1-t^2}y)^2+x^2}\sqrt{\beta\rho^2+2x^2}e^{-\frac{x^2+y^2}{2}}dxdy.
\end{aligned}$$

Since $\rho\sim 1$, it is easy to check that

$$\sup_{|t|\le 1-\eta_0}(|\partial_t h_\infty(\rho,t)|+|\partial_{tt}h_\infty(\rho,t)|+|\partial_{ttt}h_\infty(\rho,t)|\lesssim 1.$$

To show the lower bound on $|\partial_t h_\infty(\rho,t)|$, observe that $h_\infty(\rho,t)$ is an even function of $t$. Thus without loss of generality we assume $0\le t<1$. Now let $t=\sin\theta_0$ with $\theta_0\in[0,\frac{\pi}{2})$. By using polar coordinates, we obtain

$$\begin{aligned}
h_\infty(\rho,t) &= \frac{1}{2\pi}\int_0^\infty\int_0^{2\pi}\sqrt{\beta\rho^2+\rho^2\cos^2(\theta-\theta_0)+r^2\sin^2\theta}\sqrt{\beta\rho^2+2r^2\sin^2\theta}e^{-\frac{r^2}{2}}rd\theta dr \\
&= \frac{1}{\pi}\int_0^\infty\int_0^{\pi}\sqrt{\beta\rho^2+\rho^2\cos^2(\theta-\theta_0)+r^2\sin^2\theta}\sqrt{\beta\rho^2+2r^2\sin^2\theta}e^{-\frac{r^2}{2}}rd\theta dr.
\end{aligned}$$

Observe that

$$\partial_{\theta_0}\Big(h_\infty(\rho,\sin\theta_0)\Big)=(\partial_t h_\infty)(\rho,t)\Big|_{t=\sin\theta_0}\cos\theta_0. \tag{D.2}$$

By Lemma D.2 (note that $\gamma_3=r^2$) and integrating in $r$, we then obtain

$$\partial_t h_\infty(\rho,t)\gtrsim t, \qquad \forall\, 0\le t<1.$$

Finally to show that $\partial_{tt}h_\infty(\rho,t)\gtrsim 1$ for $|t|\ll 1$, it suffices for us to show (since $|\partial_{ttt}h_\infty(\rho,t)|\lesssim 1$ for $|t|\ll 1$)

$$\partial_{tt}h_\infty(\rho,0)\gtrsim 1.$$

By using (D.2), we only need to check

$$\partial_{\theta_0\theta_0}\Big(h_\infty(\rho,\sin\theta_0)\Big)\Big|_{\theta_0=0}\gtrsim 1.$$

This again follows from Lemma D.2. □

**Lemma D.3.** *Suppose $\phi_1:\mathbb{R}\to\mathbb{R}$, $\phi_2:\mathbb{R}\to\mathbb{R}$ are $C^1$ functions such that*

$$\max_{|z|\le L}(|\phi_1(z)|+|\phi_1'(z)|+|\phi_2(z)|+|\phi_2'(z)|)\le C_{L,\phi_1,\phi_2},$$

*where $C_{L,\phi_1,\phi_2}>0$ is finite for each finite $L$.*

*Suppose $0 < c_1 < c_2 < \infty$ and $\phi_3 : (\frac{c_1}{2}, 2c_2) \to \mathbb{R}$ is a smooth function such that*

$$\sup_{\frac{c_1}{2} < |z| < 2c_2} (|\phi_3(z)| + |\phi_3'(z)|) \le C_{c_1,c_2,\phi_3},$$

*where $C_{c_1,c_2,\phi_3} > 0$ depends only on $c_1$, $c_2$ and $\phi_3$.*

*Let $(d_{ij})_{1\le i\le 2, 1\le j\le 3}$ be given constants and consider*

$$I(u,w,v) = \frac{1}{m}\sum_{j=1}^m \phi_1\Big(\frac{d_{11}|u|^2 + d_{12}(a_j\cdot e_1)^2 + d_{13}(a_j\cdot u)^2}{\beta|u|^2 + (a_j\cdot u)^2 + (a_j\cdot e_1)^2}\Big)\phi_2\Big(\frac{d_{21}|u|^2 + d_{22}(a_j\cdot e_1)^2 + d_{23}(a_j\cdot u)^2}{\beta|u|^2 + (a_j\cdot u)^2 + (a_j\cdot e_1)^2}\Big)$$
$$\cdot\, \phi_3(\|u\|_2)(a_j\cdot w)(a_j\cdot v), \quad u \in \mathbb{R}^n,\ w, v \in \mathbb{S}^{n-1}.$$

*Then for any $0 < \epsilon \le 1$, if $m \gtrsim n$, then the following hold with high probability:*

$$|I(u,w,v) - \mathbb{E} I(u,w,v)| \le \epsilon, \qquad \forall\, w, v \in \mathbb{S}^{n-1}, \forall\, c_1 \le \|u\|_2 \le c_2.$$

*Proof of Lemma D.3.* We first note that, by using a polarization argument and scaling (cf. the beginning part of the proof of Lemma C.8), it suffices for us to prove the statement for $I(u,w,w)$ uniformly in $w \in \mathbb{S}^{n-1}$ and $u \in \mathbb{R}^n$ with $c_1 \le \|u\|_2 \le c_2$.

Now let $\phi \in C_c^\infty(\mathbb{R})$ such that $0 \le \phi(x) \le 1$ for all $x$, $\phi(x) = 1$ for $|x| \le 1$ and $\phi(x) = 0$ for $|x| \ge 2$. Let $\delta > 0$ be a sufficiently small constant. The smallness of $\delta$ will be specified momentarily. Then

$$|I_1(u,w)| = \Big|\frac{1}{m}\sum_{j=1}^m \phi_1\Big(\frac{d_{11}|u|^2 + d_{12}(a_j\cdot e_1)^2 + d_{13}(a_j\cdot u)^2}{\beta|u|^2 + (a_j\cdot u)^2 + (a_j\cdot e_1)^2}\Big)\phi_2\Big(\frac{d_{21}|u|^2 + d_{22}(a_j\cdot e_1)^2 + d_{23}(a_j\cdot u)^2}{\beta|u|^2 + (a_j\cdot u)^2 + (a_j\cdot e_1)^2}\Big)$$
$$\cdot\, \phi_3(\|u\|_2)(a_j\cdot w)^2\phi\Big(\frac{a_j\cdot u}{\delta\langle a_j\cdot w\rangle}\Big)\Big|$$
$$\lesssim \frac{1}{m}\sum_{j=1}^m (a_j\cdot w)^2\phi\Big(\frac{a_j\cdot u}{\delta\langle a_j\cdot w\rangle}\Big)$$
$$\lesssim \frac{1}{m}\sum_{j=1}^m (a_j\cdot w)^2\Big(1 - \phi(2\delta^{\frac18}(a_j\cdot w))\Big) + \frac{1}{m}\sum_{j=1}^m \delta^{-\frac14}\phi\Big(\frac{a_j\cdot u}{\delta\langle\delta^{-\frac18}\rangle}\Big).$$

The expectation of the above two terms are clearly small if we take $\delta > 0$ sufficiently small. Moreover they are clearly OK for union bounds and can be made small in high probability. Thus for sufficiently small $\delta$, if $m \gtrsim n$, then with high probability we have

$$|I_1(u,w) - \mathbb{E} I_1(u,w)| \le \frac{\epsilon}{3}, \qquad \forall\, w \in \mathbb{S}^{n-1},\ \forall\, c_1 \le \|u\|_2 \le c_2.$$

We now fix $\delta$ and deal with the main term

$$\begin{aligned} &I_2(u,w)\\ =&\frac{1}{m}\sum_{j=1}^m \phi_1\Big(\frac{d_{11}|u|^2 + d_{12}(a_j\cdot e_1)^2 + d_{13}(a_j\cdot u)^2}{\beta|u|^2 + (a_j\cdot u)^2 + (a_j\cdot e_1)^2}\Big)\phi_2\Big(\frac{d_{21}|u|^2 + d_{22}(a_j\cdot e_1)^2 + d_{23}(a_j\cdot u)^2}{\beta|u|^2 + (a_j\cdot u)^2 + (a_j\cdot e_1)^2}\Big)\\ &\qquad \cdot\, \phi_3(\|u\|_2)(a_j\cdot w)^2\cdot\Big(1 - \phi\Big(\frac{a_j\cdot u}{\delta\langle a_j\cdot w\rangle}\Big)\Big)\\ =&\frac{1}{m}\sum_{j=1}^m H(\|u\|_2, a_j\cdot u, a_j\cdot w, a_j\cdot e_1),\end{aligned}$$

where

$$
\begin{aligned}
&H(s,z,y,b) \\
&= \phi_1(\frac{d_{11}s^2 + d_{12}b^2 + d_{13}z^2}{\beta s^2 + z^2 + b^2})\phi_2(\frac{d_{21}s^2 + d_{22}b^2 + d_{23}z^2}{\beta s^2 + z^2 + b^2}) \cdot \phi_3(s) y^2 \left(1 - \phi(\frac{z}{\delta \langle y \rangle})\right).
\end{aligned}
$$

The main point is to check the union bounds. Note that $s = \|u\|_2 \sim 1$. We have

$$
\begin{aligned}
&|\partial_s H(s,z,y,b)| \lesssim y^2; \\
&|\partial_z H(s,z,y,b)| \lesssim |y|; \\
&|\partial_y H(s,z,y,b)| \lesssim |y|.
\end{aligned}
$$

Thus for $c_1 \le \|u\|_2, \|\tilde{u}\|_2 \le c_2$, $w, \tilde{w} \in \mathbb{S}^{n-1}$, we have

$$
\begin{aligned}
&\Big| H(\|u\|_2, a_j \cdot u, a_j \cdot w, a_j \cdot e_1) - H(\|\tilde{u}\|_2, a_j \cdot \tilde{u}, a_j \cdot \tilde{w}, a_j \cdot e_1) \Big| \\
&\lesssim \|u - \tilde{u}\|_2 (|a_j \cdot w|^2) + |a_j \cdot (u - \tilde{u})||a_j \cdot w| + |a_j \cdot (w - \tilde{w})|(|a_j \cdot w| + |a_j \cdot \tilde{w}|).
\end{aligned}
$$

Clearly then the union bounds hold for $I_2$. Thus for $m \gtrsim n$, with high probability it holds that

$$
|I_2(u,w) - \mathbb{E} I_2(u,w)| \le \frac{\epsilon}{3}, \qquad \forall\, w \in \mathbb{S}^{n-1}, \forall\, c_1 \le \|u\|_2 \le c_2.
$$

The desired estimate for $I(u,w,w)$ then easily follows. □

Consider

$$
h(\rho, t, e^{\perp}) = \frac{1}{m} \sum_{j=1}^{m} \sqrt{\beta \rho^2 + \rho^2 (a_j \cdot \hat{u})^2 + X_j^2} \cdot \sqrt{\beta \rho^2 + 2X_j^2},
$$

where

$$
\begin{aligned}
&X_j = a_j \cdot e_1, \quad u = \rho \hat{u}, \quad 0 < c_1 \le \rho \le c_2 < \infty; \\
&\hat{u} = t e_1 + \sqrt{1 - t^2} e^{\perp}, \quad |t| < 1,\, e^{\perp} \cdot e_1 = 0, e^{\perp} \in \mathbb{S}^{n-1}.
\end{aligned}
$$

Here we take $c_1 > 0$, $c_2 > 0$ as two fixed constants. The main point is that $\rho \sim 1$. We consider $h$ in the regime

$$
|t| \le 1 - \epsilon_0,
$$

where $0 < \epsilon_0 \ll 1$ is fixed.

**Lemma D.4.** *Let $0 < \epsilon_0 \ll 1$ be fixed. For any $0 < \epsilon \le 1$, if $m \gtrsim n$, then with high probability it holds that*

$$
|\partial_t h - \mathbb{E} \partial_t h| + |\partial_{tt} h - \mathbb{E} \partial_{tt} h| \le \epsilon, \qquad \forall\, |t| \le 1 - \epsilon_0,\, e^{\perp} \cdot e_1 = 0, e^{\perp} \in \mathbb{S}^{n-1}, c_1 \le \rho \le c_2.
$$

*Proof of Lemma C.3.* Denote $Y_j = a_j \cdot e^{\perp}$ and

$$
Z_j = a_j \cdot \hat{u} = t X_j + \sqrt{1 - t^2} Y_j.
$$

Clearly

$$\frac{d}{dt}Z_j = X_j - \frac{t}{\sqrt{1-t^2}}Y_j;$$
$$\frac{d^2}{dt^2}Z_j = -(1-t^2)^{-\frac{3}{2}}Y_j.$$

Using $Y_j = (1-t^2)^{-\frac{1}{2}}(Z_j - tX_j)$, we obtain

$$\frac{d}{dt}Z_j = \frac{1}{1-t^2}X_j - \frac{t}{1-t^2}Z_j;$$
$$\frac{d^2}{dt^2}Z_j = (1-t^2)^{-2}(tX_j - Z_j).$$

Therefore

$$\begin{aligned}\partial_t h &= \frac{1}{1-t^2} \cdot \frac{1}{m}\sum_{j=1}^{m} \sqrt{\frac{\beta|u|^2 + 2X_j^2}{\beta|u|^2 + (a_j \cdot u)^2 + X_j^2}}\|u\|_2 \cdot (a_j \cdot u)X_j \\ &\quad - \frac{t}{1-t^2} \cdot \frac{1}{m}\sum_{j=1}^{m} \sqrt{\frac{\beta|u|^2 + 2X_j^2}{\beta|u|^2 + (a_j \cdot u)^2 + X_j^2}}(a_j \cdot u)^2 \\ &=: \frac{1}{1-t^2}H_1 - \frac{t}{1-t^2}H_2.\end{aligned}$$

By Lemma D.3, it holds with high probability that

$$|H_1 - \mathbb{E}H_1| + |H_2 - \mathbb{E}H_2| \le (1-\epsilon_0^2) \cdot \frac{\epsilon}{3}, \qquad \forall\, \hat{u} \in \mathbb{S}^{n-1}, c_1 \le \|u\|_2 \le c_2, |t| \le 1-\epsilon_0.$$

The desired estimate for $\partial_t h$ then easily follows.

To compute $\partial_{tt} h$, we shall denote

$$A_j = \beta\rho^2 + \rho^2 Z_j^2 + X_j^2 = \beta|u|^2 + (a_j \cdot u)^2 + X_j^2;$$
$$B_j = \beta\rho^2 + 2X_j^2 = \beta|u|^2 + 2X_j^2.$$

Then

$$\partial_{tt}h = -\frac{1}{m}\sum_{j=1}^{m} A_j^{-\frac{3}{2}} B_j^{\frac{1}{2}} (\rho^2 Z_j \frac{d}{dt} Z_j)^2 + \frac{1}{m}\sum_{j=1}^{m} A_j^{-\frac{1}{2}} \sqrt{B_j}\rho^2 \cdot \Big( (\frac{d}{dt}Z_j)^2 + Z_j \frac{d^2}{dt^2} Z_j \Big)$$

$$\begin{aligned}
&= -(1-t^2)^{-2}\frac{1}{m}\sum_{j=1}^{m} A_j^{-\frac{3}{2}} B_j^{\frac{1}{2}} \|u\|_2^2 (a_j\cdot u)^2 X_j^2 \\
&\quad + 2t(1-t^2)^{-2}\frac{1}{m}\sum_{j=1}^{m} A_j^{-\frac{3}{2}} B_j^{\frac{1}{2}} \|u\|_2 (a_j\cdot u)^3 X_j \\
&\quad - t^2(1-t^2)^{-2}\frac{1}{m}\sum_{j=1}^{m} A_j^{-\frac{3}{2}} B_j^{\frac{1}{2}} (a_j\cdot u)^4 \\
&\quad + (1-t^2)^{-2}\frac{1}{m}\sum_{j=1}^{m} A_j^{-\frac{1}{2}} B_j^{\frac{1}{2}} \|u\|_2^2 X_j^2 \\
&\quad - t(1-t^2)^{-2}\frac{1}{m}\sum_{j=1}^{m} A_j^{-\frac{1}{2}} B_j^{\frac{1}{2}} \|u\|_2 X_j (a_j\cdot u) \\
&\quad - (1-t^2)^{-1}\frac{1}{m}\sum_{j=1}^{m} A_j^{-\frac{1}{2}} B_j^{\frac{1}{2}} (a_j\cdot u)^2.
\end{aligned}$$

It is then a bit tedious but not difficult to verify that the above terms can be treated with the help of Lemma D.3. Thus with high probability it holds that

$$|\partial_{tt}h - \mathbb{E}\partial_{tt}h| \le \frac{\epsilon}{5}, \qquad \forall\, \hat{u} \in \mathbb{S}^{n-1},\ c_1 \le \|u\|_2 \le c_2, |t| \le 1-\epsilon_0.$$

□

**Lemma D.5.** *Let $X \sim \mathcal{N}(0,1)$, $Y \sim \mathcal{N}(0,1)$ be independent. Define*

$$H(\rho,s) = \mathbb{E}\sqrt{\beta\rho^2 + \rho^2(\sqrt{1-s^2}X + sY)^2 + X^2}\sqrt{\beta\rho^2 + 2X^2};$$
$$h(\rho,s) = \frac{1}{2}(1+2\beta)\rho^2 - H(\rho,s).$$

*Then it holds that*

$$\sup_{|\rho-1|\ll 1, |s|\ll 1} \sum_{j=1}^{3} (|\partial^j H| + |\partial^j h|) \lesssim 1.$$

*where $\partial = \partial_\rho$ or $\partial_s$.*

*Proof.* For $H(\rho,s)$, this is obvious since the integrand inside the expectation is smooth. The estimate for $h(\rho,s)$ also follows easily. □

**Lemma D.6** (Calculation of $\partial^2 h$ at $(\rho = 1,\ s = 0)$)**.** *Let*

$$H(\rho,s) = \mathbb{E}\sqrt{\beta\rho^2 + \rho^2(\sqrt{1-s^2}X + sY)^2 + X^2}\sqrt{\beta\rho^2 + 2X^2};$$
$$h(\rho,s) = \frac{1}{2}(1+2\beta)\rho^2 - H(\rho,s).$$

*Then at $\rho = 1$, $s = 0$, we have*

$$
\begin{aligned}
&(\partial_{\rho\rho} H)(1,0) = \gamma_1 > 0, \quad (\partial_{\rho s} H)(\rho, 0) = 0, \, \forall\, \rho > 0; \\
&(\partial_{ss} H)(1,0) = -\gamma_2 < 0; \\
&(\partial_s h)(\rho, 0) = 0, \forall\, \rho > 0, \quad (\partial_\rho h)(1,0) = 0; \\
&(\partial_{\rho\rho} h)(1,0) = \gamma_3 > 0, \quad (\partial_{\rho s} h)(\rho, 0) = 0, \, \forall\, \rho > 0; \\
&(\partial_{ss} h)(1,0) = \gamma_4 > 0,
\end{aligned}
$$

*where $\gamma_i > 0$, $i = 1, \cdots, 4$ are constants depending on $\beta$.*

*Proof of Lemma D.6.* Firstly by using parity it is easy to check that $(\partial_s H)(\rho, 0) = 0$ for any $\rho > 0$. It follows easily that $(\partial_{\rho s} h)(\rho, 0) = (\partial_{\rho s} H)(\rho, 0) = 0$ for any $\rho > 0$. It is also easy to check that

$$
\begin{aligned}
(\partial_\rho H)(1,0) &= \partial_\rho \mathbb{E}(\sqrt{\beta\rho^2 + (\rho^2+1)X^2}\sqrt{\beta\rho^2 + 2X^2})\Big|_{\rho=1} \\
&= \mathbb{E}(2\beta + X^2) = 2\beta + 1.
\end{aligned}
$$

Clearly $(\partial_\rho h)(1,0) = 0$. One should note that we can also deduce this directly (and easily) from the fact that the original loss function attains a minimum at $u = e_1$.

<u>`Calculation of` $\partial_{ss} H$</u>. By a tedious computation, we have

$$
\begin{aligned}
2\pi(\partial_{ss} H)(1,0) &= 2\pi \partial_{ss}\Big(\mathbb{E}\sqrt{\beta + X^2 + (\sqrt{1-s^2}X + sY)^2}\sqrt{\beta + 2X^2}\Big)\Big|_{s=0} \\
&= \int_{\mathbb{R}^2} \frac{-2x^4 - \beta x^2 + (\beta + x^2)y^2}{\beta + 2x^2} e^{-\frac{x^2+y^2}{2}} dxdy \\
&= \int_{\mathbb{R}} \frac{-2x^4 - \beta x^2 + \beta + x^2}{\beta + 2x^2} e^{-\frac{x^2}{2}} dx \\
&= \int_{\mathbb{R}} \Big(-x^2 \frac{\beta + 1 + 2x^2}{\beta + 2x^2} + 1\Big) e^{-\frac{x^2}{2}} dx \\
&= -\int_{\mathbb{R}} \frac{x^2}{\beta + 2x^2} e^{-\frac{x^2}{2}} dx < 0.
\end{aligned}
$$

<u>`Calculation of` $\partial_{\rho\rho} H$</u>. By a tedious computation, we have

$$
\begin{aligned}
(\partial_{\rho\rho} H)(1,0) &= \partial_{\rho\rho}\Big(\mathbb{E}\sqrt{\beta\rho^2 + (\rho^2+1)X^2}\sqrt{\beta\rho^2 + 2X^2}\Big)\Big|_{\rho=1} \\
&= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \frac{2\beta^2 + 5\beta x^2 + x^4}{\beta + 2x^2} e^{-\frac{x^2}{2}} dx \\
&= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} (2\beta + \frac{\beta + x^2}{\beta + 2x^2} x^2) e^{-\frac{x^2}{2}} dx.
\end{aligned}
$$

It follows that

$$
\begin{aligned}
(\partial_{\rho\rho} h)(1,0) &= 1 + 2\beta - (\partial_{\rho\rho} H)(1,0) \\
&= -\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \frac{x^4}{\beta + 2x^2} e^{-\frac{x^2}{2}} dx < 0.
\end{aligned}
$$

$\square$

**Lemma D.7.** *Let $0 < c_1 < c_2 < \infty$ be fixed. Consider*

$$f_0(u) = -\frac{1}{m} \sum_{k=1}^{m} \sqrt{\beta |u|^2 + (a_k \cdot u)^2 + (a_k \cdot e_1)^2} \sqrt{\beta |u|^2 + 2(a_k \cdot e_1)^2}.$$

*For any $0 < \epsilon \le 1$, if $m \gtrsim n$, then it holds with high probability that*

$$\Big| \sum_{i,j=1}^{n} \xi_i \xi_j (\partial_{ij} f_0)(u) - \sum_{i,j=1}^{n} \xi_i \xi_j \mathbb{E}(\partial_{ij} f_0)(u) \Big| \le \epsilon, \qquad \forall \xi \in \mathbb{S}^{n-1}, \quad \forall c_1 \le \|u\|_2 \le c_2.$$

*Proof of Lemma D.7.* To simplify the notation, write $a_k$ as $a$, and denote

$$\begin{aligned}
&A = \beta |u|^2 + (a \cdot u)^2 + (a \cdot e_1)^2, \qquad B = \beta |u|^2 + 2(a \cdot e_1)^2; \\
&\partial_i A = 2\beta u_i + 2(a \cdot u) a_i, \qquad \partial_{ij} A = 2\beta \delta_{ij} + 2 a_i a_j; \\
&\partial_i B = 2\beta u_i, \qquad \partial_{ij} B = 2\beta \delta_{ij}.
\end{aligned}$$

We need to compute $\partial_{ij} \tilde{F}$ for

$$\tilde{F} = A^{\frac{1}{2}} B^{\frac{1}{2}}.$$

Clearly

$$\begin{aligned}
\partial_i \tilde{F} &= \frac{1}{2} A^{-\frac{1}{2}} \partial_i A B^{\frac{1}{2}} + \frac{1}{2} B^{-\frac{1}{2}} \partial_i B A^{\frac{1}{2}}; \\
\partial_{ij} \tilde{F} &= -\frac{1}{4} A^{-\frac{3}{2}} \partial_i A \partial_j A B^{\frac{1}{2}} \\
&\quad + \frac{1}{2} A^{-\frac{1}{2}} \partial_{ij} A B^{\frac{1}{2}} \\
&\quad + \frac{1}{2} A^{-\frac{1}{2}} \partial_i A \frac{1}{2} B^{-\frac{1}{2}} \partial_j B \\
&\quad - \frac{1}{4} B^{-\frac{3}{2}} \partial_j B \partial_i B A^{\frac{1}{2}} \\
&\quad + \frac{1}{2} B^{-\frac{1}{2}} \partial_{ij} B A^{\frac{1}{2}} \\
&\quad + \frac{1}{4} B^{-\frac{1}{2}} A^{-\frac{1}{2}} \partial_i B \partial_j A.
\end{aligned}$$

We then have

$$\begin{aligned}
&\sum_{i,j=1}^{n} \xi_i \xi_j (\partial_{ij} f_0)(u) \\
=& \frac{1}{4} \cdot \frac{1}{m} \sum_{k=1}^{m} A_k^{-\frac{3}{2}} |\xi \cdot \nabla A_k|^2 B_k^{\frac{1}{2}} \qquad \text{(D.3)}\\
&+ \frac{1}{4} \cdot \frac{1}{m} \sum_{k=1}^{m} B_k^{-\frac{3}{2}} |\xi \cdot \nabla B_k|^2 A_k^{\frac{1}{2}} \qquad \text{(D.4)}\\
&- \frac{1}{2} \cdot \frac{1}{m} \sum_{k=1}^{m} A_k^{-\frac{1}{2}} \langle \xi, (\nabla^2 A_k) \xi \rangle B_k^{\frac{1}{2}} \qquad \text{(D.5)}\\
&- \frac{1}{2} \cdot \frac{1}{m} \sum_{k=1}^{m} A_k^{-\frac{1}{2}} B_k^{-\frac{1}{2}} (\nabla A_k \cdot \xi)(\nabla B_k \cdot \xi) \qquad \text{(D.6)}\\
&- \frac{1}{2} \cdot \frac{1}{m} \sum_{k=1}^{m} B_k^{-\frac{1}{2}} A_k^{\frac{1}{2}} \langle \xi, (\nabla^2 B_k) \xi \rangle, \qquad \text{(D.7)}
\end{aligned}$$

where $A_k = \beta|u|^2 + (a_k \cdot u)^2 + (a_k \cdot e_1)^2$, $B_k = \beta |u|^2 + 2(a_k \cdot e_1)^2$, and we have denoted

$$\langle \xi, (\nabla^2 A_k) \xi \rangle = \sum_{i,j=1}^{n} \xi_i \xi_j \partial_{ij} A_k.$$

Thanks to the strong damping provided by $A_k$, it is tedious but not difficult to check that the terms (D.3), (D.5), (D.6) can be easily controlled with the help of Lemma D.3. The term (D.7) can be estimated in a similar way as in the estimate of (C.9) in the proof of Lemma C.9 (note that this is done in high probability therein!). The term (D.4) is also easy to handle. We omit further details. □

## Appendix E. Technical estimates for Section 5

**Lemma E.1.** *Let $\phi \in C_c^\infty(\mathbb{R})$ satisfies $0 \le \phi(x) \le 1$ for all $x$, $\phi(x) = 1$ for $|x| \le 1$ and $\phi(x) = 0$ for $|x| \ge 2$. There exist $\epsilon > 0$ sufficiently small, and $N$ sufficiently large such that*

$$\mathbb{E} \frac{(a \cdot \xi)^2 (a \cdot e_1)^2}{\epsilon + (a \cdot e_1)^2} \phi(\frac{a \cdot \xi}{N}) \ge 0.99, \qquad \forall \xi \in \mathbb{S}^{n-1},$$

*where $a \sim \mathcal{N}(0, \mathrm{I_n})$.*

*Proof of Lemma E.1.* We first show that there exist $\epsilon > 0$, such that

$$\mathbb{E} \frac{(a \cdot \xi)^2 (a \cdot e_1)^2}{\epsilon + (a \cdot e_1)^2} \ge 0.995, \qquad \forall \xi \in \mathbb{S}^{n-1}. \tag{E.1}$$

Clearly it suffices for us to show

$$\sup_{\xi \in \mathbb{S}^{n-1}} \mathbb{E} (a \cdot \xi)^2 \frac{\epsilon}{\epsilon + (a \cdot e_1)^2} \to 0, \quad \text{as } \epsilon \to 0. \tag{E.2}$$

Observe that $\xi = se_1 + \sqrt{1-s^2}e_1^{\perp}$, $|s| \le 1$, $e^{\perp} \cdot e_1 = 0$. Thus denoting $X$ and $Y$ as two independent standard Gaussian random variables with mean zero and unit variance, we have

$$\begin{aligned}&\sup_{\xi \in \mathbb{S}^{n-1}} \mathbb{E}(a\cdot\xi)^2 \frac{\epsilon}{\epsilon + (a\cdot e_1)^2} \\ &\lesssim \mathbb{E}X^2 \frac{\epsilon}{\epsilon+X^2} + \mathbb{E}Y^2 \frac{\epsilon}{\epsilon+X^2} \\ &\lesssim \epsilon + \mathbb{E}\frac{\epsilon}{\epsilon+X^2} \lesssim \sqrt{\epsilon},\end{aligned}$$

where in the last inequality we used the fact that

$$\int_{|x|\le 1} \frac{\epsilon}{\epsilon+x^2} dx \sim \sqrt{\epsilon}.$$

Thus (E.2) and (E.1) hold. Now $\epsilon$ is fixed. To show the final inequality, we note that

$$\begin{aligned}&\mathbb{E}\frac{(a\cdot\xi)^2(a\cdot e_1)^2}{\epsilon+(a\cdot e_1)^2}\chi_{|a\cdot\xi|\ge N} \\ &\le \mathbb{E}(a\cdot\xi)^2\chi_{|a\cdot\xi|\ge N} \\ &\le \mathbb{E}X^2\chi_{|X|\ge N} \to 0,\end{aligned}$$

as $N$ tend to infinity. Thus the desired inequality easily follows. $\square$

**Lemma E.2.** *Let $0 < \eta_0 \ll 1$ be given. Then if $m \gtrsim n$, then the following hold with high probability:*

$$\frac{1}{m}\sum_{k=1}^m \frac{(a_k\cdot\hat{u})^4}{(a_k\cdot e_1)^2} \ge 100, \qquad \forall\, \hat{u} \in \mathbb{S}^{n-1} \text{ with } ||\hat{u}\cdot e_1| - 1| \ge \eta_0.$$

*Proof of Lemma E.2.* Without loss of generality we write

$$\hat{u} = se^{\perp} \pm \sqrt{1-s^2}e_1, \qquad e^{\perp} \in \mathbb{S}^{n-1} \text{ with } e^{\perp}\cdot e_1 = 0.$$

Clearly $|s| \ge s_0 = s_0(\eta_0) > 0$ where $s_0(\eta_0)$ is a constant depending only on $\eta_0$. Take $a \sim \mathcal{N}(0, \mathrm{I_n})$ and observe that

$$\begin{aligned}&\mathbb{E}\frac{(a\cdot\hat{u})^4}{\epsilon(1+(a\cdot e^{\perp})^2)+(a\cdot e_1)^2} \\ &\ge \mathbb{E}\frac{s^4(a\cdot e^{\perp})^4}{\epsilon(1+(a\cdot e^{\perp})^2)+(a\cdot e_1)^2} \\ &\ge s_0^4\frac{1}{2\pi}\int_{1\le y\le 2,\, x\in\mathbb{R}} \frac{y^4}{\epsilon(1+y^2)+x^2}e^{-\frac{x^2+y^2}{2}}dxdy \\ &\ge s_0^4\frac{1}{200}\int_{|x|\le 1}\frac{1}{5\epsilon+x^2}dx \ge s_0^4\cdot O(\epsilon^{-\frac{1}{2}}) \ge 200,\end{aligned}$$

if $\epsilon > 0$ is taken sufficiently small. Now we fix this $\epsilon$. Clearly for $m \gtrsim n$ with high probability it holds that

$$\begin{aligned}\frac{1}{m}\sum_{k=1}^m \frac{(a_k\cdot\hat{u})^4}{(a_k\cdot e_1)^2} &\ge \frac{1}{m}\sum_{k=1}^m \frac{(a_k\cdot\hat{u})^4}{\epsilon(1+(a_k\cdot e^{\perp})^2)+(a_k\cdot e_1)^2} \\ &\ge 100, \qquad \forall\, \hat{u}\in\mathbb{S}^{n-1} \text{ with } ||\hat{u}\cdot e_1|-1| \le \eta_0.\end{aligned}$$

$\square$

**Lemma F.1.** *For any $\epsilon > 0$, there exists $R_0 = R_0(\beta, \epsilon) > 0$ sufficiently small, such that if $m \gtrsim n$, then the following hold with high probability:*

$$R\frac{1}{m}\sum_{k=1}^{m}\frac{(a_k \cdot \hat{u})^2}{\beta R + (a_k \cdot e_1)^2} < \epsilon, \qquad \forall\, \hat{u} \in \mathbb{S}^{n-1}, \quad \forall\, 0 < R \le R_0.$$

*Proof.* Let $\phi \in C_c^\infty(\mathbb{R})$ be such that $0 \le \phi(x) \le 1$ for all $x$, $\phi(x) = 1$ for $|x| \le 1$ and $\phi(x) = 0$ for $|x| \ge 2$. We then split the sum as

$$\begin{aligned}
&R\frac{1}{m}\sum_{k=1}^{m}\frac{(a_k \cdot \hat{u})^2}{\beta R + (a_k \cdot e_1)^2} \\
\le&\frac{1}{\beta}\frac{1}{m}\sum_{k=1}^{m}(a_k \cdot \hat{u})^2\phi(\frac{a_k \cdot e_1}{\eta_0}) + R \cdot \eta_0^{-2}\frac{1}{m}\sum_{k=1}^{m}(a_k \cdot \hat{u})^2.
\end{aligned}$$

Clearly the first term is amenable to union bounds, and we can make it sufficiently small with high probability by taking $\eta_0$ small (depending only on $\beta$ and $\epsilon$). The second term is trivial since we can take $R$ sufficiently small. Thus the desired the result follows. $\square$

**Lemma F.2.** *There exists $R_1 = R_1(\beta) > 0$ sufficiently small, such that if $m \gtrsim n$, then the following hold with high probability:*

$$c_1 \le \frac{1}{m}\sum_{k=1}^{m}\frac{(a_k \cdot \hat{u})(a_k \cdot e_1)^3}{\beta R + (a_k \cdot e_1)^2} \le c_2, \qquad \forall\, \hat{u} \in \mathbb{S}^{n-1} \text{ with } \widehat{u}_1 \cdot e_1 \ge \frac{1}{10}, \quad \forall\, 0 < R \le R_1.$$

*In the above $c_1$, $c_2 > 0$ are constants depending only on $\beta$.*

*Proof.* Denote $X_k = a_k \cdot e_1$. Write $\hat{u} = s e_1 + \sqrt{1 - s^2}e^\perp$, where $s \ge \frac{1}{10}$ and $e^\perp \in \mathbb{S}^{n-1}$ satisfies $e^\perp \cdot e_1 = 0$. We then write

$$\begin{aligned}
&\frac{1}{m}\sum_{k=1}^{m}\frac{(a_k \cdot \hat{u})X_k^3}{\beta R + X_k^2} \\
=&s\frac{1}{m}\sum_{k=1}^{m}\frac{X_k^4}{\beta R + X_k^2} + \sqrt{1-s^2}\frac{1}{m}\sum_{k=1}^{m}\frac{(a_k \cdot e^\perp)X_k^3}{\beta R + X_k^2} \\
=&s\frac{1}{m}\sum_{k=1}^{m}\frac{X_k^4}{\beta R + X_k^2} + \sqrt{1-s^2}\frac{1}{m}\sum_{k=1}^{m}(a_k \cdot e^\perp)X_k - \sqrt{1-s^2}\frac{1}{m}\sum_{k=1}^{m}(a_k \cdot e^\perp)\frac{\beta R X_k}{\beta R + X_k^2}.
\end{aligned}$$

For the first term we note that for $0 < R \le 1$,

$$\frac{1}{m}\sum_{k=1}^{m}\frac{X_k^4}{\beta + X_k^2} \le \frac{1}{m}\sum_{k=1}^{m}\frac{X_k^4}{\beta R + X_k^2} \le \frac{1}{m}\sum_{k=1}^{m}X_k^2.$$

Thus we clearly have for all $0 < R \le 1$, $\frac{1}{10} \le s \le 1$, with high probability it holds that

$$2c_1 \le s\frac{1}{m}\sum_{k=1}^{m}\frac{X_k^4}{\beta R + X_k^2} \le \frac{1}{2}c_2.$$

The second term is clearly OK for union bounds and with high probability it can be made sufficiently small. For the last term, observe that with high probability,

$$\frac{1}{m}\sum_{k=1}^{m}|a_k\cdot e^{\perp}|\frac{\beta R|X_k|}{\beta R+X_k^2}\lesssim\sqrt{\beta R}\frac{1}{m}\sum_{k=1}^{m}|a_k\cdot e^{\perp}|\ll 1,\quad \forall\, e^{\perp}\in\mathbb{S}^{n-1},$$

if $R\le R_1$ and $R_1$ is sufficiently small. The desired result then clearly follows. $\square$

*Proof of Lemma 6.5.* Without loss of generality we consider the situation $\hat{u}=e_1\cos\theta+e^{\perp}\sin\theta$ with $\epsilon_1\le\theta\le\frac{\pi}{2}-\epsilon_2$, where $0<\epsilon_1,\epsilon_2\ll 1$. The point is that $\theta$ stays away from the end-points $0$ and $\frac{\pi}{2}$.

Denote $X_k=a_k\cdot e_1$, $Y_k=a_k\cdot e^{\perp}$ and $Z_k=a_k\cdot\hat{u}$. Then

$$\begin{aligned}
&Z_k=\cos\theta X_k+\sin\theta Y_k\ \Rightarrow\ Y_k=\frac{1}{\sin\theta}Z_k-\frac{\cos\theta}{\sin\theta}X_k;\\
&\partial_\theta Z_k=-\sin\theta X_k+\cos\theta Y_k=\cot\theta Z_k-\frac{1}{\sin\theta}X_k.
\end{aligned}$$

We then obtain

$$\begin{aligned}
\partial_\theta f=\;&4R^2\cot\theta\underbrace{\frac{1}{m}\sum_{k=1}^{m}\frac{Z_k^4}{\beta R+X_k^2}}_{=:H_0}-4R^2\csc\theta\frac{1}{m}\sum_{k=1}^{m}\frac{Z_k^3X_k}{\beta R+X_k^2}\\
&-4R\cot\theta\frac{1}{m}\sum_{k=1}^{m}\frac{Z_k^2X_k^2}{\beta R+X_k^2}+4R\csc\theta\frac{1}{m}\sum_{k=1}^{m}\frac{Z_kX_k^3}{\beta R+X_k^2}.
\end{aligned}$$

Since $R\sim 1$, it is not difficult to check that the third and fourth terms above are amenable to union bounds[2], i.e. with high probability (for $m\gtrsim n$) we have

$$\Big|\frac{1}{m}\sum_{k=1}^{m}\frac{Z_k^2X_k^2}{\beta R+X_k^2}-\text{mean}\Big|+\Big|\frac{1}{m}\sum_{k=1}^{m}\frac{Z_kX_k^3}{\beta R+X_k^2}-\text{mean}\Big|\ll 1,\qquad\forall\, c_1\le R\le c_2,\forall\,\hat{u}\in\mathbb{S}^{n-1}.$$

Next we treat the second term. Let $\phi\in C_c^\infty(\mathbb{R})$ be such that $0\le\phi(x)\le 1$ for all $x$, $\phi(x)=1$ for $|x|\le 1$ and $\phi(x)=0$ for $|x|\ge 2$. We have

$$\frac{1}{m}\sum_{k=1}^{m}\frac{Z_k^3X_k}{\beta R+X_k^2}=\underbrace{\frac{1}{m}\sum_{k=1}^{m}\frac{Z_k^3X_k}{\beta R+X_k^2}\phi\Big(\frac{Z_k}{M\langle X_k\rangle}\Big)}_{=:H_1}+\underbrace{\frac{1}{m}\sum_{k=1}^{m}\frac{Z_k^3X_k}{\beta R+X_k^2}\Big(1-\phi\Big(\frac{Z_k}{M\langle X_k\rangle}\Big)\Big)}_{=:H_2},$$

where $\langle z\rangle=(1+|z|^2)^{\frac{1}{2}}$. It is not difficult to check that $H_1$ is OK for union bounds, and with high probability it holds that

$$\Big|H_1-\mathbb{E}H_1\Big|\ll 1,\qquad\forall\,\hat{u}\in\mathbb{S}^{n-1},\ \forall\, c_1\le R\le c_2.$$

[2]The union bound includes covering in $\hat{u}$ and $R$.

For $H_2$ we have ($\eta_0$ will be taken sufficiently small)

$$\begin{aligned} H_2 &\le \eta_0 \frac{1}{m}\sum_{k=1}^m \frac{Z_k^4}{\beta R + X_k^2} + \eta_0^{-3}\frac{1}{m}\sum_{k=1}^m \frac{X_k^4}{\beta R + X_k^2}\Big(1-\phi(\frac{Z_k}{M\langle X_k\rangle})\Big) \\ &\le \underbrace{\eta_0 \frac{1}{m}\sum_{k=1}^m \frac{Z_k^4}{\beta R + X_k^2}}_{=:H_{2,a}} + \underbrace{\eta_0^{-3}\frac{1}{m}\sum_{k=1}^m X_k^2\Big(1-\phi(\frac{Z_k}{M\langle X_k\rangle})\Big)}_{=:H_{2,b}}. \end{aligned}$$

We first take $\eta_0$ sufficiently small so that $H_{2,a}$ can be included in the estimate of $H_0$ without affecting too much the main order. On the other hand, once $\eta_0$ is fixed, we can take $M$ sufficiently large such that

$$|H_{2,b}| + |\mathbb{E}H_{2,b}| \ll 1, \qquad \forall\, \hat{u}\in\mathbb{S}^{n-1},\ \forall\, c_1\le R\le c_2.$$

Finally we treat $H_0$. Clearly

$$H_0 \ge \underbrace{\frac{1}{m}\sum_{k=1}^m \frac{Z_k^4}{\beta R + X_k^2}\phi(\frac{Z_k}{K})}_{=:H_{0,a}}.$$

By taking $K$ large, it can be easily checked that

$$\sup_{\hat{u}\in\mathbb{S}^{n-1}, c_1\le R\le c_2} |\mathbb{E}H_0 - \mathbb{E}H_{0,a}| \ll 1.$$

On the other hand, for fixed $K$, clearly $H_{0,a}$ is OK for union bounds. It holds with high probability that

$$|H_{0,a} - \mathbb{E}H_{0,a}| \ll 1.$$

Collecting all the estimates, we obtain

$$\partial_\theta f \ge \mathbb{E}\partial_\theta f + \text{Error},$$

where $|\,\text{Error}\,| \ll 1$. The desired lower bound for $\partial_\theta f$ then easily follows from Lemma F.3 below. $\square$

**Lemma F.3.** *Let $u=\sqrt{R}\hat{u}$ with $0<c_1\le R\le c_2<\infty$ and $\hat{u}\in\mathbb{S}^{n-1}$. Assume $\hat{u}=\cos\theta e_1+\sin\theta e^{\perp}$, where $\theta\in[0,\pi]$ and $e^{\perp}\in\mathbb{S}^{n-1}$ satisfies $e^{\perp}\cdot e_1=0$. We have*

$$\mathbb{E}f(u) = h(\beta, R, \cos^2\theta),$$

*where*

$$\max_{0\le s\le 1}\partial_s h(\beta,R,s) \le -\gamma_1 < 0,$$
$$\min_{0\le s\le 1}\partial_{ss} h(\beta,R,s) \ge \gamma_2 > 0.$$

*Here $\gamma_i=\gamma_i(\beta,c_1,c_2)$, $i=1,2$ depend only on $(\beta,\ c_1,\ c_2)$. It follows that*

$$\mathbb{E}\partial_\theta f = a_1(\beta,R,\cos^2\theta)\sin(2\theta);$$
$$\mathbb{E}\partial_{\theta\theta} f = 2a_1(\beta,R,\cos^2\theta)\cos(2\theta) + a_2(\beta,R,\theta)\sin^2(2\theta),$$

*where*

$$\gamma_3 < a_i(\beta, R, s) \leq \gamma_4, \forall\, s \in [0,1],\ i = 1, 2;$$

*and* $\gamma_3 > 0$*,* $\gamma_4 > 0$ *are constants depending only on* $(\beta,\, c_1,\, c_2)$*.*

*Proof of Lemma F.3.* We have

$$\begin{aligned}
\mathbb{E} f(u) &= \frac{1}{2\pi} \int_{\mathbb{R}^2} \frac{(R(x\cos\theta + y\sin\theta)^2 - x^2)^2}{\beta R + x^2} e^{-\frac{x^2+y^2}{2}} dxdy \\
&= \frac{1}{\pi} \int_0^\infty \frac{1}{\beta R + x^2} e^{-\frac{x^2}{2}} \cdot \sqrt{2\pi} h_1(R, x, \cos^2\theta) dx,
\end{aligned}$$

where

$$h_1(R, x, s) = 3R^2 - 2Rx^2 + x^4 + s(6R^2 - 2Rx^2)(-1 + x^2) + R^2 s^2 (3 - 6x^2 + x^4)$$

Integrating further in $x$ then gives

$$\mathbb{E} f(u) = \sqrt{2\pi} \cdot \frac{1}{\pi} \cdot R\Big(c_1 s^2 + 2c_2 s + c_3\Big), \qquad s = \cos^2\theta,$$

where the value of $c_3$ is unimportant for us, and

$$\begin{aligned}
c_1 &= R \int_0^\infty \frac{1}{\beta R + x^2} e^{-\frac{x^2}{2}} (3 - 6x^2 + x^4) dx; \\
c_2 &= \int_0^\infty \frac{1}{\beta R + x^2} e^{-\frac{x^2}{2}} (3R - x^2)(-1 + x^2) dx.
\end{aligned}$$

First we show that $c_2 < 0$. By a short computation, we have

$$c_2 = \frac{3+\beta}{2\beta} \cdot \Big(\beta R \sqrt{2\pi} - e^{\frac{\beta R}{2}} \pi \sqrt{\beta R}(1 + \beta R)\,\mathrm{Erfc}(\sqrt{\frac{\beta R}{2}})\Big),$$

where

$$\mathrm{Erfc}(y) = \frac{2}{\sqrt{\pi}} \int_y^\infty e^{-t^2} dt.$$

We then reduce the matter to showing

$$y < e^{y^2}(1 + 2y^2) \int_y^\infty e^{-t^2} dt, \quad \forall\, y > 0. \tag{F.1}$$

This follows easily from the usual bound on $\mathrm{Erfc}(y)$:

$$\frac{1}{y + \sqrt{y^2 + 2}} < \mathrm{Erfc}(y) \cdot e^{y^2} \cdot \frac{\sqrt{\pi}}{2} \leq \frac{1}{y + \sqrt{y^2 + \frac{4}{\pi}}}, \quad \forall\, y > 0. \tag{F.2}$$

Thus $c_2 < 0$.

Next we show that $c_1 > 0$. We have

$$2\beta c_1 = -\sqrt{2\pi}\beta R(5 + \beta R) + e^{\frac{\beta R}{2}} \pi \sqrt{\beta R}(3 + \beta R(6 + \beta R)) \cdot \mathrm{Erfc}(\sqrt{\frac{\beta R}{2}}).$$

It amounts to checking

$$e^{y^2} \int_y^\infty e^{-t^2} dt > \frac{y(5 + 2y^2)}{3 + 4y^2(3 + y^2)}, \quad \forall\, y > 0.$$

This follows from Lemma F.4 below.

Finally we show $c_1 + c_2 < 0$. We have

$$
\begin{aligned}
&2(c_1+c_2)\\
&=\sqrt{2\pi}R(-2+\beta-\beta R) - e^{\frac{\beta R}{2}}\pi\cdot(-\beta^{\frac{3}{2}}R^{\frac{5}{2}}+(\beta R)^{\frac{3}{2}}+\sqrt{\beta R}-3R\sqrt{\beta R})\,\mathrm{Erfc}(\sqrt{\frac{\beta R}{2}}).
\end{aligned}
$$

Denote $y = \sqrt{\frac{\beta R}{2}} > 0$. We then reduce matters to showing

$$
2y^2 - 2R(1+y^2) < e^{y^2}\cdot 2y\cdot(-2y^2R - 3R + 1 + 2y^2)\int_y^\infty e^{-t^2}dt.
$$

Since we have shown (F.1), we then only need to check

$$
1 + y^2 > e^{y^2}y(2y^2+3)\int_y^\infty e^{-t^2}dt.
$$

This in turn follows from Lemma F.4.

Finally we consider the polynomial

$$
\tilde{h}(s) = c_1 s^2 + 2c_2 s.
$$

Since $\tilde{h}'(s) = 2c_1 s + 2c_2$ and $\tilde{h}'(0) = 2c_2 < 0$, $\tilde{h}'(1) = 2c_1 + 2c_2 < 0$, we have $\tilde{h}'(s) < 0$ for all $s \in [0,1]$. Since $c_1 > 0$, we have $\tilde{h}''(s) > 0$. The desired result then easily follows. $\square$

**Lemma F.4** (Refined upper and lower bounds on the Complementary Error function)**.** *Let* $\mathrm{Erfc}(x) = \frac{2}{\sqrt{\pi}}\int_x^\infty e^{-t^2}dt$ *for* $x > 0$. *Then*

$$
\begin{aligned}
e^{x^2}\cdot\mathrm{Erfc}(x)\cdot\frac{\sqrt{\pi}}{2} &> \frac{x(5+2x^2)}{3+4x^2(3+x^2)}, \quad \forall\, x > 0;\\
e^{x^2}\cdot\mathrm{Erfc}(x)\cdot\frac{\sqrt{\pi}}{2} &< \frac{1+x^2}{x(3+2x^2)}, \quad \forall\, x > 0.
\end{aligned}
$$

*Remark* F.1. In the regime $y \geq 1$, one can check that the upper and lower bounds here are sharper than (F.2). One should also recall that the usual way to derive the lower bound in (F.2) through conditional expectation. Namely one can regard $e^{-y^2}/(\sqrt{\pi}\,\mathrm{Erfc}(y))$ as the conditional mean $\mu_1(y) = \mathbb{E}(X|X > y)$ where $X$ has the p.d.f. $\frac{1}{\sqrt{\pi}}e^{-x^2}$. Then evaluating the variance $\mathbb{E}((X-\mu_1)^2|X > y) > 0$ gives $y\mu_1 + \frac{1}{2} - \mu_1^2 > 0$. This yields the upper bound for $\mu_1$ which in turn is the desired lower bound in (F.2). An interesting question is to derive a sharper two-sided bounds via more careful conditioning. However we shall not dwell on this issue here.

*Proof of Lemma F.4.* We focus on the regime $x > 1$. By performing successive simple change of variables, we have

$$\begin{aligned} g(x) := e^{x^2} \int_x^\infty e^{-t^2} dt &= \int_0^\infty e^{-2xs} e^{-s^2} ds \\ &= \frac{1}{2x} \int_0^\infty e^{-s} e^{-(\frac{s}{2x})^2} ds \\ &\sim \sum_{k=0}^{\infty} (-1)^k x^{-(2k+1)} \cdot \frac{1}{2} \cdot \frac{(2k)!}{4^k k!} \\ &\sim \sum_{k=0}^{\infty} (-1)^k x^{-(2k+1)} \cdot \frac{1}{2} \cdot \left(\frac{1}{2}\right)_k, \end{aligned}$$

where in the last line we adopted Pochhammer's symbol $(a)_n = a(a+1)\cdots(a+n-1)$. Note that the above is an asymptotic series, and it is not difficult to check that

$$\left| g(x) - \sum_{k=0}^{m} (-1)^k x^{-(2k+1)} \cdot \frac{1}{2} \cdot \left(\frac{1}{2}\right)_k \right| \le x^{-2m-3} \cdot \frac{1}{2} \cdot \left(\frac{1}{2}\right)_{m+1}, \qquad \forall m \ge 1, \forall x > 0.$$

Moreover, if $m$ is an even integer, then

$$g(x) < \sum_{k=0}^{m} (-1)^k x^{-(2k+1)} \cdot \frac{1}{2} \cdot \left(\frac{1}{2}\right)_k, \qquad \forall x > 0;$$

and if $m$ is odd, then

$$g(x) > \sum_{k=0}^{m} (-1)^k x^{-(2k+1)} \cdot \frac{1}{2} \cdot \left(\frac{1}{2}\right)_k, \qquad \forall x > 0.$$

Now taking $m = 4$, we have

$$g(x) < \frac{1}{2}x^{-1} - \frac{1}{4}x^{-3} + \frac{3}{8}x^{-5} - \frac{15}{16}x^{-7} + \frac{105}{32}x^{-9}.$$

For $x \ge 3$, it is not difficult to verify that

$$\frac{1}{2}x^{-1} - \frac{1}{4}x^{-3} + \frac{3}{8}x^{-5} - \frac{15}{16}x^{-7} + \frac{105}{32}x^{-9} < \frac{1+x^2}{x(3+2x^2)}.$$

Hence the upper bound is OK for $x \ge 3$.

Next taking $m = 5$, we have

$$g(x) > \frac{1}{2}x^{-1} - \frac{1}{4}x^{-3} + \frac{3}{8}x^{-5} - \frac{15}{16}x^{-7} + \frac{105}{32}x^{-9} - \frac{945}{64}x^{-11}.$$

It is not difficult to verify that for $x \ge 4$, we have

$$\frac{1}{2}x^{-1} - \frac{1}{4}x^{-3} + \frac{3}{8}x^{-5} - \frac{15}{16}x^{-7} + \frac{105}{32}x^{-9} - \frac{945}{64}x^{-11} > \frac{x(5+2x^2)}{3+12x^2+4x^4}.$$

Hence the lower bound is OK for $x \ge 4$.

Finally for the regime $x \in [0, 4]$, we use rigorous numerics to verify the inequality. Since we are on a compact interval, this can be done by a rigorous computation with controllable numerical errors. See figure 6 for the plot of the corresponding functions. $\square$

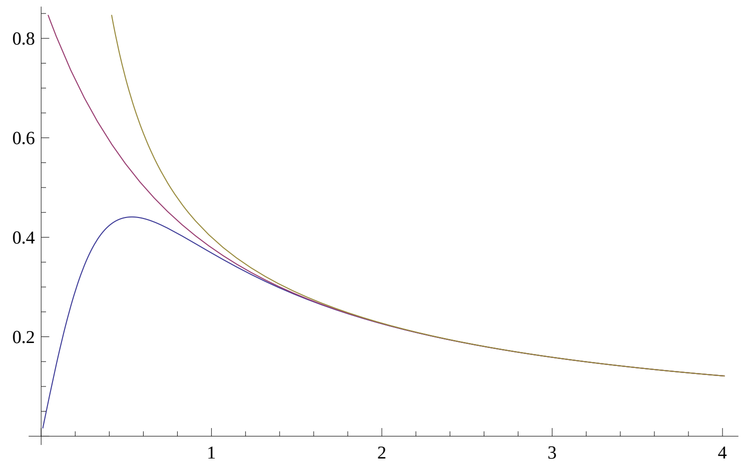



FIGURE 6. The functions $\frac{1+x^2}{x(3+2x^2)}$, $e^{x^2}\operatorname{Erfc}(x)\cdot\frac{\sqrt{\pi}}{2}$ and $\frac{x(5+2x^2)}{3+12x^2+4x^4}$

*Proof of Lemma 6.6.* Again denote $X_k = a_k \cdot e_1$ and $Z_k = a_k \cdot \hat{u}$. Without loss of generality we assume $\theta \in [\frac{\pi}{2}-\eta, \frac{\pi}{2}+\eta]$ for some sufficiently small $\eta > 0$. By a tedious computation, we have

$$\begin{aligned}\partial_{\theta\theta} f = {} & 4R^2(1+2\cos 2\theta)\csc^2\theta \frac{1}{m}\sum_{k=1}^m \frac{Z_k^4}{\beta R + X_k^2} - 24R^2(\cot\theta\csc\theta)\frac{1}{m}\sum_{k=1}^m \frac{X_k Z_k^3}{\beta R + X_k^2} \\ & + 4R(\csc^2\theta)(3R-\cos 2\theta)\frac{1}{m}\sum_{k=1}^m \frac{Z_k^2 X_k^2}{\beta R + X_k^2} + 8R(\cot\theta\csc\theta)\frac{1}{m}\sum_{k=1}^m \frac{X_k^3 Z_k}{\beta R + X_k^2} \\ & - 4R\csc^2\theta \frac{1}{m}\sum_{k=1}^m \frac{X_k^4}{\beta R + X_k^2}.\end{aligned}$$

Note that the third, fourth and fifth terms are OK for union bounds. The second and the first term can be handled in a similar way as in the proof of Lemma 6.5. The only difference is that the sign is now negative in the regime $\theta \to \frac{\pi}{2}$. Using Lemma F.3 it follows that $\partial_{\theta\theta} f < 0$ in this regime. We omit the repetitive details. $\square$

*Proof of Theorem 6.4.* Without loss of generality we consider the regime $\|u - e_1\|_2 \ll 1$. Before we work out the needed estimates for the restricted convexity, we explain the main difficulty in connection with the full Hessian matrix. Denote $X_k = a_k \cdot e_1$. Then for any $\xi \in \mathbb{S}^{n-1}$, we have

$$\begin{aligned} H_{\xi\xi} &= \sum_{i,j} \xi_i \xi_j (\partial_{u_i u_j} f)(u) \\ &= 12\frac{1}{m}\sum_{k=1}^m \frac{(a_k\cdot\xi)^2(a_k\cdot u)^2}{\beta|u|^2+X_k^2} && \text{(F.3)} \\ &\quad - 4\frac{1}{m}\sum_{k=1}^m \frac{(a_k\cdot\xi)^2X_k^2}{\beta|u|^2+X_k^2} && \text{(F.4)} \\ &\quad - 16\beta\frac{1}{m}\sum_{k=1}^m \frac{(a_k\cdot u)^3(a_k\cdot\xi)(u\cdot\xi)}{(\beta|u|^2+X_k^2)^2} && \text{(F.5)} \\ &\quad + 16\beta\frac{1}{m}\sum_{k=1}^m \frac{X_k^2(a_k\cdot u)(a_k\cdot\xi)(u\cdot\xi)}{(\beta|u|^2+X_k^2)^2} && \text{(F.6)} \\ &\quad - 2\beta\frac{1}{m}\sum_{k=1}^m \frac{((a_k\cdot u)^2-X_k^2)^2}{(\beta|u|^2+X_k^2)^2} && \text{(F.7)} \\ &\quad + 8\beta^2(\xi\cdot u)^2\frac{1}{m}\sum_{k=1}^m \frac{((a_k\cdot u)^2-X_k^2)^2}{(\beta|u|^2+X_k^2)^3}. && \text{(F.8)} \end{aligned}$$

First observe that if $u = e_1$, then the Hessian can be controlled rather easily thanks to the damping $\beta|u|^2 + X_k^2$.

On the other hand, for $u \neq e_1$, as far as the lower bound is concerned, the main difficult terms are (F.7) and (F.5) which are out of control if we do not impose any condition on $\xi$ (i.e. using (F.3) to control it). On the other hand, if we restrict $\xi$ to the direction $u - e_1$, then we can control these difficult terms by using the main good term (F.3). Namely, introduce the decomposition

$$u = e_1 + t\xi,$$

where $t = \|u - e_1\|_2 \ll 1$. Then for (F.5) we write

$$(a_k\cdot u)^3(a_k\cdot\xi) = (a_k\cdot u)^2(a_k\cdot e_1)(a_k\cdot\xi) + t(a_k\cdot u)^2(a_k\cdot\xi)^2$$

Since $t \ll 1$, the term $t(a_k\cdot u)^2(a_k\cdot e_1)^2$ (together with the pre-factor term in (F.5)) can be included into (F.3) which still has a good lower bound by using localization. On the other hand, the term $(a_k\cdot u)^2(a_k\cdot e_1)(a_k\cdot\xi)$ can be split as

$$\begin{aligned} &(a_k\cdot u)^2(a_k\cdot e_1)(a_k\cdot\xi) \\ =&(a_k\cdot u)^2(a_k\cdot e_1)(a_k\cdot\xi)\phi(\frac{a_k\cdot u}{K}) && \text{(F.9)} \\ &+ (a_k\cdot u)^2(a_k\cdot e_1)(a_k\cdot\xi)\Big(1-\phi(\frac{a_k\cdot u}{K})\Big), && \text{(F.10)} \end{aligned}$$

where $\phi$ is a smooth cut-off function satisfying $0 \le \phi(z) \le 1$ for all $z \in \mathbb{R}$, $\phi(z) = 1$ for $|z| \le 1$ and $\phi(z) = 0$ for $|z| \ge 2$. Clearly the contribution of (F.9) in (F.5) is OK for union bounds. On the other hand, for (F.10) we have

$$\begin{aligned}&(a_k\cdot u)^2|a_k\cdot e_1||a_k\cdot\xi|\cdot\Bigl(1-\phi(\frac{a_k\cdot u}{K})\Bigr)\\ &\le (a_k\cdot u)^2\epsilon(a_k\cdot\xi)^2+\epsilon^{-1}(a_k\cdot u)^2(a_k\cdot e_1)^2\Bigl(1-\phi(\frac{a_k\cdot u}{K})\Bigr).\end{aligned}$$

Clearly this is under control (the first term can again be controlled using (F.3)).

Now we turn to (F.7). The main term is $(a_k\cdot u)^4$. We write

$$(a_k\cdot u)^2(a_k\cdot u)^2=(a_k\cdot u)^2(a_k\cdot e_1)^2+t^2(a_k\cdot u)^2(a_k\cdot\xi)^2+2t(a_k\cdot u)^2(a_k\cdot e_1)(a_k\cdot\xi).$$

Clearly then this is also under control.

By further using localization, we can then show that with high probability, it holds that

$$H_{\xi\xi}\ge\mathbb{E}H_{\xi\xi}+\text{Error},$$

where $|\,\text{Error}\,|\ll 1$. The desired conclusion then follows from Lemma F.5. $\square$

*Remark* F.2. Introduce the parametrization $u=\sqrt{R}(e_1\cos\theta+e^{\perp}\sin\theta)$ where $e^{\perp}\in e_1=0$, $|R-1|\ll 1$ and $|\theta|\ll 1$. One might hope to prove that the Hessian matrix

$$\begin{pmatrix}\partial_{RR}f & \partial_{R\theta}f\\ \partial_{R\theta}f & \partial_{\theta\theta}f\end{pmatrix}$$

is positive definite near $u=e_1$ under the mere assume $m\gtrsim n$ and with high probability. However there is a subtle issue which we explain as follows. Consider the main term (write $X=a_k\cdot e_1$ and $Y=a_k\cdot e^{\perp}$)

$$\tilde f=\tilde f_k=\frac{\Bigl(R(X\cos\theta+Y\sin\theta)^2-X^2\Bigr)^2}{\beta R+X^2}.$$

The most troublesome piece come from quartic and cubic terms in $Y$, and we consider

$$\tilde h_1=\frac{R^2Y^4\sin^4\theta}{\beta R+X^2},\qquad \tilde h_2=\frac{R^2\Bigl(4Y^3X\sin^3\theta\cos\theta\Bigr)}{\beta R+X^2}.$$

For $\tilde h_2$ we do not have a favorable sign and the only hope is to control it via $\tilde h_1$. On the other hand, for $\tilde h_1$, we can take $X=Y=1$, $\beta=1$, and compute

$$(\partial_{RR}\tilde h_1)\cdot(\partial_{\theta\theta}\tilde h_1)-(\partial_{R\theta}\tilde h_1)^2=-\frac{8R^2}{(1+R)^4}\sin^6\theta\cdot\Bigl(3+4R+R^2+(2+4R+R^2)\cos2\theta\Bigr).$$

In yet other words, the sign is not favorable and this renders the Hessian out of control (before taking the expectation).

**Lemma F.5.** *Let $u=e_1+t\xi$ where $\xi\in\mathbb{S}^{n-1}$. Then for $|t|\ll 1$, we have*

$$\mathbb{E}\partial_{tt}f(u)\ge c_0>0,\qquad\forall\,\xi\in\mathbb{S}^{n-1},$$

*where $c_0>0$ is a constant depending only on $\beta$.*

*Proof of Lemma F.5.* Introduce the parametrization $\xi = se_1 + \sqrt{1-s^2}e^{\perp}$ where $e^{\perp} \cdot e_1 = 0$, $|s| \le 1$. Then $u = e_1 + t(se_1 + \sqrt{1-s^2}e^{\perp}) = (1+ts)e_1 + t\sqrt{1-s^2}e^{\perp}$. Thus

$$\mathbb{E}f(u) = \frac{1}{2\pi}\int_{\mathbb{R}^2} \underbrace{\frac{\Big(((1+ts)x + t\sqrt{1-s^2}y)^2 - x^2\Big)^2}{\beta(1+2ts+t^2)+x^2}}_{=:h(t,s,x,y)} e^{-\frac{x^2+y^2}{2}} dxdy.$$

It is not difficult to check that

$$\partial_{tt}h(t,s,x,y)\Big|_{t=0} = \frac{8x^2(sx+\sqrt{1-s^2}y)^2}{\beta+x^2}.$$

Thus it follows that

$$\mathbb{E}\partial_{tt}f(u)\Big|_{t=0,|s|\le 1} \gtrsim 1.$$

The desired result then follows by a simple perturbation argument using the fact that $\mathbb{E}\partial_{ttt}f$ is uniformly bounded and taking $|t|$ sufficiently small. $\square$

## Appendix G. Technical estimates for Section 7

**Lemma G.1.** *Let $u = \sqrt{R}\hat{u}$ with $0 < c_1 \le R \le c_2 < \infty$ and $\hat{u} \in \mathbb{S}^{n-1}$. Assume $\hat{u} = \cos\theta e_1 + \sin\theta e^{\perp}$, where $\theta \in [0,\pi]$ and $e^{\perp} \in \mathbb{S}^{n-1}$ satisfies $e^{\perp}\cdot e_1 = 0$. We have*

$$\mathbb{E}\partial_\theta f = a_1(\beta_1,\beta_2,R,\theta)\sin(2\theta);$$

*where*

$$\gamma_1 < a_1(\beta_1,\beta_2,R,\theta) \le \gamma_2, \quad \forall\,\theta\in[0,\pi],\, c_1\le R\le c_2;$$

*and $\gamma_1 > 0$, $\gamma_2 > 0$ are constants depending only on $(\beta_1, \beta_2, c_1, c_2)$. Furthermore for some sufficiently small constants $\theta_0 = \theta_0(\beta_1,\beta_2,c_1,c_2) > 0$, $\theta_1 = \theta_1(\beta_1,\beta_2,c_1,c_2) > 0$, we have*

$$\begin{aligned} &\gamma_3 < \mathbb{E}\partial_{\theta\theta}f < \gamma_4, && \text{if } 0\le\theta\le\theta_0 \text{ or } \pi-\theta_0\le\theta\le\pi,\\ &\gamma_5 < -\mathbb{E}\partial_{\theta\theta}f < \gamma_6, && \text{if } \Big|\theta-\frac{\pi}{2}\Big| < \theta_1, \end{aligned}$$

*where $\gamma_i > 0$, $i = 3,\cdots,6$ depend only on $(\beta_1, \beta_2, c_1, c_2)$.*

*Proof.* We have

$$\mathbb{E}f(u) = \frac{1}{2\pi}\int_{\mathbb{R}^2}\frac{\Big(R(x\cos\theta+y\sin\theta)^2-x^2\Big)^2}{R+\beta_1R(x\cos\theta+y\sin\theta)^2+\beta_2x^2}e^{-\frac{x^2+y^2}{2}}dxdy.$$

Denote

$$h(a,b) = \frac{(Ra^2-b)^2}{R+\beta_1Ra^2+\beta_2b}.$$

Then

$$
\begin{aligned}
&\partial_\theta\Big(h(x\cos\theta+y\sin\theta,x^2)\Big)=(-x\sin\theta+y\cos\theta)\partial_a h;\\
&\partial_x\Big(h(x\cos\theta+y\sin\theta,x^2)\Big)=\partial_a h\cdot\cos\theta+2x\partial_b h;\\
&\partial_y\Big(h(x\cos\theta+y\sin\theta,x^2)\Big)=\partial_a h\cdot\sin\theta;\\
&\partial_\theta\Big(h(x\cos\theta+y\sin\theta,x^2)\Big)=(y\partial_x-x\partial_y)\Big(h(x\cos\theta+y\sin\theta,x^2)\Big)-2xy\partial_b h.
\end{aligned}
$$

By using integration by parts, we then obtain

$$
\begin{aligned}
\mathbb{E}\partial_\theta f&=\frac{1}{\pi}\int_{\mathbb{R}^2}(-xy)(\partial_b h)(x\cos\theta+y\sin\theta,x^2)e^{-\frac{x^2+y^2}{2}}dxdy\\
&=\frac{2}{\pi}\int_{x>0,y>0}\Big((\partial_b h)(x\cos\theta-y\sin\theta,x^2)-(\partial_b h)(x\cos\theta+y\sin\theta,x^2)\Big)xye^{-\frac{x^2+y^2}{2}}dxdy.
\end{aligned}
$$

Now denote Denote

$$
h_1(a,b)=\frac{(Ra-b)^2}{R+\beta_1 Ra+\beta_2 b}.
$$

It is not difficult to check that for $a\ge 0$, $b\ge 0$, $\beta_1,\beta_2>0$, $R>0$,

$$
\partial_{ab}h_1=-2R^2\frac{(1+a(\beta_1+\beta_2))\cdot(b(\beta_1+\beta_2)+R)}{(\beta_2 b+R+\beta_1 aR)^3}<0.
$$

Observe that

$$
(\partial_b h)(a,b)=(\partial_b h_1)(a^2,b).
$$

Then if $x,y>0$ and $\theta\in[0,\pi]$, then

$$
\begin{aligned}
&(\partial_b h)(x\cos\theta-y\sin\theta,x^2)-(\partial_b h)(x\cos\theta+y\sin\theta,x^2)\\
=&(\partial_b h_1)((x\cos\theta-y\sin\theta)^2,x^2)-(\partial_b h_1)((x\cos\theta+y\sin\theta)^2,x^2)\\
=&-2\int_0^1(\partial_{ab}h_1)\Big((x\cos\theta+y\sin\theta)^2-4\tau xy\cos\theta\sin\theta,x^2\Big)d\tau\cdot xy\cdot\sin(2\theta).
\end{aligned}
$$

Integrating in $x$ and $y$, we then obtain

$$
\mathbb{E}\partial_\theta f=a_1(\beta_1,\beta_2,R,\theta)\sin(2\theta),
$$

where $a_1\sim 1$ and is a smooth function of $\theta$. Differentiating in $\theta$ then gives

$$
\mathbb{E}\partial_{\theta\theta}f=2a_1(\beta_1,\beta_2,R,\theta)\cos(2\theta)+\partial_\theta a_1(\beta_1,\beta_2,R,\theta)\sin(2\theta).
$$

Then second term clearly vanishes near $\theta=0,\frac{\pi}{2},\pi$. Thus the desired estimate for $\mathbb{E}\partial_{\theta\theta}f$ follows. $\square$

**Lemma G.2** (Strong convexity of $\mathbb{E}f$ when $\|u\pm e_1\|\ll 1$)**.** *Let $h(u)=\mathbb{E}f(u)$. There exists $0<\epsilon_0\ll 1$ such that the following hold:*

*(1) If $\|u-e_1\|_2\le\epsilon_0$, then for any $\xi\in\mathbb{S}^{n-1}$, we have*

$$
\sum_{i,j=1}^n\xi_i\xi_j(\partial_i\partial_j h)(u)\ge\gamma_1>0,
$$

*where $\gamma_1$ is a constant.*

*(2) If $\|u+e_1\|_2 \le \epsilon_0$, then for any $\xi \in \mathbb{S}^{n-1}$, we have*

$$\sum_{i,j=1}^{n} \xi_i \xi_j (\partial_i \partial_j h)(u) \ge \gamma_1 > 0.$$

*Proof.* We shall employ the same approach as in the proof of Theorem 3.5 and sketch only the needed modifications. Without loss of generality consider the regime $\|u-e_1\| \ll 1$ and introduce the change of variables:

$$\begin{aligned} &u = \rho \hat{u}; \\ &\hat{u} = \sqrt{1-s^2} e_1 + s e^{\perp}, \qquad e^{\perp} \cdot e_1 = 0,\, e^{\perp} \in \mathbb{S}^{n-1}, \end{aligned}$$

where $|\rho - 1| \ll 1$ and $0 \le s \ll 1$. Denote

$$h_1(\rho, s) = h(u) = h(\rho(\sqrt{1-s^2} e_1 + s e^{\perp})),$$

where we note that the value of $h(u)$ depends only on $(\rho,\, s)$. Clearly

$$h_1(\rho, s) = \frac{1}{2\pi} \int_{\mathbb{R}^2} \underbrace{\frac{\left(\rho^2 (\sqrt{1-s^2} x + s y)^2 - x^2\right)^2}{\rho^2 + \beta_1 \rho^2 (\sqrt{1-s^2} x + s y)^2 + \beta_2 x^2}}_{=: h_2(\rho, s, x, y)} e^{-\frac{x^2+y^2}{2}} dx dy.$$

It is easy to check that

$$\max_{\frac{1}{2} \le \rho \le 2, |s| \le \frac{1}{2}} \sum_{i,j \le 4} |\partial_\rho^i \partial_s^j h_1(\rho, s)| \lesssim 1.$$

By a tedious computation, we have

$$\begin{aligned} &\partial_{\rho\rho} h_2(\rho, 0, x, y) \\ &= \frac{2(3\rho^2 + \rho^6) x^4 + k_1 \cdot x^6 + k_2 x^8}{(\beta_2 x^2 + \rho^2 (1 + \beta_1 x^2))^3}, \end{aligned}$$

where

$$\begin{aligned} &k_1 = 2(-\beta_2 + 6\beta_1 \rho^2 + 6\beta_2 \rho^2 + 3\beta_2 \rho^4 + 2\beta_1 \rho^6); \\ &k_2 = 2(-\beta_1 \beta_2 - 2\beta_2^2 + 3\beta_1^2 \rho^2 + 6\beta_1 \beta_2 \rho^2 + 6\beta_2^2 \rho^2 + 3\beta_1 \beta_2 \rho^4 + \beta_1^2 \rho^6). \end{aligned}$$

Since $\rho \to 1$, it is clear that $k_1 > 0$ and $k_2 > 0$, and thus

$$\partial_{\rho\rho} h_1(1, 0) \gtrsim 1.$$

It is not difficult to check that $\partial_s h_1(\rho, 0) = 0$ for any $\rho > 0$. Clearly also $\partial_{\rho s} h_1(\rho, 0) = 0$ for any $\rho > 0$. To compute $\partial_{ss} h_1(1, 0)$ we shall use Lemma G.1. Observe that ($s = \sin\theta$ with $\theta \to 0+$)

$$\begin{aligned} h_1(\rho, \sin\theta) &= \mathbb{E} f(u); \\ \cos\theta \partial_s h_1(\rho, \sin\theta) &= \mathbb{E} \partial_\theta f; \\ -\sin\theta \partial_s h_1(\rho, \sin\theta) + \cos^2\theta \partial_{ss} h_1(\rho, \sin\theta) &= \mathbb{E} \partial_{\theta\theta} f. \end{aligned}$$

Clearly it follows that

$$\partial_{ss} h_1(1, 0) \gtrsim 1.$$

The rest of the argument is then essentially the same as in the proof of Theorem 3.5. We omit further details. $\square$

*Proof of Theorem 7.3.* We rewrite

$$f(u) = \frac{1}{m}\sum_{k=1}^m G(|u|^2, (a_k\cdot u)^2, X_k^2),$$

where

$$G(a,b,c) = \frac{(b-c)^2}{a+\beta_1 b+\beta_2 c}.$$

Clearly for any $\xi\in\mathbb{S}^{n-1}$,

$$\begin{aligned}
&\sum_{i,j=1}^n \xi_i\xi_j\partial_{u_iu_j} f \\
&= \frac{1}{m}\sum_{k=1}^m \partial_a G\cdot 2|\xi|^2 && \text{(G.1)}\\
&\quad + \frac{1}{m}\sum_{k=1}^m \partial_{aa} G\cdot 4(u\cdot\xi)^2 && \text{(G.2)}\\
&\quad + \frac{1}{m}\sum_{k=1}^m \partial_{ab} G\cdot 8(a_k\cdot u)(a_k\cdot\xi)(\xi\cdot u) && \text{(G.3)}\\
&\quad + \frac{1}{m}\sum_{k=1}^m \partial_{bb} G\cdot 4(a_k\cdot u)^2(a_k\cdot\xi)^2 && \text{(G.4)}\\
&\quad + \frac{1}{m}\sum_{k=1}^m \partial_{b} G\cdot 2(a_k\cdot\xi)^2. && \text{(G.5)}
\end{aligned}$$

In the above, $\partial_a G = (\partial_a G)(|u|^2, (a_k\cdot u)^2, X_k^2)$ and similar notation is used for $\partial_{aa}G$, $\partial_{bb}G$, $\partial_b G$.

`Estimate of` (G.1) `and` (G.2). Clearly these two terms are OK for union bounds, and we have (for $m\gtrsim n$ and with high probability)

$$|(\text{G.1}) - \text{mean}\,| + |(\text{G.2}) - \text{mean}\,| \ll 1, \qquad \forall\,\xi\in\mathbb{S}^{n-1}, \qquad \forall\,\frac{1}{2}\le\|u\|_2\le 2.$$

`Estimate of` (G.3). We have

$$(\partial_{ab}G)(a,b,c) = -\frac{2(b-c)(a+(\beta_1+\beta_2)c)}{(a+\beta_1 b+\beta_2 c)^3}.$$

Consider the function

$$\tilde{G}_1(a,y,c) = -y\frac{2(y^2-c)(a+(\beta_1+\beta_2)c)}{(a+\beta_1 y^2+\beta_2 c)^3}.$$

Clearly for $\frac{1}{10}\le a,\tilde{a}\le 10$, $y,\tilde{y}\in\mathbb{R}$, $c\ge 0$, we have $|\tilde{G}_1|\lesssim 1$ and

$$|\tilde{G}_1(a,y,c) - \tilde{G}_1(\tilde{a},\tilde{y},c)| \lesssim |a-\tilde{a}| + |y-\tilde{y}|.$$

Then for any $(u,\, \tilde{u})$ with $\frac{1}{2} \le \|u\|_2, \|\tilde{u}\|_2 \le 2$ and $(\xi,\, \tilde{\xi})$ with $\xi, \tilde{\xi} \in \mathbb{S}^{n-1}$, we have

$$\begin{aligned}&\Big|(\partial_{ab}G)(|u|^2, (a_k\cdot u)^2, X_k^2)(a_k\cdot u)(a_k\cdot\xi)\\&\quad - (\partial_{ab}G)(|\tilde{u}|^2, (a_k\cdot \tilde{u})^2, X_k^2)(a_k\cdot \tilde{u})(a_k\cdot\tilde{\xi})\Big|\\&\lesssim |a_k\cdot(\xi-\tilde{\xi})| + |a_k\cdot\xi|\cdot(|a_k\cdot(u-\tilde{u})| + \|u-\tilde{u}\|_2).\end{aligned}$$

Thus the union bound is also OK for this term, and we have

$$|(\text{G.3}) - \text{mean}\,| \ll 1, \qquad \forall\, \xi\in\mathbb{S}^{n-1}, \qquad \forall\, \frac{1}{2}\le \|u\|_2 \le 2.$$

`Estimate of` (G.4) `and` (G.5). We begin by noting that (G.4) and (G.5) can be combined into one term. Namely, observe that

$$\begin{aligned}&(\partial_{bb}G)(a,b,c)\cdot 2b + (\partial_b G)(a,b,c)\\=&\frac{H_1}{(a+\beta_1 b+\beta_2 c)^3},\end{aligned}$$

where

$$\begin{aligned}H_1 =& \beta_1^2 b^3 + a^2(6b-2c) + 3\beta_1\beta_2 b^2 c + 3b(\beta_1^2+2\beta_1\beta_2+2\beta_2^2)c^2 - \beta_2(\beta_1+2\beta_2)c^3\\&+ a\Big(3\beta_1 b^2 + 6(\beta_1+2\beta_2)bc - (\beta_1+4\beta_2)c^2\Big).\end{aligned}$$

We can then write

$$\begin{aligned}&(\text{G.4}) + (\text{G.5})\\=&\frac{1}{m}\sum_{k=1}^m (a_k\cdot\xi)^2 h_3(u, a_k\cdot u, X_k),\end{aligned}$$

where $h_3$ is a bounded smooth function with bounded derivatives in all of its arguments. Now let $\phi\in C_c^\infty$ be such that $0\le\phi(x)\le 1$ for all $x$, $\phi(x)=1$ for $|x|\le 1$ and $\phi(x)=0$ for $|x|\ge 2$. We then split the sum as

$$\begin{aligned}&\frac{1}{m}\sum_{k=1}^m (a_k\cdot\xi)^2 h_3(u, a_k\cdot u, X_k),\\=&\frac{1}{m}\sum_{k=1}^m (a_k\cdot\xi)^2\phi(\frac{a_k\cdot\xi}{K}) h_3(u, a_k\cdot u, X_k)\\&\quad+\frac{1}{m}\sum_{k=1}^m (a_k\cdot\xi)^2(1-\phi(\frac{a_k\cdot\xi}{K}))\cdot h_3(u, a_k\cdot u, X_k),\end{aligned}$$

where $K$ will be taken sufficiently large. Clearly the first term will be OK for union bounds. On the other hand, the second term can be dominated by

$$\text{const}\cdot\frac{1}{m}\sum_{k=1}^m (a_k\cdot\xi)^2(1-\phi(\frac{a_k\cdot\xi}{K})),$$

which can be made small by taking $K$ large. Thus we have

$$|(\text{G.4}) + (\text{G.5}) - \text{mean}\,| \ll 1, \quad \forall\, \xi\in\mathbb{S}^{n-1}, \quad \forall\, \frac{1}{2}\le\|u\|_2\le 2.$$

Collecting the estimates, we have for $m \gtrsim n$ and with high probability,

$$\Big|\sum_{i,j=1}^{n} \xi_i \xi_j \partial_{u_i u_j} f(u) - \text{mean}\Big| \ll 1, \quad \forall\, \xi \in \mathbb{S}^{n-1}, \quad \forall\, \frac{1}{2} \le \|u\|_2 \le 2.$$

The desired result then follows from Lemma G.2. □

## References


[1] S. Bhojanapalli, N. Behnam, and N. Srebro, "Global optimality of local search for low rank matrix recovery," *Advances in Neural Information Processing Systems*, pp. 3873–3881, 2016.

[2] E. J. Candès and X. Li, "Solving quadratic equations via PhaseLift when there are about as many equations as unknowns," *Found. Comut. Math.*, vol. 14, no. 5, pp. 1017–1026, 2014.

[3] E. J. Candès, X. Li, and M. Soltanolkotabi, "Phase retrieval via Wirtinger flow: Theory and algorithms," *IEEE Trans. Inf. Theory*, vol. 61, no. 4, pp. 1985–2007, 2015.

[4] E. J. Candès, T. Strohmer, and V. Voroninski, "Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming," *Commun. Pure Appl. Math.*, vol. 66, no. 8, pp. 1241–1274, 2013.

[5] J. Cai, M. Huang, D. Li and Y. Wang, " Solving phase retrieval with random initial guess is nearly as good as by spectral initialization," Preprint.

[6] Y. Chen and E. J. Candès, " Solving random quadratic systems of equations is nearly as easy as solving linear systems," *Commun. Pure Appl. Math.*, vol. 70, no. 5, pp. 822–883, 2017.

[7] J.C. Dainty and J.R. Fienup, "Phase retrieval and image reconstruction for astronomy," *Image Recovery: Theory and Application*, vol. 231, pp. 275, 1987.

[8] S. S. Du, C. Jin, J. D. Lee, and M. I. Jordan, "Gradient descent can take exponential time to escape saddle points," *Advances in Neural Information Processing Systems*, pp. 1067–1077, 2017.

[9] J. R. Fienup, "Phase retrieval algorithms: a comparison," *Appl. Opt.*, vol. 21, no. 15, pp. 2758–2769, 1982.

[10] B. Gao, Y. Wang, and Z. Xu, Solving a perturbed amplitude-based model for phase retrieval, 2019 [Online]. Available: http://arxiv.org/abs/1904.10307

[11] B. Gao and Z. Xu, "Phaseless recovery using the Gauss–Newton method," *IEEE Trans. Signal Process.*, vol. 65, no. 22, pp. 5885–5896, 2017.

[12] R. Ge, F. Huang, C. Jin, and Y. Yuan, "Escaping from saddle pointsonline stochastic gradient for tensor decomposition," *Conference on Learning Theory*, pp. 797–842, 2015.

[13] R. Ge, J. Lee, C. Jin, and T. Ma, "Matrix completion has no spurious local minimum," *Advances in Neural Information Processing Systems*, pp. 2973–2981, 2016.

[14] R. W. Gerchberg, A practical algorithm for the determination of phase from image and diffraction plane pictures, *Optik*, vol. 35, pp. 237–246, 1972.

[15] R. W. Gerchberg and W. O. Saxton, "A practical algorithm for the determination of the phase from image and diffraction plane pictures," *Optik*, vol. 35, pp. 237-246, 1972.

[16] R. W. Harrison, "Phase problem in crystallography," *JOSA A*, vol. 10, no. 5, pp. 1046–1055, 1993.

[17] C. Jin, R. Ge, P. Netrapalli, S. M. Kakade, and M. I. Jordan, "How to escape saddle points efficiently," *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pp. 1724–1732, 2017.

[18] C. Jin, P. Netrapalli, and M. I. Jordan, Accelerated gradient descent escapes saddle points faster than gradient descent, 2017 [Online]. Available: http://arxiv.org/abs/1711.10456

[19] Z. Li, J. F. Cai, and K. Wei, "Towards the optimal construction of a loss function without spurious local minima for solving quadratic equations," *IEEE Trans. Inf. Theory*, vol. 66, no. 5, pp. 3242–3260, 2020.

[20] J. Miao, T. Ishikawa, Q. Shen, and T. Earnest, "Extending x-ray crystallography to allow the imaging of noncrystalline materials, cells, and single protein complexes," *Annu. Rev. Phys. Chem.*, vol. 59, pp. 387–410, 2008.

[21] R. P. Millane, "Phase retrieval in crystallography and optics," *J. Optical Soc. America A*, vol. 7, no. 3, pp. 394-411, 1990.

[22] P. Netrapalli, P. Jain, and S. Sanghavi, "Phase retrieval using alternating minimization," *IEEE Trans. Signal Process.*, vol. 63, no. 18, pp. 4814–4826, 2015.

[23] D. Park, A. Kyrillidis, and C. Caramanis, Non-square matrix sensing without spurious local minima via the Burer-Monteiro approach, 2016 [Online]. Available: http://arxiv.org/abs/1609.03240

[24] H. Sahinoglou and S. D. Cabrera, "On phase retrieval of finite-length sequences using the initial time sample," *IEEE Trans. Circuits and Syst.*, vol. 38, no. 8, pp. 954–958, 1991.

[25] 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 Process. Mag.*, vol. 32, no. 3, pp. 87–109, 2015.

[26] J. Sun, Q. Qu, and J, Wright, "A geometric analysis of phase retrieval," *Found. Comput. Math.*, vol. 18, no. 5, pp. 1131–1198, 2018.

[27] J. Sun, Q. Qu, and J, Wright, “Complete dictionary recovery over the sphere I: Overview and the geometric picture,” *IEEE Trans. Inf. Theory*, vol. 63, no. 2, pp. 853–884, 2016.

[28] R. Vershynin, *High-dimensional probability: An introduction with applications in data science.* U.K.:Cambridge Univ. Press, 2018.

[29] I. Waldspurger, A. d'Aspremont, and S. Mallat, “Phase recovery, maxcut and complex semidefinite programming,” *Math. Prog.*, vol. 149, no. 1-2, pp. 47–81, 2015.

[30] A. Walther, “The question of phase retrieval in optics,” *J. Mod. Opt.*, vol. 10, no. 1, pp. 41–49, 1963.

[31] G. Wang, G. B. Giannakis, and Y. C. Eldar, “Solving systems of random quadratic equations via truncated amplitude flow,” *IEEE Trans. Inf. Theory*, vol. 64, no. 2, pp. 773–794, 2018.

[32] H. Zhang, Y. Zhou, Y. Liang, and Y. Chi, “A nonconvex approach for phase retrieval: Reshaped wirtinger flow and incremental algorithms,” *The Journal of Machine Learning Research*, vol. 18, no. 1, pp. 5164–5198, 2017.