Solving Systems of Random Quadratic Equations via Truncated Amplitude Flow

Gang Wang, *Student Member, IEEE,* Georgios B. Giannakis, *Fellow, IEEE,* and Yonina C. Eldar, *Fellow, IEEE*

Abstract

This paper puts forth a new algorithm, termed *truncated amplitude flow* (TAF), to recover an unknown n-dimensional real-/complex-valued vector $\boldsymbol{x}$ from m quadratic equations of the form $y_i = |\langle \boldsymbol{a}_i, \boldsymbol{x}\rangle|^2$. This problem is known to be *NP-hard* in general. We prove that as soon as the number of equations m is on the order of the number of unknowns n, TAF recovers the solution exactly (up to a global unimodular constant) with high probability and complexity growing linearly with the time required to read the data. Our method adopts the *amplitude-based* cost function and proceeds in two stages: In stage one, we introduce an *orthogonality-promoting* initialization that is obtained with a few simple power iterations. Stage two refines the initial estimate by successive updates of scalable *truncated generalized gradient iterations*. The former is in sharp contrast to existing spectral initializations, while the latter handles the rather challenging nonconvex and nonsmooth amplitude-based cost function. In particular for real-valued vectors, our gradient truncation rule provably eliminates the erroneously estimated signs with high probability to markedly improve upon its untruncated version. Numerical tests demonstrate that our initialization method returns more accurate and robust estimates relative to its spectral counterparts. Furthermore, even under the same initialization, our amplitude-based refinement outperforms Wirtinger-based alternatives, corroborating the superior performance of TAF over state-of-the-art algorithms.

Index terms— Nonconvex optimization, phase retrieval, amplitude-based cost function, orthogonality-promoting initialization.

I. Introduction

Consider a system of m quadratic equations

$$y_i = |\langle \boldsymbol{a}_i, \boldsymbol{x}\rangle|^2, \quad 1 \le i \le m \tag{1}$$

where the data vector $\boldsymbol{y} := [y_1 \ \cdots \ y_m]^{\mathcal{T}}$ and feature vectors $\boldsymbol{a}_i \in \mathbb{R}^n$ or $\mathbb{C}^n$, collected in the $m \times n$ matrix $\boldsymbol{A} := [\boldsymbol{a}_1 \ \cdots \ \boldsymbol{a}_m]^{\mathcal{H}}$ are known, whereas the vector $\boldsymbol{x} \in \mathbb{R}^n$ or $\mathbb{C}^n$ is the wanted unknown. When $\{\boldsymbol{a}_i\}_{i=1}^m$ and/or $\boldsymbol{x}$ are complex, the amplitudes of their inner-products $\{\langle \boldsymbol{a}_i, \boldsymbol{x}\rangle\}$ are given but phase information is lacking; in the real case only the signs of $\{\langle \boldsymbol{a}_i, \boldsymbol{x}\rangle\}$ are unknown. Assuming that the system of quadratic equations in (1) admits a unique solution $\boldsymbol{x}$ (up to a global unimodular constant), our objective is to reconstruct $\boldsymbol{x}$ from m phaseless quadratic equations, or equivalently, to recover the missing signs/phases of $\{\langle \boldsymbol{a}_i, \boldsymbol{x}\rangle\}$ under real-/complex-valued settings. Indeed, it has been established that $m \ge 2n-1$ or $m \ge 4n-4$ generic measurements $\{(\boldsymbol{a}_i;\, y_i)\}_{i=1}^m$ as in (1) suffice for uniquely determining an n-dimensional real- or complex-valued vector $\boldsymbol{x}$ [1], [2], [3], respectively.

The problem in (1) constitutes an instance of nonconvex quadratic programming, that is generally known to be *NP-hard* [4]. Specifically for real-valued vectors $\{\boldsymbol{a}_i\}$ and $\boldsymbol{x}$, this can be understood as

Work in this paper was supported in part by NSF grants 1500713 and 1514056. G. Wang and G. B. Giannakis are with the Digital Technology Center and the ECE Dept., University of Minnesota, Minneapolis, MN 55455, USA. G. Wang is also with the School of Automation, Beijing Institute of Technology, Beijing 100081, P. R. China. Y. C. Eldar is with EE Dept., Technion – Israel Institute of Technology, Haifa 32000, Israel. Emails: {gangwang,georgios}@umn.edu; yonina@ee.technion.ac.il.

a combinatorial optimization since one seeks a series of signs $\{s_i = \pm 1\}_{i=1}^m$, such that the solution to the system of linear equations $\langle \boldsymbol{a}_i, \boldsymbol{x} \rangle = s_i \psi_i$, where $\psi_i := \sqrt{y_i}$, obeys the given quadratic system (1). Concatenating all amplitudes $\{\psi_i\}_{i=1}^m$ to form the vector $\boldsymbol{\psi} := [\psi_1 \ \cdots \ \psi_m]^{\mathcal{T}}$, apparently there are a total of 2^m different combinations of $\{s_i\}_{i=1}^m$, among which only two lead to $\boldsymbol{x}$ up to a global sign. The complex case becomes even more complicated, where instead of a set of signs $\{s_i\}_{i=1}^m$, one must determine for uniqueness a collection of unimodular complex scalars $\{\sigma_i \in \mathbb{C}\}_{i=1}^m$. Special cases with $\boldsymbol{a}_i > \boldsymbol{0}$ (entry-wise inequality), $x_i^2 = 1$, and $y_i = 0$, $1 \leq i \leq m$ correspond to the so-called *stone problem* [5, Section 3.4.1], [6]. In many fields of physical sciences and engineering, the problem of recovering the phase from intensity/magnitude-only measurements is commonly referred to as *phase retrieval* [7], [8], [9]. The plethora of applications include X-ray crystallography [10], optics [11], [12], as well as array and high-power coherent diffractive imaging [13], [14], to astronomy [15], and microscopy [16], where due to physical limitations, optical sensors/detectors such as charge-coupled device (CCD) cameras, photosensitive films, and human eyes can record only (squared) modulus of the Fresnel or Fraunhofer diffraction pattern, while losing the phase of the incident light reaching the object. It has been shown that reconstructing a discrete, finite-duration signal from its Fourier transform magnitudes is generally *NP-complete* [17]. Even checking quadratic feasibility (i.e., whether a solution to a given quadratic system exists or not) is itself an *NP-hard* problem [18, Theorem 2.6]. Thus, despite its simple form and practical relevance across various fields, tackling the quadratic system in (1) under real-/complex-valued settings is challenging and *NP-hard* in general.

A. Prior Art

Adopting the least-squares criterion, the task of recovering $\boldsymbol{x}$ can be recast as that of minimizing the following *intensity-based* empirical loss

$$\underset{\boldsymbol{z} \in \mathbb{C}^n}{\text{minimize}} \quad f(\boldsymbol{z}) := \frac{1}{2m} \sum_{i=1}^m \left(y_i - |\boldsymbol{a}_i^{\mathcal{H}} \boldsymbol{z}|^2 \right)^2 \tag{2}$$

or, the *amplitude-based* one

$$\underset{\boldsymbol{z} \in \mathbb{C}^n}{\text{minimize}} \quad h(\boldsymbol{z}) := \frac{1}{2m} \sum_{i=1}^m \left(\psi_i - |\boldsymbol{a}_i^{\mathcal{H}} \boldsymbol{z}| \right)^2 \tag{3}$$

where $(\cdot)^{\mathcal{H}}$ stands for Hermitian transpose. Unfortunately, the presence of quadratic terms in (2) or the modulus in (3) renders the corresponding objective function nonconvex. Minimizing nonconvex objectives, which may exhibit many stationary points, is in general *NP-hard* [19]. It is worth stressing that even checking whether a given point is a local minimum or establishing convergence to a local minimum turns out to be *NP-complete* [19].

In the classical discretized one-dimensional (1D) phase retrieval, the amplitude vector $\boldsymbol{\psi}$ corresponds to the n-point Fourier transform of the n-dimensional signal $\boldsymbol{x}$ [20]. It has been shown based on spectral factorization that there is no unique solution in the 1D phase retrieval, even if we disregard trivial ambiguities [21]. To overcome this ill-posedness, several approaches have been suggested. One possibility is to assume additional constraints on the unknown signal such as sparsity [22], [23], [14], [24]. Other approaches rely on introducing redundancy into measurements using for example, the short-time Fourier transform, or masks [25], [26]. Finally, recent works assume random measurements (e.g., Gaussian $\{\boldsymbol{a}_i\}$ designs) [27], [8], [24], [28], [29], [6], [30]. Henceforth, this paper focuses on random measurements $\{\psi_i\}$ obtained from independently and identically distributed (i.i.d.) Gaussian $\{\boldsymbol{a}_i\}$ designs.

Existing approaches to solving (2) (or related ones using the Poisson likelihood; see, e.g., [6]) or (3) fall under two categories: nonconvex and convex ones. Popular nonconvex solvers include the alternating projection such as Gerchberg-Saxton [31] and Fineup [7], [32], AltMinPhase [27], (Truncated) Wirtinger flow (WF/TWF) [29], [6], [33], and Karzmarz variants [34] as well as trust-region methods [35]. Inspired by WF, other relevant judiciously initialized counterparts have also been developed for faster semidefinite optimization [36], [37], blind deconvolution [38], [39], and matrix completion [40]. Convex counterparts on the other hand rely on the so-called *matrix-lifting* technique or *Shor*'s semidefinite relaxation [41], to obtain the solvers abbreviated as PhaseLift [28], PhaseCut [42], and CoRK [43]. Other approaches dealing with noisy phase retrieval with outliers are discussed in [22], [30], [44].

In terms of sample complexity, it is proved that[1] $\mathcal{O}(n)$ noise-free random measurements suffice for uniquely determining a general signal [45]. It is also self-evident that recovering a general n-dimensional $\boldsymbol{x}$ requires at least $\mathcal{O}(n)$ measurements. Convex approaches enable exact recovery from the optimal bound $\mathcal{O}(n)$ of noiseless Gaussian measurements [46], while they require solving a semidefinite program of a matrix variable with size $n \times n$, thus incurring worst-case computational complexity on the order of $\mathcal{O}(n^{4.5})$ [42], [47] that does not scale well with the dimensionality n. Upon exploiting the underlying problem structure, $\mathcal{O}(n^{4.5})$ can be reduced to $\mathcal{O}(n^3)$ [42]. Solving for vector variables, nonconvex approaches achieve significantly improved computational performance. Using formulation (3) and adopting a spectral initialization commonly employed in matrix completion [48], AltMinPhase establishes exact recovery with sample complexity $\mathcal{O}(n \log^3 n)$ under i.i.d. Gaussian $\{\boldsymbol{a}_i\}$ designs with resampling [27]. Concerning formulation (2), WF iteratively refines the spectral initial estimate by means of a gradient-like update, which can be approximately interpreted as a stochastic gradient descent variant [29]. More details on WF can be found in [33]. The follow-up TWF improves upon WF through a truncation procedure to separate gradient components of excessively extreme (large or small) sizes. Likewise, at the initialization stage, since the terms $\{(\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x})^2\boldsymbol{a}_i\boldsymbol{a}_i^{\mathcal{H}}\}$ involving fourth-order moments of Gaussian $\{\boldsymbol{a}_i\}$ responsible for the spectral initialization are heavy-tailed, data $\{y_i\}_{i=1}^m$ are pre-screened to yield improved initial estimates in the so-termed truncated spectral initialization method [6]. WF allows exact recovery from $\mathcal{O}(n \log n)$ measurements in $\mathcal{O}(mn^2 \log(1/\epsilon))$ time/flops to yield an ϵ-accurate solution for any given $\epsilon > 0$ [29], while TWF advances these to $\mathcal{O}(n)$ measurements and $\mathcal{O}(mn \log(1/\epsilon))$ time [6]. Interestingly, the truncation procedure in the gradient stage turns out to be useful in avoiding spurious stationary points in the context of nonconvex optimization. Albeit for large-scale linear regressions, similar ideas including censoring have been studied [49], [50]. It is also worth mentioning that when $m \geq Cn \log^3 n$ for some sufficiently large universal constant $C > 0$, the objective function in (3) is shown to admit benign geometric structure that allows certain iterative algorithms (e.g., trust-region methods) to efficiently find a global minimizer with random initializations [35]. Hence, the challenge of solving systems of random quadratic equations lies in the case where a near-optimal number of equations are involved, e.g., $m = 2n - 1$ in the real-valued setting.

Although achieving a linear (in the number of unknowns n) sample and computational complexity, the state-of-the-art TWF approach still requires at least $4n \sim 5n$ equations to yield a stable empirical success rate (e.g., $\geq 99\%$) under the noiseless real-valued Gaussian model [6, Section 3], which are more than twice the known information-limit of $m = 2n - 1$ [1]. Similar though less obvious results hold also in the complex-valued scenario. On the other hand, even though the truncated spectral initialization in [6] improves upon the "plain vallina" spectral initialization, its performance still suffers when the number of measurements is relatively small and its advantage (over the untruncated one) narrows as the number of measurements grows; see more details on this in Fig. 4 and Section II. Furthermore, it is worth stressing

[1]The notation $\phi(n) = \mathcal{O}(g(n))$ means that there is a constant $c > 0$ such that $|\phi(n)| \leq c|g(n)|$.

that extensive numerical and experimental validation confirms that the *amplitude-based* cost function performs significantly better than the *intensity-based* one [51]; that is, formulation (3) is superior over (2). Hence, besides enhancing initialization, markedly improved performance in the gradient stage could be expected by re-examining the amplitude-based cost function and also incorporating judiciously designed gradient regularization rules.

B. This Paper

Along the lines of suitably initialized nonconvex schemes [29], [6] and inspired by [51], the present paper develops a linear-time (in both dimensions m and n) algorithm to minimize the amplitude-based cost function, referred to as *truncated amplitude flow* (TAF). Our approach provably recovers an n-dimensional unknown real-/complex-valued signal $\boldsymbol{x}$ exactly from a near-optimal number of noiseless random measurements, while also featuring a near-perfect statistical performance in the noisy setting. TAF operates in two stages: In stage one, we introduce an orthogonality-promoting initialization that is computable using a few power iterations. Stage two refines the initial estimate by successive updates of truncated generalized gradient iterations. Specifically, our initialization is built upon the hidden orthogonality characteristics of high-dimensional random vectors [52], which is in stark contrast to spectral alternatives starting from the strong law of large numbers (SLLN) [14], [29], [6]. On the other hand, the challenge of phase retrieval lies in reconstructing the signs/phases of $\langle \boldsymbol{a}_i, \boldsymbol{x} \rangle$ in the real-/complex-valued settings. Our refinement stage leverages a simple yet effective regularization rule to eliminate the erroneously estimated phases in the generalized gradient components with high probability. Simulated tests corroborate that the proposed initialization returns more accurate and robust initial estimates than its spectral counterparts in the noiseless and noisy settings. In addition, our TAF (with gradient truncation) markedly improves upon its "plain-vallina" version termed amplitude flow (AF). Empirical results corroborate the advantage of TAF over its competing alternatives.

The remainder of this paper is organized as follows. The amplitude-based cost function, as well as the two algorithmic stages is described and analyzed in detail in Section II. Section III summarizes the TAF algorithm and establishes its theoretical performance. Extensive simulated tests comparing TAF with Wirtinger-based approaches are presented in Section IV. Finally, main proofs are given in Section V, while technical details are deferred to Appendix.

II. Algorithm: Truncated Amplitude Flow

In this section, the two stages of our TAF algorithm are detailed. First, the challenge of handling the nonconvex and nonsmooth amplitude-based cost function is analyzed, and addressed by a carefully designed gradient regularization rule. Limitations of (truncated) spectral initializations are then pointed out, followed by a simple motivating example to inspire our orthogonality-promoting initialization method. For concreteness, our analysis will focus on the real-valued Gaussian model with $\boldsymbol{x} \in \mathbb{R}^n$ and i.i.d. design vectors $\boldsymbol{a}_i \in \mathbb{R}^n \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_n)$, whereas numerical implementations for the complex-valued Gaussian model having $\boldsymbol{x} \in \mathbb{C}^n$ and i.i.d. $\boldsymbol{a}_i \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I}_n) := \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_n/2) + j\mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_n/2)$ will be discussed briefly.

To start, define the Euclidean distance of any estimate $\boldsymbol{z}$ to the solution set: $\mathrm{dist}(\boldsymbol{z}, \boldsymbol{x}) := \min \|\boldsymbol{z} \pm \boldsymbol{x}\|$ for real signals, and $\mathrm{dist}(\boldsymbol{z}, \boldsymbol{x}) := \mathrm{minimize}_{\phi \in [0, 2\pi)} \|\boldsymbol{z} - \boldsymbol{x}\mathrm{e}^{i\phi}\|$ for complex ones [29], where $\|\cdot\|$ denotes the Euclidean norm. Define also the indistinguishable global phase constant in the real-valued setting as

$$\phi(\boldsymbol{z}) := \begin{cases} 0, & \|\boldsymbol{z} - \boldsymbol{x}\| \le \|\boldsymbol{z} + \boldsymbol{x}\|, \\ \pi, & \text{otherwise.} \end{cases} \tag{4}$$

Henceforth, fixing $\boldsymbol{x}$ to be any solution of the given quadratic system (1), we always assume that $\phi\left(\boldsymbol{z}\right)=0$; otherwise, $\boldsymbol{z}$ is replaced by $\mathrm{e}^{-j\phi(\boldsymbol{z})}\boldsymbol{z}$, but for simplicity of presentation, the constant phase adaptation term $\mathrm{e}^{-j\phi(\boldsymbol{z})}$ will be dropped whenever it is clear from the context.

State-of-the-art solution algorithms for (2) or (3) including WF/TWF are two-staged. Therefore, to fully demonstrate the power of our TAF approach, numerical tests comparing all stages of (T)AF and (T)WF will be presented throughout our analysis. Toward this end, the basic test settings are depicted next. Simulated estimates will be averaged over 100 independent Monte Carlo (MC) realizations without mentioning this explicitly each time. Performance of different schemes is evaluated in terms of the relative root mean-square error, i.e.,

$$\text{Relative error} := \frac{\text{dist}(\boldsymbol{z},\,\boldsymbol{x})}{\|\boldsymbol{x}\|}, \tag{5}$$

and the success rate among 100 trials, where a success will be claimed for a trial if the returned estimate incurs a relative error less than 10^{-5} [6]. Simulated tests under both noiseless and noisy Gaussian models are performed, corresponding to $\psi_i=\left|\boldsymbol{a}_i^{\mathcal{H}}\boldsymbol{x}+\eta_i\right|$ [27] with $\eta_i=0$ and $\eta_i\sim\mathcal{N}(0,\sigma^2)$, respectively, with i.i.d. $\boldsymbol{a}_i\sim\mathcal{N}(\boldsymbol{0},\boldsymbol{I}_n)$ or $\boldsymbol{a}_i\sim\mathcal{CN}(\boldsymbol{0},\boldsymbol{I}_n)$.

A. Truncated Generalized Gradient Stage

Let us rewrite the amplitude-based cost function in a matrix-vector representation as

$$\underset{\boldsymbol{z}\in\mathbb{R}^n}{\text{minimize}}\;\;\ell(\boldsymbol{z})=\frac{1}{2m}\big\|\boldsymbol{\psi}-|\boldsymbol{A}\boldsymbol{z}|\,\big\|^2 \tag{6}$$

where, with a slight abuse of notation, $|\boldsymbol{A}\boldsymbol{z}|:=[|\boldsymbol{a}_1^{\mathcal{T}}\boldsymbol{z}|\;\;\cdots\;\;|\boldsymbol{a}_m^{\mathcal{T}}\boldsymbol{z}|]^{\mathcal{T}}$. Apart from being nonconvex, $\ell(\boldsymbol{z})$ is also nondiffentiable, hence challenging the algorithmic design and analysis. In the presence of smoothness or convexity, convergence analysis of iterative algorithms relies either on continuity of the gradient (ordinary gradient methods) [53], or, on the convexity of the objective functional (subgradient methods) [54]. Although subgradient methods have found widespread applicability in nonsmooth optimization, they are *limited* to the class of convex functions [55, Page 4]. In nonconvex nonsmooth optimization settings, the so-termed *generalized gradient* broadens the scope of the (sub)gradient to the class of *almost everywhere* differentiable functions [56]. Consider now a continuous but not necessarily differentiable function $h(\boldsymbol{z})\in\mathbb{R}$ defined over an open region $\mathcal{S}\subseteq\mathbb{R}^n$.

Definition 1. *[57, Definition 1.1] The generalized gradient of a function h at $\boldsymbol{z}$, denoted by ∂h, is the convex hull of the set of limits of the form $\lim\nabla h(\boldsymbol{z}_k)$, where $\boldsymbol{z}_k\to\boldsymbol{z}$ as $k\to+\infty$, i.e.,*

$$\partial h(\boldsymbol{z}):=\text{conv}\Big\{\lim_{k\to+\infty}\nabla h(\boldsymbol{z}_k):\boldsymbol{z}_k\to\boldsymbol{z},\;\boldsymbol{z}_k\notin\mathcal{G}_\ell\Big\}$$

*where the symbol '*conv*' signifies the convex hull of a set, and $\mathcal{G}_\ell$ denotes the set of points in $\mathcal{S}$ at which h fails to be differentiable.*

Having introduced the notion of the generalized gradient, and with t denoting the iteration count, our approach to solving (6) amounts to iteratively refining the initial guess $\boldsymbol{z}_0$ (returned by our orthogonality-promoting initialization method to be detailed shortly) by means of the ensuing *truncated* generalized gradient iterations

$$\boldsymbol{z}_{t+1}=\boldsymbol{z}_t-\mu_t\,\partial\ell_{\text{tr}}(\boldsymbol{z}_t) \tag{7}$$

where $\mu_t>0$ is the step size, and a piece of the (truncated) generalized gradient $\partial\ell_{\text{tr}}(\boldsymbol{z}_t)$ is given by

$$\partial \ell_{\rm tr}(\boldsymbol{z}_t) := \frac{1}{m}\sum_{i\in\mathcal{I}_{t+1}} \left(\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}_t - \psi_i \frac{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}_t}{|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}_t|}\right)\boldsymbol{a}_i \tag{8}$$

for some index set $\mathcal{I}_{t+1} \subseteq [m] := \{1, 2, \ldots, m\}$ to be designed next. The convention $\frac{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}_t}{|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}_t|} := 0$ is adopted, if $\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}_t = 0$. It is easy to verify that the update in (7) with a full generalized gradient in (8) monotonically decreases the objective function value in (6).

Any stationary point $\boldsymbol{z}^*$ of $\ell(\boldsymbol{z})$ can be characterized by the following fixed-point equation [58], [59]

$$\boldsymbol{A}^{\mathcal{T}}\left(\boldsymbol{A}\boldsymbol{z}^* - \boldsymbol{\psi}\odot\frac{\boldsymbol{A}\boldsymbol{z}^*}{|\boldsymbol{A}\boldsymbol{z}^*|}\right) = \boldsymbol{0} \tag{9}$$

for entry-wise product $\odot$, which may have many solutions. Clearly, if $\boldsymbol{z}^*$ is a solution, so is $-\boldsymbol{z}^*$. Further, both solutions/global minimizers $\boldsymbol{x}$ and $-\boldsymbol{x}$ satisfy (9) due to the fact $\boldsymbol{A}\boldsymbol{x} - \boldsymbol{\psi}\odot\frac{\boldsymbol{A}\boldsymbol{x}}{|\boldsymbol{A}\boldsymbol{x}|} = \boldsymbol{0}$. Considering any stationary point $\boldsymbol{z}^* \neq \pm\boldsymbol{x}$ that has been adapted such that $\phi(\boldsymbol{z}^*) = 0$, one can write

$$\boldsymbol{z}^* = \boldsymbol{x} + (\boldsymbol{A}^{\mathcal{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\mathcal{T}}\left[\boldsymbol{\psi}\odot\left(\tfrac{\boldsymbol{A}\boldsymbol{z}^*}{|\boldsymbol{A}\boldsymbol{z}^*|} - \tfrac{\boldsymbol{A}\boldsymbol{x}}{|\boldsymbol{A}\boldsymbol{x}|}\right)\right]. \tag{10}$$

Thus, a necessary condition for $\boldsymbol{z}^* \neq \boldsymbol{x}$ in (10) is $\frac{\boldsymbol{A}\boldsymbol{z}^*}{|\boldsymbol{A}\boldsymbol{z}^*|} \neq \frac{\boldsymbol{A}\boldsymbol{x}}{|\boldsymbol{A}\boldsymbol{x}|}$. Expressed differently, there must be sign differences between $\boldsymbol{A}\boldsymbol{z}^*$ and $\boldsymbol{A}\boldsymbol{x}$ whenever one gets stuck with an undesirable stationary point $\boldsymbol{z}^*$. Inspired by this observation, it is reasonable to devise algorithms that can detect and separate out the generalized gradient components corresponding to mistakenly estimated signs $\left\{\frac{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}_t}{|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}_t|}\right\}$ along the iterates $\{\boldsymbol{z}_t\}$.

Precisely, if $\boldsymbol{z}_t$ and $\boldsymbol{x}$ lie in different sides of the hyperplane $\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z} = 0$, then the sign of $\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}_t$ will be different than that of $\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}$; that is, $\frac{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}}{|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}|} \neq \frac{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}}{|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}|}$. Specifically, one can re-write the i-th generalized gradient component as

$$\begin{aligned}\partial\ell_i(\boldsymbol{z}) &= \Big(\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z} - \psi_i\frac{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}}{|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}|}\Big)\boldsymbol{a}_i = \Big(\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z} - |\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}|\cdot\frac{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}}{|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}|}\Big)\boldsymbol{a}_i + \Big(\frac{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}}{|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}|} - \frac{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}}{|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}|}\Big)\psi_i\boldsymbol{a}_i \\ &= \boldsymbol{a}_i\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h} + \underbrace{\Big(\frac{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}}{|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}|} - \frac{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}}{|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}|}\Big)\psi_i\boldsymbol{a}_i}_{\triangleq\,\boldsymbol{r}_i}\end{aligned} \tag{11}$$

where $\boldsymbol{h} := \boldsymbol{z} - \boldsymbol{x}$. Intuitively, the SLLN asserts that averaging the first term $\boldsymbol{a}_i\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}$ over m instances approaches $\boldsymbol{h}$, which qualifies it as a desirable search direction. However, certain generalized gradient entries involve erroneously estimated signs of $\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}$; hence, nonzero $\boldsymbol{r}_i$ terms exert a negative influence on the search direction $\boldsymbol{h}$ by dragging the iterate away from $\boldsymbol{x}$, and they typically have sizable magnitudes as will be further elaborated in Remark (2) shortly.

Figure 1 demonstrates this from a geometric perspective, where the black dot denotes the origin, and the red dot the solution $\boldsymbol{x}$ whereas $-\boldsymbol{x}$ is omitted for ease of exposition. Assume without loss of generality that the i-th missing sign is positive, i.e., $\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x} = \psi_i$. As will be demonstrated in Theorem 1, with high probability, the initial estimate returned by our orthogonality-promoting method obeys $\|\boldsymbol{h}\| \leq \rho\|\boldsymbol{x}\|$ for some sufficiently small constant $\rho > 0$. Therefore, all points lying on or within the circle (or sphere in high-dimensional spaces) in Fig. 1 satisfy $\|\boldsymbol{h}\| \leq \rho\|\boldsymbol{x}\|$. If $\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z} = 0$ does not intersect with the circle, all points within the circle satisfy $\frac{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}}{|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}|} = \frac{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}}{|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}|}$ qualifying the i-th generalized gradient a desirable search (descent) direction in (11). If, on the other hand, $\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z} = 0$ intersects with the circle, then points lying with $\boldsymbol{x}$ on the same side of $\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z} = 0$ in Fig. 1 admit correctly estimated signs while points lying on

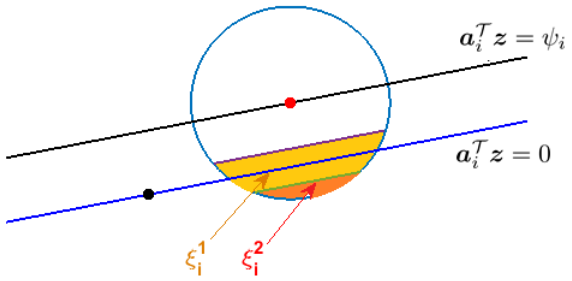

Fig. 1. The geometric understanding of the proposed truncation rule on the i-th gradient component involving $\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}=\psi_i$, where the red dot denotes the solution $\boldsymbol{x}$ and the black one is the origin. Hyperplanes $\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}=\psi_i$ and $\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}=0$ (of $\boldsymbol{z}\in\mathbb{R}^n$) passing through points $\boldsymbol{z}=\boldsymbol{x}$ and $\boldsymbol{z}=\boldsymbol{0}$, respectively, are shown.

different sides of $\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}=0$ with $\boldsymbol{x}$ have $\frac{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}}{|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}|}\neq\frac{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}}{|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}|}$, giving rise to a corrupted search direction in (11), so the corresponding generalized gradient component should be eliminated.

Nevertheless, it is difficult or even impossible to check whether the sign of $\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}_t$ equals that of $\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}$. Fortunately, as demonstrated in Fig. 1, most spurious generalized gradient components (those corrupted by nonzero $\boldsymbol{r}_i$ terms) hover around the watershed hyperplane $\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}_t=0$. For this reason, TAF includes only those components having $\boldsymbol{z}_t$ sufficiently away from its watershed, i.e.,

$$\mathcal{I}_{t+1}:=\left\{1\le i\le m\,\middle|\,\frac{|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}_t|}{|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}|}\ge\frac{1}{1+\gamma}\right\},\quad t\ge 0 \tag{12}$$

for an appropriately selected threshold $\gamma>0$. To be more specific, the light yellow color-coded area denoted by ξ_i^1 in Fig. 1 signifies the truncation region of $\boldsymbol{z}$, i.e., if $\boldsymbol{z}\in\xi_i^1$ obeying the condition in (12), the corresponding generalized gradient component $\partial\ell_i(\boldsymbol{z};\psi_i)$ will be thrown out. However, the truncation rule may mis-reject the 'good' gradients if $\boldsymbol{z}_t$ lies in the upper part of ξ_i^1; and 'bad' gradients may be missed as well if $\boldsymbol{z}_t$ belongs to the spherical cap ξ_i^2. Fortunately, as we will show in Lemmas 5 and 6, the probabilities of the miss and the mis-rejection are provably very small, hence precluding a noticeable influence on the descent direction. Although not perfect, it turns out that such a regularization rule succeeds in detecting and eliminating most corrupted generalized gradient components and hence maintaining a well-behaved search direction.

Regarding our gradient regularization rule in (12), two observations are in order.

Remark 1. The truncation rule in (12) includes only relatively sizable $\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}_t$'s, hence enforcing the smoothness of the (truncated) objective function $\ell_{\mathrm{tr}}(\boldsymbol{z}_t)$ at $\boldsymbol{z}_t$. Therefore, the truncated generalized gradient $\partial\ell_{\mathrm{tr}}(\boldsymbol{z})$ employed in (7) and (8) boils down to the ordinary gradient/Wirtinger derivative $\nabla\ell_{\mathrm{tr}}(\boldsymbol{z}_t)$ in the real-/complex-valued case.

Remark 2. As will be elaborated in (81) and (83), the quantities $(1/m)\sum_{i=1}^m\psi_i$ and $\max_{i\in[m]}\psi_i$ in (11) have magnitudes on the order of $\sqrt{\pi/2}\|\boldsymbol{x}\|$ and $\sqrt{m}\|\boldsymbol{x}\|$, respectively. In contrast, Proposition 1 asserts that the first term in (11) obeys $\|\boldsymbol{a}_i\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}\|\approx\|\boldsymbol{h}\|\le\rho\|\boldsymbol{x}\|$ for a sufficiently small $\rho\ll\sqrt{\pi/2}$. So spurious generalized gradient components typically have large magnitudes. It turns out that our gradient regularization rule in (12) also throws out gradient components of large sizes. To see this, for all $\boldsymbol{z}\in\mathbb{R}^n$

such that $\|\boldsymbol{h}\| \le \rho\|\boldsymbol{x}\|$ in (26), one can re-express

$$\sum_{i=1}^{m} \partial \ell_i(\boldsymbol{z}) = \sum_{i=1}^{m} \underbrace{\left(1 - \frac{|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}|}{|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}|}\right)}_{\triangleq \beta_i} \boldsymbol{a}_i\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z} \tag{13}$$

for some weight $\beta_i \in [-\infty, 1)$ assigned to the direction $\boldsymbol{a}_i\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z} \approx \boldsymbol{z}$ due to $\mathbb{E}[\boldsymbol{a}_i\boldsymbol{a}_i^{\mathcal{T}}] = \boldsymbol{I}_n$. Then $\partial\ell_i(\boldsymbol{z})$ of an excessively large size corresponds to a large $|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}|/|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}|$ in (13), or equivalently a small $|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}|/|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}|$ in (12), thus rendering the corresponding $\partial\ell_i(\boldsymbol{z})$ to be eliminated according to the truncation rule in (12).

We note that our truncation rule deviates from the intuition behind TWF, which also throws away gradient components corresponding to large-size $\{|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}_t|/|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}|\}$ in (12). As demonstrated by our analysis in Appendix E, it rarely happens that a gradient component having a large $|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}_t|/|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}|$ yields an incorrect sign of $\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}$ under a sufficiently accurate initialization. Further, discarding too many samples (those $i \notin \mathcal{T}_{t+1}$ in TWF [6, Section 2.1]) introduces large bias into $(1/m)\sum_{i\in\mathcal{T}_{t+1}}^{m} \boldsymbol{a}_i\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}$, so that TWF does not work well when m/n is small close to the information-limit of $m/n \approx 2$. In sharp contrast, the motivation and objective of our truncation rule in (12) is to directly sense and eliminate gradient components that involve mistakenly estimated signs with high probability. Numerical comparison depicted in Fig. 2 suggests that even when starting with the *same truncated spectral initialization*, TAF's refinement outperforms those of TWF and WF, demonstrating the merits of our gradient update rule over TWF/WF. Further, comparing TAF (gradient iterations in (7)-(8) with truncation in (12) initialized by the truncated spectral estimate) and AF (gradient iterations in (7)-(8) initialized by the truncated spectral estimate) corroborates the extreme power of our truncation rule in (12).

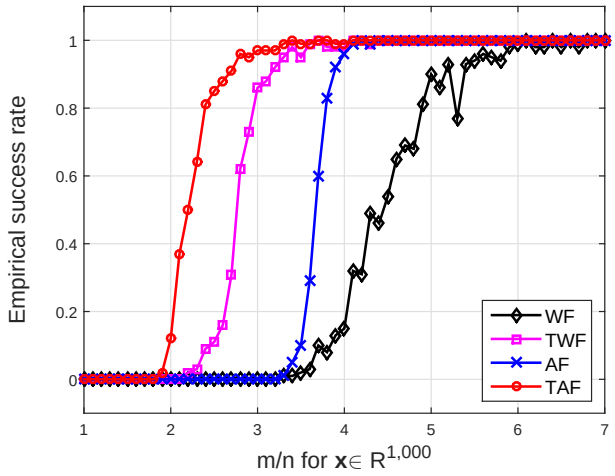

Fig. 2. Empirical success rate for WF, TWF, AF, and TAF with the same truncated spectral initialization under the noiseless real-valued Gaussian model.

B. Orthogonality-promoting Initialization Stage

Leveraging the SLLN, spectral initialization methods estimate $\boldsymbol{x}$ as the (appropriately scaled) leading eigenvector of $\boldsymbol{Y} := \frac{1}{m}\sum_{i\in\mathcal{T}_0} y_i\boldsymbol{a}_i\boldsymbol{a}_i^{\mathcal{T}}$, where $\mathcal{T}_0$ is an index set accounting for possible data truncation.

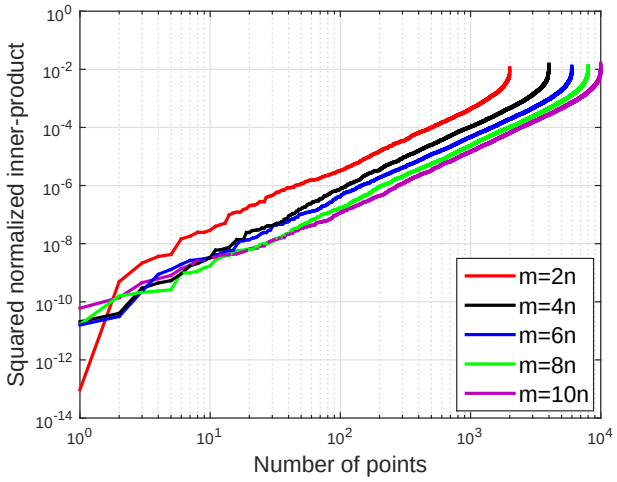

Fig. 3. Ordered squared normalized inner-product for pairs $\boldsymbol{x}$ and $\boldsymbol{a}_i$, $\forall i \in [m]$ with m/n varying by 2 from 2 to 10, and $n = 1,000$.

As asserted in [6], each summand $(\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x})^2\boldsymbol{a}_i\boldsymbol{a}_i^{\mathcal{T}}$ follows a heavy-tail probability density function lacking a moment generating function. This causes major performance degradation especially when the number of measurements is small. Instead of spectral initializations, we shall take another route to bypass this hurdle. To gain intuition for our initialization, a motivating example is presented first that reveals fundamental characteristics of high-dimensional random vectors.

A curious experiment: Fixing any nonzero vector $\boldsymbol{x} \in \mathbb{R}^n$, generate data $\psi_i = |\langle \boldsymbol{a}_i, \boldsymbol{x}\rangle|$ using i.i.d. $\boldsymbol{a}_i \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_n)$, $1 \le i \le m$. Then evaluate the following squared normalized inner-product

$$\cos^2\theta_i := \frac{|\langle \boldsymbol{a}_i, \boldsymbol{x}\rangle|^2}{\|\boldsymbol{a}_i\|^2\|\boldsymbol{x}\|^2} = \frac{\psi_i^2}{\|\boldsymbol{a}_i\|^2\|\boldsymbol{x}\|^2}, \quad 1 \le i \le m, \tag{14}$$

where θ_i is the angle between vectors $\boldsymbol{a}_i$ and $\boldsymbol{x}$. Consider ordering all $\{\cos^2\theta_i\}$ in an ascending fashion, and collectively denote them as $\boldsymbol{\xi} := [\cos^2\theta_{[m]} \ \cdots \ \cos^2\theta_{[1]}]^{\mathcal{T}}$ with $\cos^2\theta_{[1]} \ge \cdots \ge \cos^2\theta_{[m]}$. Figure 3 plots the ordered entries in $\boldsymbol{\xi}$ for m/n varying by 2 from 2 to 10 with $n = 1,000$. Observe that almost all $\{\boldsymbol{a}_i\}$ vectors have a squared normalized inner-product with $\boldsymbol{x}$ smaller than 10^{-2}, while half of the inner-products are less than 10^{-3}, which implies that $\boldsymbol{x}$ is nearly orthogonal to a large number of $\boldsymbol{a}_i$'s.

This example corroborates the folklore that random vectors in high-dimensional spaces are almost always nearly orthogonal to each other [52]. This inspired us to pursue an *orthogonality-promoting initialization method*. Our key idea is to approximate $\boldsymbol{x}$ by a vector that is most orthogonal to a subset of vectors $\{\boldsymbol{a}_i\}_{i\in\mathcal{I}_0}$, where $\mathcal{I}_0$ is an index set with cardinality $|\mathcal{I}_0| < m$ that includes indices of the smallest squared normalized inner-products $\{\cos^2\theta_i\}$. Since $\|\boldsymbol{x}\|$ appears in all inner-products, its exact value does not influence their ordering. Henceforth, we assume with no loss of generality that $\|\boldsymbol{x}\| = 1$.

Using data $\{(\boldsymbol{a}_i; \psi_i)\}$, evaluate $\cos^2\theta_i$ according to (14) for each pair $\boldsymbol{x}$ and $\boldsymbol{a}_i$. Instrumental for the ensuing derivations is noticing from the inherent near-orthogonal property of high-dimensional random vectors that the summation of $\cos^2\theta_i$ over indices $i \in \mathcal{I}_0$ should be very small; rigorous justification is deferred to Section V. It holds that $\sum_{i\in\mathcal{I}_0}\cos^2\theta_i = \sum_{i\in\mathcal{I}_0}\frac{|\langle \boldsymbol{a}_i,\boldsymbol{x}\rangle|^2}{\|\boldsymbol{a}_i\|^2\|\boldsymbol{x}\|^2} = \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}\left(\sum_{i\in\mathcal{I}_0}\frac{\boldsymbol{a}_i\boldsymbol{a}_i^{\mathcal{T}}}{\|\boldsymbol{a}_i\|^2}\right)\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}$ is negligibly small, yet $\boldsymbol{x}$ is unknown. Therefore, a meaningful approximation of $\boldsymbol{x}$, henceforth denoted by

$\boldsymbol{z}_0 \in \mathbb{R}^n$, can be obtained via minimizing the former with $\boldsymbol{x}$ replaced by the optimization variable $\boldsymbol{z}$, i.e.,

$$\underset{\|\boldsymbol{z}\|=1}{\text{minimize}} \quad \boldsymbol{z}^{\mathcal{T}} \left(\frac{1}{|\mathcal{I}_0|} \sum_{i \in \mathcal{I}_0} \frac{\boldsymbol{a}_i \boldsymbol{a}_i^{\mathcal{T}}}{\|\boldsymbol{a}_i\|^2} \right) \boldsymbol{z} \tag{15}$$

which amounts to finding the smallest eigenvalue and the associated eigenvector of $\boldsymbol{Y}_0 := \frac{1}{|\mathcal{I}_0|} \sum_{i \in \mathcal{I}_0} \frac{\boldsymbol{a}_i \boldsymbol{a}_i^{\mathcal{T}}}{\|\boldsymbol{a}_i\|^2} \succeq \boldsymbol{0}$ (The symbol $\succeq$ means positive semidefinite). Finding the smallest eigenvalue calls for eigen-decomposition or matrix inversion, each typically requiring computational complexity on the order of $\mathcal{O}(n^3)$. Such a computational burden can be intractable when n grows large. Applying a standard concentration result, we show how the computation can be significantly reduced.

Since $\boldsymbol{a}_i/\|\boldsymbol{a}_i\|$ has unit norm and is uniformly distributed on the unit sphere, it is uniformly spherically distributed.[2] Spherical symmetry implies that $\boldsymbol{a}_i/\|\boldsymbol{a}_i\|$ has zero mean and covariance matrix $\boldsymbol{I}_n/n$ [60]. Appealing again to the SLLN, the sample covariance matrix $\frac{1}{m} \sum_{i=1}^m \frac{\boldsymbol{a}_i \boldsymbol{a}_i^{\mathcal{T}}}{\|\boldsymbol{a}_i\|^2}$ approaches $\boldsymbol{I}_n/n$ as m grows. Simple derivations lead to

$$\sum_{i \in \mathcal{I}_0} \frac{\boldsymbol{a}_i \boldsymbol{a}_i^{\mathcal{T}}}{\|\boldsymbol{a}_i\|^2} = \sum_{i=1}^m \frac{\boldsymbol{a}_i \boldsymbol{a}_i^{\mathcal{T}}}{\|\boldsymbol{a}_i\|^2} - \sum_{i \in \bar{\mathcal{I}}_0} \frac{\boldsymbol{a}_i \boldsymbol{a}_i^{\mathcal{T}}}{\|\boldsymbol{a}_i\|^2} \approxeq \frac{m}{n} \boldsymbol{I}_n - \sum_{i \in \bar{\mathcal{I}}_0} \frac{\boldsymbol{a}_i \boldsymbol{a}_i^{\mathcal{T}}}{\|\boldsymbol{a}_i\|^2} \tag{16}$$

where $\bar{\mathcal{I}}_0$ is the complement of $\mathcal{I}_0$ in the set $[m]$.

Define $\boldsymbol{S} := [\boldsymbol{a}_1/\|\boldsymbol{a}_1\| \ \cdots \ \boldsymbol{a}_m/\|\boldsymbol{a}_m\|]^{\mathcal{T}} \in \mathbb{R}^{m \times n}$, and form $\bar{\boldsymbol{S}}_0$ by removing the rows of $\boldsymbol{S}$ if their indices do not belong to $\bar{\mathcal{I}}_0$. The task of seeking the smallest eigenvalue of $\boldsymbol{Y}_0 = \frac{1}{|\mathcal{I}_0|} \boldsymbol{S}_0^{\mathcal{T}} \boldsymbol{S}_0$ reduces to computing the largest eigenvalue of the designed matrix

$$\bar{\boldsymbol{Y}}_0 := \frac{1}{|\bar{\mathcal{I}}_0|} \bar{\boldsymbol{S}}_0^{\mathcal{T}} \bar{\boldsymbol{S}}_0, \tag{17}$$

namely,

$$\tilde{\boldsymbol{z}}_0 := \arg \max_{\|\boldsymbol{z}\|=1} \quad \boldsymbol{z}^{\mathcal{T}} \bar{\boldsymbol{Y}}_0 \boldsymbol{z} \tag{18}$$

which can be efficiently solved via simple power iterations. If, on the other hand, $\|\boldsymbol{x}\| \neq 1$, then the estimate $\tilde{\boldsymbol{z}}_0$ from (17) is further scaled so that its norm matches approximately that of $\boldsymbol{x}$, which is estimated to be $\sqrt{\frac{1}{m} \sum_{i=1}^m y_i}$, or more accurately $\sqrt{\frac{n \sum_{i=1}^m y_i}{\sum_{i=1}^m \|\boldsymbol{a}_i\|^2}}$. To see this, using the rotational invariance property of normal distributions as detailed in (66), it suffices to consider the case where $\boldsymbol{x} = \|\boldsymbol{x}\| \boldsymbol{e}_1$, with $\boldsymbol{e}_1$ denoting the first canonical vector of $\mathbb{R}^n$. Then it can be easily verified that

$$\frac{1}{m} \sum_{i=1}^m y_i = \frac{1}{m} \sum_{i=1}^m a_{i,1}^2 \|\boldsymbol{x}\|^2 \approx \|\boldsymbol{x}\|^2, \tag{19}$$

where the last arises from the concentration result $(1/m) \sum_{i=1}^m a_{i,1}^2 \approx \mathbb{E}[a_{i,1}^2] = \mathrm{cov}(a_{i,1}) = 1$ using the SLLN. Regarding the second estimate, one can rewrite its square as

$$\frac{n \sum_{i=1}^m y_i}{\sum_{i=1}^m \|\boldsymbol{a}_i\|^2} = \frac{1}{m} \sum_{i=1}^m y_i \cdot \frac{n}{(1/m) \cdot \sum_{i=1}^m \|\boldsymbol{a}_i\|^2}. \tag{20}$$

It is clear from (19) that the first term on the right hand side of (20) approximates $\|\boldsymbol{x}\|^2$. The second term approaches 1 because the denominator $(1/m) \cdot \sum_{i=1}^m \|\boldsymbol{a}_i\|^2 \approx n$ appealing again to the SLLN and

[2]A random vector $\boldsymbol{z} \in \mathbb{R}^n$ is said to be spherical (or spherically symmetric) if its distribution does not change under rotations of the coordinate system; that is, the distribution of $\boldsymbol{P}\boldsymbol{z}$ coincides with that of $\boldsymbol{z}$ for any given orthogonal $n \times n$ matrix $\boldsymbol{P}$.

$\boldsymbol{a}_i \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_n)$, thus rendering $\frac{n\sum_{i=1}^m y_i}{\sum_{i=1}^m \|\boldsymbol{a}_i\|^2} \approx \|\boldsymbol{x}\|^2$. Nevertheless, for simplicity, we choose to work with the first norm estimate to yield

$$\boldsymbol{z}_0 = \sqrt{\frac{\sum_{i=1}^m y_i}{m}}\tilde{\boldsymbol{z}}_0. \tag{21}$$

It is worth stressing that, comparing to $\boldsymbol{Y} := \frac{1}{m}\sum_{i\in\mathcal{T}_0} y_i \boldsymbol{a}_i\boldsymbol{a}_i^{\mathcal{T}}$ in spectral methods, our constructed matrix $\bar{\boldsymbol{Y}}_0$ in (17) does not depend on the observed data $\{y_i\}$ explicitly; the dependence is only through the choice of $\mathcal{I}_0$. Our orthogonality-promoting initialization thus enjoys two advantages over its spectral alternatives: a1) it does not suffer from heavy-tails of the fourth-order moments of Gaussian $\{\boldsymbol{a}_i\}$ vectors common in spectral initialization schemes; and, a2) it is less sensitive to the noisy data.

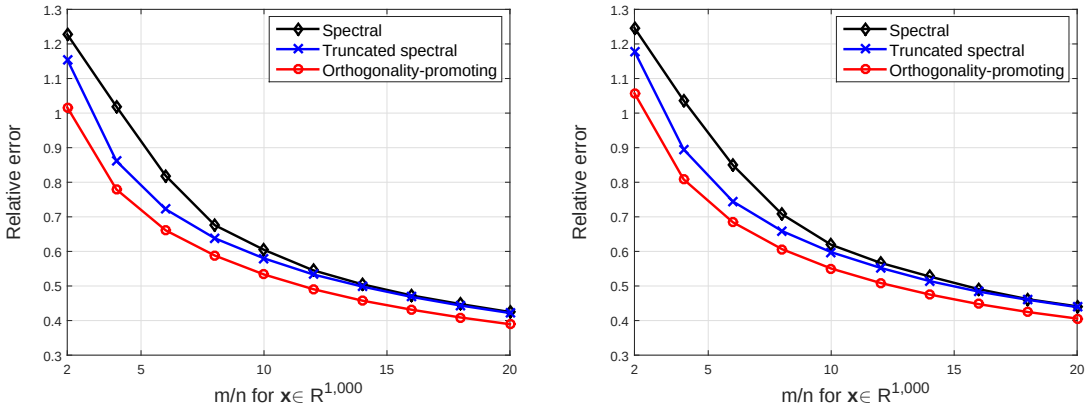

Fig. 4. Relative error of initial estimates versus m/n for: i) the spectral method [29]; ii) the truncated spectral method [6]; and iii) our orthogonality-promoting method with $n = 1,000$, and m/n varying by 2 from 2 to 20. Left: Noiseless real-valued Gaussian model with $\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_n)$, $\boldsymbol{a}_i \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_n)$, and $\eta_i = 0$. Right: Noisy real-valued Gaussian model with $\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_n)$, $\boldsymbol{a}_i \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_n)$, and $\sigma^2 = 0.2^2\|\boldsymbol{x}\|^2$.

Figure 4 compares three different initialization schemes including the spectral initialization [27], [29], the truncated spectral initialization [6], and the proposed orthogonality-promoting initialization methods. The relative error of their returned initial estimates versus the measurement/unknown ratio m/n is depicted under the noiseless and noisy real-valued Gaussian models, where $\boldsymbol{x} \in \mathbb{R}^{1,000}$ was randomly generated and m/n increases by 2 from 2 to 20. Apparently, all schemes enjoy improved performance as m/n increases in both noiseless and noisy settings. In particular, the proposed initialization method outperforms its spectral alternatives. Interestingly, the spectral and truncated spectral schemes exhibit similar performance when m/n becomes sufficiently large (e.g., $m/n \geq 14$ in the noiseless setup or $m/n \geq 16$ in the noisy one). This confirms that the truncation helps only if m/n is relatively small. Indeed, the truncation is effected by discarding measurements of excessively large or small sizes emerging from the heavy tails of the data distribution. Hence, its resulting advantage over the non-truncated spectral initialization narrows as the number of measurements increases, thus straightening out the heavy tails. In contrast, the orthogonality-promoting initialization method achieves consistently superior performance over its competing spectral alternatives under both noiseless and noisy Gaussian data.

Algorithm 1 Truncated amplitude flow (TAF) solver

1: **Input:** Data $\{\psi_i := |\langle \boldsymbol{a}_i, \boldsymbol{x}\rangle|\}_{i=1}^m$ and feature vectors $\{\boldsymbol{a}_i\}_{i=1}^m$; the maximum number of iterations T; by default, take constant step sizes $\mu = 0.6/1$ for real-/complex-valued models, truncation thresholds $|\overline{\mathcal{I}}_0| = \lceil \frac{1}{6}m \rceil$, [3] and $\gamma = 0.7$.

2: **Find** $\overline{\mathcal{I}}_0$ comprising of indices corresponding to the $|\overline{\mathcal{I}}_0|$ largest values of $\{\psi_i/\|\boldsymbol{a}_i\|\}$.

3: **Initialize** $\boldsymbol{z}_0$ to $\sqrt{\frac{\sum_{i=1}^m \psi_i^2}{m}}\tilde{\boldsymbol{z}}_0$, where $\tilde{\boldsymbol{z}}_0$ is the normalized leading eigenvector of $\overline{\boldsymbol{Y}}_0 := \frac{1}{|\overline{\mathcal{I}}_0|}\sum_{i\in\overline{\mathcal{I}}_0} \frac{\boldsymbol{a}_i\boldsymbol{a}_i^{\mathcal{T}}}{\|\boldsymbol{a}_i\|^2}$.

4: **Loop: for** $t = 0$ **to** $T-1$

$$\boldsymbol{z}_{t+1} = \boldsymbol{z}_t - \frac{\mu}{m}\sum_{i\in\mathcal{I}_{t+1}} \left(\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}_t - \psi_i \frac{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}_t}{|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}_t|}\right)\boldsymbol{a}_i$$

where $\mathcal{I}_{t+1} := \left\{1 \le i \le m \,\middle|\, \left|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}_t\right| \ge \frac{1}{1+\gamma}\psi_i\right\}$.

5: **Output:** $\boldsymbol{z}_T$.

III. Main Results

The TAF algorithm is summarized in Algorithm 1. Default values are set for pertinent algorithmic parameters. Postulating independent data samples $\{(\boldsymbol{a}_i; \psi_i)\}$ drawn from the noiseless real-valued Gaussian model, the following result establishes theoretical performance of TAF.

Theorem 1 (Exact recovery)**.** *Let $\boldsymbol{x} \in \mathbb{R}^n$ be an arbitrary signal vector, and consider (noise-free) measurements $\psi_i = |\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}|$, in which $\boldsymbol{a}_i \overset{i.i.d.}{\sim} \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_n)$, $1 \le i \le m$. Then with probability at least $1 - (m+5)\mathrm{e}^{-n/2} - \mathrm{e}^{-c_0 m} - 3/n^2$ for some universal constant $c_0 > 0$, the initialization $\boldsymbol{z}_0$ returned by the orthogonality-promoting method in Algorithm 1 satisfies*

$$\operatorname{dist}(\boldsymbol{z}_0, \boldsymbol{x}) \le \rho\,\|\boldsymbol{x}\| \tag{22}$$

with $\rho = 1/10$ (or any sufficiently small positive constant), provided that $m \ge c_1|\overline{\mathcal{I}}_0| \ge c_2 n$ for some numerical constants c_1, $c_2 > 0$, and sufficiently large n. Further, choosing a constant step size $\mu \le \mu_0$ along with a truncation level $1/2 \le \gamma \le 4$, and starting from any initial guess $\boldsymbol{z}_0$ satisfying (22)*, successive estimates of the TAF solver (tabulated in Algorithm 1) obey*

$$\operatorname{dist}(\boldsymbol{z}_t, \boldsymbol{x}) \le \rho\,(1-\nu)^t\,\|\boldsymbol{x}\|, \quad t = 0, 1, 2, \ldots \tag{23}$$

for some $0 < \nu < 1$, which holds with probability exceeding $1 - (m+5)\mathrm{e}^{-n/2} - 8\mathrm{e}^{-c_0 m} - 3/n^2$. Typical parameter values are $\mu_0 = 0.6$, and $\gamma = 0.7$.

Proof. The proof of Theorem 1 is relegated to Section V. □

Theorem 1 asserts that: i) TAF reconstructs the solution $\boldsymbol{x}$ exactly as soon as the number of equations is about the number of unknowns, which is theoretically order optimal. Our numerical tests demonstrate that for the real-valued Gaussian model, TAF achieves a success rate of 100% when m/n is as small as 3, which is slightly larger than the information limit of $m/n = 2$. This is a significant reduction in the sample complexity ratio, which is 5 for TWF and 7 for WF. Surprisingly, TAF enjoys also a success rate of over 50% when m/n is the information limit 2, which has not yet been presented for any existing algorithm;

[3] The symbol $\lceil\cdot\rceil$ is the ceiling operation returning the smallest integer greater than or equal to the given number.

see further discussion in Section IV; and, ii) TAF converges exponentially fast with the convergence rate independent of the dimension n. Specifically, TAF requires at most $\mathcal{O}(\log(1/\epsilon))$ iterations to achieve any given solution accuracy $\epsilon > 0$ (a.k.a., $\mathrm{dist}(\boldsymbol{z}_t, \boldsymbol{x}) \leq \epsilon \|\boldsymbol{x}\|$), with iteration cost $\mathcal{O}(mn)$. Since the truncation takes time on the order of $\mathcal{O}(m)$, the computational burden of TAF per iteration is dominated by the evaluation of the gradient components. The latter involves two matrix-vector multiplications that are computable in $\mathcal{O}(mn)$ flops, namely, $\boldsymbol{A}\boldsymbol{z}_t$ yields $\boldsymbol{u}_t$, and $\boldsymbol{A}^{\mathcal{T}}\boldsymbol{v}_t$ the gradient, where $\boldsymbol{v}_t := \boldsymbol{u}_t - \boldsymbol{\psi} \odot \frac{\boldsymbol{u}_t}{|\boldsymbol{u}_t|}$. Hence, the total running time of TAF is $\mathcal{O}(mn \log(1/\epsilon))$, which is proportional to the time taken to read the data $\mathcal{O}(mn)$.

Regarding stability of TAF, it is worth mentioning that TAF is stable under additive noise. To be more specific, under the noisy data model $\psi_i = |\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x} + \eta_i|$, it can be shown that the truncated amplitude flow estimates in Algorithm 1 satisfy

$$\mathrm{dist}\left(\boldsymbol{z}_t, \boldsymbol{x}\right) \lesssim (1-\nu)^t \left\|\boldsymbol{x}\right\| + \frac{1}{\sqrt{m}} \left\|\boldsymbol{\eta}\right\|, \quad t = 0,\, 1,\, 2,\, \ldots \tag{24}$$

with high probability for all $\boldsymbol{x} \in \mathbb{R}^n$, provided that $m \geq c_1|\overline{\mathcal{I}}_0| \geq c_2 n$ for sufficiently large n as well as bounded noise $\|\boldsymbol{\eta}\|_\infty \leq c_3 \|\boldsymbol{x}\|$ with $\boldsymbol{\eta} := [\eta_1 \ \cdots \ \eta_n]^{\mathcal{T}}$, where $0 < \nu < 1$, and $c_1,\, c_2,\, c_3 > 0$ are some universal constants. The proof can be directly adapted from those of Theorem 1 above and Theorem 2 in [6].

IV. Simulated Tests

Additional numerical tests evaluating performance of the proposed scheme relative to TWF/WF are presented in this section. For fairness, all pertinent algorithmic parameters involved in each scheme were set to their default values. The initial estimate was found based on 50 power iterations, and was subsequently refined by $T = 1,000$ gradient-like iterations in each scheme. The Matlab implementations of TAF are available at http://www.tc.umn.edu/~gangwang/TAF.

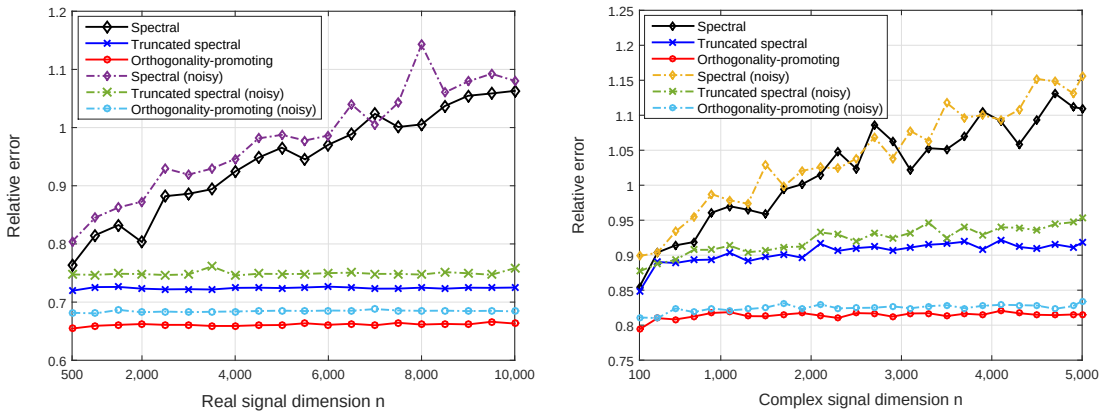

Fig. 5. The average relative error of estimates obtained from 100 MC trials using: i) the spectral method [27], [29]; ii) the truncated spectral method [6]; and iii) the proposed orthogonality-promoting method on noise-free (solid lines) and noisy (dotted lines) instances with $m/n = 6$, and n varying from $500/100$ to $10,000/5,000$ for real-/complex-valued vectors. Left: Real-valued Gaussian model with $\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_n)$, $\boldsymbol{a}_i \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_n)$, and $\sigma^2 = 0.2^2 \|\boldsymbol{x}\|^2$. Right: Complex-valued Gaussian model with $\boldsymbol{x} \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I}_n)$, $\boldsymbol{a}_i \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I}_n)$, and $\sigma^2 = 0.2^2 \|\boldsymbol{x}\|^2$.

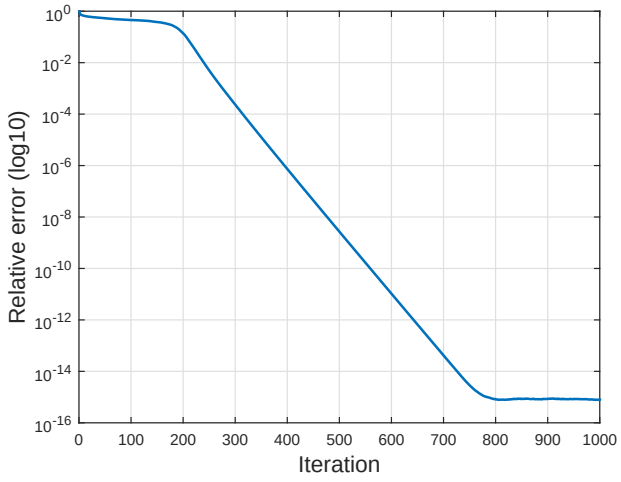

Fig. 6. Relative error versus iteration for TAF for a noiseless real-valued Gaussian model under the information-limit of $m = 2n-1$.

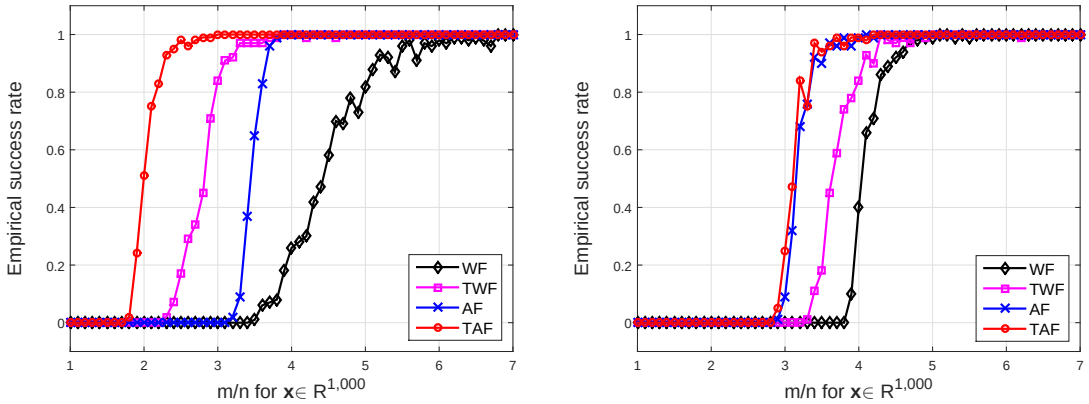

Fig. 7. Empirical success rate for WF, TWF, AF, and TAF with $n = 1,000$ and m/n varying by 0.1 from 1 to 7. Left: Noiseless real-valued Gaussian model with $\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_n)$ and $\boldsymbol{a}_i \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_n)$; Right: Noiseless complex-valued Gaussian model with $\boldsymbol{x} \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I}_n)$ and $\boldsymbol{a}_i \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I}_n)$.

Left panel in Fig. 5 presents the average relative error of three initialization methods on a series of noiseless/noisy real-valued Gaussian problems with $m/n = 6$ fixed, and n varying from 500 to 10^4, while those for the corresponding complex-valued Gaussian instances are shown in the right panel. Apparently, the proposed initialization method returns more accurate and robust estimates than the spectral ones. To demonstrate the extreme power of TAF, Fig. 6 plots the relative error of recovering a real-valued signal in logarithmic scale versus the iteration count under the information-limit of $m = 2n - 1$ noiseless i.i.d. Gaussian measurements [1]. In this case, since the returned initial estimate is relatively far from the optimal solution (see Fig. 4), TAF converges slowly for the first 200 iterations or so due to elimination

of a significant amount of 'bad' generalized gradient components (corrupted by mistakenly estimated signs). As the iterate gets more accurate and lands within a small-size neighborhood of $\boldsymbol{x}$, TAF converges exponentially fast to the globally optimal solution. It is worth emphasizing that no existing method succeeds in this case. Figure 7 compares the empirical success rate of three schemes under both real-valued and complex-valued Gaussian models with $n = 10^3$ and m/n varying by 0.1 from 1 to 7. Moreover, for real-valued vectors, TAF achieves a success rate of over 50% when $m/n = 2$, and guarantees perfect recovery from about $3n$ measurements; while for complex-valued ones, TAF enjoys a success rate of 95% when $m/n = 3.4$, and ensures perfect recovery from about $4.5n$ measurements.

The next experiment further evaluates efficacy of the proposed initialization method, simulating all schemes initialized by the truncated spectral initial estimate [6] and the orthogonality-promoting initial estimate. Apparently, all schemes except WF admit a significant performance improvement when initialized by the proposed orthogonality-promoting initialization relative to the truncated spectral initialization. Nevertheless, TAF with the orthogonality-promoting initialization enjoys the superior performance over all simulated schemes.

Finally, to test the effectiveness and scalability of TAF in real-world conditions, the Milky Way Galaxy image [4] $\boldsymbol{X} \in \mathbb{R}^{1080\times 1920\times 3}$ shown in Fig. 9 is involved. The first two indices encode the pixel locations, and the third the RGB (red, green, blue) color bands. Consider a type of physically realizable measurements termed coded diffraction patterns (CDP) with random masks [26], [29], [6]. Letting $\boldsymbol{x} \in \mathbb{R}^n$ be a vectorization of a certain band of $\boldsymbol{X}$ and postulating a number K of random masks, one can write

$$\boldsymbol{\psi}^{(k)} = \left|\boldsymbol{F}\boldsymbol{D}^{(k)}\boldsymbol{x}\right|, \quad 1 \le k \le K, \tag{25}$$

where $\boldsymbol{F}$ denotes the $n\times n$ discrete Fourier transform matrix, and $\boldsymbol{D}^{(k)}$ is a diagonal matrix holding entries sampled uniformly at random from $\{1, -1, j, -j\}$ (phase delays) on its diagonal, with j denoting the imaginary unit. Each $\boldsymbol{D}^{(k)}$ represents a random mask placed after the object to modulate the illumination patterns [26]. When $K = 6$ masks were employed in our experiment, the total number of quadratic measurements becomes $m = nK$. Specifically, the algorithm was run independently on each of the three bands. A number 100 of power iterations were used to obtain an initialization, which was refined by 100 gradient-type iterations. The relative errors after our orthogonality-promoting initialization and after 100 TAF iterations are 0.6807 and 9.8631×10^{-5}, respectively, and the recovered images are displayed in Fig. 9. In sharp contrast, TWF returns images of corresponding relative errors 1.3801 and 1.3409, which are far away from the ground truth.

Regarding running times in all performed experiments, TAF converges slightly faster than TWF, while both are markedly faster than WF. All experiments were implemented using MATLAB on an Intel CPU @ 3.4 GHz (32 GB RAM) computer.

V. Proofs

To prove Theorem 1, this section establishes a few lemmas and the main ideas, while technical details are deferred to the Appendix. Relative to WF and TWF, our objective function involves nonsmoothness and nonconvexity, rendering the proof of exact recovery of TAF nontrivial. In addition, our initialization method starts from a rather different perspective than the spectral alternatives, so the thoughts and tools involved in proving performance of our initialization deviate from those of the spectral methods [27], [29], [6]. Part of the proof is adapted from [29], [6] and [58].

The proof of Theorem 1 consists of two parts: Section V-A justifies the performance of the proposed orthogonality-promoting initialization, which essentially achieves any given constant relative error as soon

[4]The Milky Way Galaxy image is downloaded from http://pics-about-space.com/milky-way-galaxy.

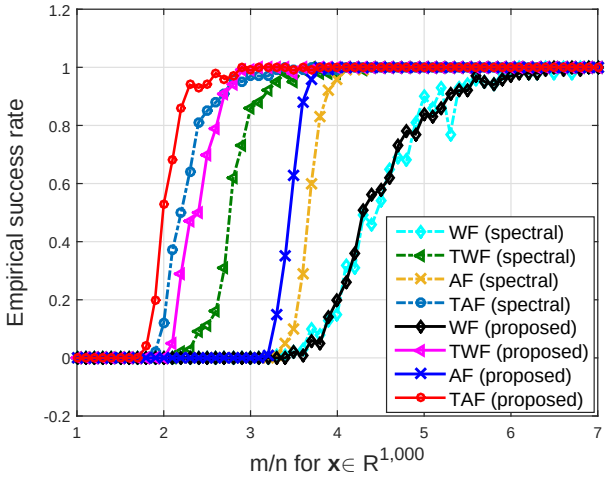

Fig. 8. Empirical success rate for WF, TWF, AF, and TAF initialized by the truncated spectral and the orthogonality-promoting initializations with $n = 1,000$ and m/n varying by 0.1 from 1 to 7.

as the number of equations is on the order of the number of unknowns, namely, $m \asymp n$.[5] Section V-B demonstrates theoretical convergence of TAF to the solution of the quadratic system in (1) at a geometric rate provided that the initial estimate has a sufficiently small constant relative error as in (22). The two stages of TAF can be performed independently, meaning that other better initialization methods, if available, could be adopted to initialize our truncated generalized gradient iterations; likewise, our initialization method can also be applied to initialize other iterative optimization algorithms.

A. Constant Relative Error by Orthogonality-promoting Initialization

This section concentrates on proving guaranteed performance of the proposed orthogonality-promoting initialization method, as asserted in the following proposition.

Proposition 1. *Fix $\boldsymbol{x} \in \mathbb{R}^n$ arbitrarily, and consider the noiseless case $\psi_i = |\boldsymbol{a}_i^{\mathcal{T}} \boldsymbol{x}|$, where $\boldsymbol{a}_i \overset{i.i.d.}{\sim} \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_n)$, $1 \le i \le m$. Then with probability at least $1 - (m+5)\mathrm{e}^{-n/2} - \mathrm{e}^{-c_0 m} - 3/n^2$ for some universal constant $c_0 > 0$, the initialization $\boldsymbol{z}_0$ returned by the orthogonality-promoting method satisfies*

$$\operatorname{dist}(\boldsymbol{z}_0, \boldsymbol{x}) \le \rho \|\boldsymbol{x}\| \tag{26}$$

for $\rho = 1/10$ or any positive constant, with the proviso that $m \ge c_1|\overline{\mathcal{I}}_0| \ge c_2 n$ for some numerical constants c_1, $c_2 > 0$ and sufficiently large n.

Due to homogeneity in (26), it suffices to work with the case where $\|\boldsymbol{x}\| = 1$. Assume for the moment that $\|\boldsymbol{x}\| = 1$ is known and $\boldsymbol{z}_0$ has been scaled such that $\|\boldsymbol{z}_0\| = 1$ in (21). Subsequently, the error between the employed $\boldsymbol{x}$'s norm estimate $\sqrt{\frac{1}{m}\sum_{i=1}^m y_i}$ and the unknown norm $\|\boldsymbol{x}\| = 1$ will be accounted for at the end of this Section. Instrumental in proving Proposition 1 is the following result, whose proof is deferred to Appendix A.

[5]The notations $\phi(n) = \mathcal{O}(g(n))$ or $\phi(n) \gtrsim g(n)$ (respectively, $\phi(n) \lesssim g(n)$) means there exists a numerical constant $c > 0$ such that $\phi(n) \le cg(n)$, while $\phi(n) \asymp g(n)$ means $\phi(n)$ and $g(n)$ are orderwise equivalent.

Fig. 9. The recovered Milky Way Galaxy images after i) truncated spectral initialization (top); ii) orthogonality-promoting initialization (middle); and iii) 100 TAF gradient iterations refining the orthogonality-promoting initialization (bottom).

Lemma 1. *Consider the noiseless data* $\psi_i = |\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}|$*, where* $\boldsymbol{a}_i \overset{i.i.d.}{\sim} \mathcal{N}(\boldsymbol{0}, \boldsymbol{I}_n)$*,* $1 \le i \le m$*. For any unit vector* $\boldsymbol{x} \in \mathbb{R}^n$*, there exists a vector* $\boldsymbol{u} \in \mathbb{R}^n$ *with* $\boldsymbol{u}^{\mathcal{T}}\boldsymbol{x} = 0$ *and* $\|\boldsymbol{u}\| = 1$ *such that*

$$\frac{1}{2}\left\|\boldsymbol{x}\boldsymbol{x}^{\mathcal{T}} - \boldsymbol{z}_0\boldsymbol{z}_0^{\mathcal{T}}\right\|_F^2 \le \frac{\left\|\overline{\boldsymbol{S}}_0\boldsymbol{u}\right\|^2}{\left\|\overline{\boldsymbol{S}}_0\boldsymbol{x}\right\|^2} \tag{27}$$

for $\boldsymbol{z}_0 = \tilde{\boldsymbol{z}}_0$*, where the unit vector* $\tilde{\boldsymbol{z}}_0$ *is given in* (18)*, and* $\overline{\boldsymbol{S}}_0$ *is formed by removing the rows of* $\boldsymbol{S} := \left[\boldsymbol{a}_1/\left\|\boldsymbol{a}_1\right\| \ \cdots \ \boldsymbol{a}_m/\left\|\boldsymbol{a}_m\right\|\right]^{\mathcal{T}} \in \mathbb{R}^{m\times n}$*, if their indices do not belong to the set* $\overline{\mathcal{I}}_0$ *specified in Algorithm 1.*

We now turn to prove Proposition 1. The first step consists in upper-bounding the term on the right-hand-side of (27). Specifically, its numerator term will be upper bounded, and the denominator term lower

bounded, which are summarized in Lemma 2 and Lemma 3, whose proofs can be found in Appendix B and Appendix C, respectively.

Lemma 2. *In the setup of Lemma 1, if $|\overline{\mathcal{I}}_0| \geq c_1' n$, then the next*

$$\left\|\overline{\boldsymbol{S}}_0 \boldsymbol{u}\right\|^2 \leq 1.01|\overline{\mathcal{I}}_0|/n \tag{28}$$

holds with probability at least $1-2\mathrm{e}^{-c_K n}$, where c_2' and c_K are some universal constants.

Lemma 3. *In the setup of Lemma 1, the following holds with probability at least $1-(m+1)\mathrm{e}^{-n/2}-\mathrm{e}^{-c_0 m}-3/n^2$,*

$$\left\|\overline{\boldsymbol{S}}_0 \boldsymbol{x}\right\|^2 \geq \frac{0.99|\overline{\mathcal{I}}_0|}{2.3n}\left[1+\log\left(m/|\overline{\mathcal{I}}_0|\right)\right] \tag{29}$$

provided that $|\overline{\mathcal{I}}_0| \geq c_1' n$, $m \geq c_2'|\overline{\mathcal{I}}_0|$, and $m \geq c_3' n$ for some absolute constants c_1', c_2', $c_3' > 0$, and sufficiently large n.

Therefore, putting the upper and lower bounds in (28) and (29) together, one arrives at

$$\frac{\left\|\overline{\boldsymbol{S}}_0 \boldsymbol{u}\right\|^2}{\left\|\overline{\boldsymbol{S}}_0 \boldsymbol{x}\right\|^2} \leq \frac{2.4}{1+\log\left(m/|\overline{\mathcal{I}}_0|\right)} \triangleq \kappa \tag{30}$$

which holds with probability at least $1-(m+3)\mathrm{e}^{-n/2}-\mathrm{e}^{-c_0 m}-3/n^2$, with the proviso that $m \geq c_1'|\overline{\mathcal{I}}_0|$, and $m \geq c_2' n$, $|\overline{\mathcal{I}}_0| \geq c_3' n$ for some absolute constants c_1', c_2', $c_3' > 0$, and sufficiently large n.

Apparently, the bound κ in (30) is meaningful only when the ratio $\log(m/|\overline{\mathcal{I}}_0|) > 1.4$, i.e., $m/|\overline{\mathcal{I}}_0| > 4$, because the left hand side expressible in terms of $\sin^2\theta$ enjoys a trivial upper bound 1. Henceforth, we will work with the case where $m/|\overline{\mathcal{I}}_0| > 4$. Empirically, $\lfloor m/|\overline{\mathcal{I}}_0| \rfloor = 6$ or equivalently $|\overline{\mathcal{I}}_0| = \lceil \frac{1}{6} m \rceil$ in Algorithm 1 works well when m/n is relatively small. Note further that the bound κ can be made arbitrarily small by letting $m/|\overline{\mathcal{I}}_0|$ be large enough. Without any loss of generality, let us take $\kappa := 0.001$. An additional step leads to the wanted bound on the distance between $\tilde{\boldsymbol{z}}_0$ and $\boldsymbol{x}$; similar arguments can be found in [29, Section 7.8]. Recall that

$$|\boldsymbol{x}^{\mathcal{T}} \tilde{\boldsymbol{z}}_0|^2 = \cos^2\theta = 1-\sin^2\theta \geq 1-\kappa, \tag{31}$$

so one has

$$\begin{aligned} \mathrm{dist}^2(\tilde{\boldsymbol{z}}_0, \boldsymbol{x}) &\leq \|\tilde{\boldsymbol{z}}_0\|^2 + \|\boldsymbol{x}\|^2 - 2|\boldsymbol{x}^{\mathcal{T}} \tilde{\boldsymbol{z}}_0| \\ &\leq \left(2-2\sqrt{1-\kappa}\right)\|\boldsymbol{x}\|^2 \\ &\approx \kappa \|\boldsymbol{x}\|^2 . \end{aligned} \tag{32}$$

Coming back to the case in which $\|\boldsymbol{x}\|$ is unknown stated prior to Lemma 1, the unit eigenvector $\tilde{\boldsymbol{z}}_0$ is scaled by the estimate of $\|\boldsymbol{x}\|$ to yield the initial guess $\boldsymbol{z}_0 = \sqrt{\frac{1}{m}\sum_{i=1}^m y_i}\tilde{\boldsymbol{z}}_0$. Using the results in Lemma 7.8 in [29], the following holds with high probability

$$\|\boldsymbol{z}_0 - \tilde{\boldsymbol{z}}_0\| = |\|\boldsymbol{z}_0\| - 1| \leq (1/20)\|\boldsymbol{x}\| . \tag{33}$$

Summarizing the two inequalities, we conclude that

$$\mathrm{dist}(\boldsymbol{z}_0, \boldsymbol{x}) \leq \|\boldsymbol{z}_0 - \tilde{\boldsymbol{z}}_0\| + \mathrm{dist}(\tilde{\boldsymbol{z}}_0, \boldsymbol{x}) \leq (1/10)\|\boldsymbol{x}\| . \tag{34}$$

The initialization thus obeys $\mathrm{dist}(\boldsymbol{z}_0, \boldsymbol{x})/\|\boldsymbol{x}\| \leq 1/10$ for any $\boldsymbol{x} \in \mathbb{R}^n$ with high probability provided that $m \geq c_1|\overline{\mathcal{I}}_0| \geq c_2 n$ holds for some universal constants c_1, $c_2 > 0$ and sufficiently large n.

B. Exact Recovery from Noiseless Data

We now prove that with accurate enough initial estimates, TAF converges at a geometric rate to $\boldsymbol{x}$ with high probability (i.e., the second part of Theorem 1). To be specific, with initialization obeying (26) in Proposition 1, TAF reconstructs the solution $\boldsymbol{x}$ exactly in linear time. In this direction, it suffices to demonstrate that the TAF's update rule (i.e., Step 4 in Algorithm 1) is locally contractive within a sufficiently small neighborhood of $\boldsymbol{x}$, as asserted in the following proposition.

Proposition 2 (Local error contraction)**.** *Consider the noise-free measurements $\psi_i = \left|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}\right|$ with i.i.d. Gaussian design vectors $\boldsymbol{a}_i \sim \mathcal{N}(\boldsymbol{0},\, \boldsymbol{I}_n)$, $1 \le i \le m$, and fix any $1/2 \le \gamma \le 4$. Then there exist universal constants $c_0,\, c_1 > 0$ and $0 < \nu < 1$ such that with probability at least $1 - 7\mathrm{e}^{-c_0 m}$, the following holds*

$$\mathrm{dist}^2\left(\boldsymbol{z} + \frac{\mu}{m}\nabla\ell_{\mathrm{tr}}(\boldsymbol{z}),\, \boldsymbol{x}\right) \le (1-\nu)\mathrm{dist}^2\left(\boldsymbol{z},\, \boldsymbol{x}\right) \tag{35}$$

for all $\boldsymbol{x},\, \boldsymbol{z} \in \mathbb{R}^n$ obeying the condition (26) *for sufficiently small $\rho > 0$ with the proviso that $m \ge c_1 n$ and that the constant step size μ satisfying $0 < \mu \le \mu_0$ for some $\mu_0 > 0$.*

Proposition 2 demonstrates that the distance of TAF's successive iterates to $\boldsymbol{x}$ is monotonically decreasing once the algorithm enters a small-size neighborhood around $\boldsymbol{x}$. This neighborhood is commonly referred to as the *basin of attraction*; see further discussions in [29], [33], [6], [37], [40]. In other words, as soon as one lands within the basin of attraction, TAF's iterates remain in this region and will be attracted to $\boldsymbol{x}$ exponentially fast. To substantiate Proposition 2, recall the concept of the *local regularity condition*, which was first developed in [29] and plays a fundamental role in establishing linear convergence to global optimum of nonconvex optimization approaches such as WF/TWF [29], [33], [6], [30]. Now consider the update rule of TAF

$$\boldsymbol{z}_{t+1} = \boldsymbol{z}_t - \frac{\mu}{m}\nabla\ell_{\mathrm{tr}}(\boldsymbol{z}_t), \quad \forall t \ge 0, \tag{36}$$

where the truncated gradient $\nabla\ell_{\mathrm{tr}}(\boldsymbol{z}_t)$ (as elaborated in Remark 1) evaluated at some point $\boldsymbol{z}_t \in \mathbb{R}^n$ is given by

$$\frac{1}{m}\nabla\ell_{\mathrm{tr}}(\boldsymbol{z}_t) \triangleq \frac{1}{m}\sum_{i\in\mathcal{I}}\left(\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}_t - \psi_i\frac{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}_t}{|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}_t|}\right)\boldsymbol{a}_i.$$

The truncated gradient $\nabla\ell_{\mathrm{tr}}(\boldsymbol{z})$ is said to satisfy the local regularity condition, or $\mathrm{LRC}(\mu,\lambda,\epsilon)$ for some constant $\lambda > 0$, provided that

$$\left\langle \frac{1}{m}\nabla\ell_{\mathrm{tr}}(\boldsymbol{z}),\, \boldsymbol{h}\right\rangle \ge \frac{\mu}{2}\left\|\frac{1}{m}\nabla\ell_{\mathrm{tr}}(\boldsymbol{z})\right\|^2 + \frac{\lambda}{2}\left\|\boldsymbol{h}\right\|^2 \tag{37}$$

holds for all $\boldsymbol{z} \in \mathbb{R}^n$ such that $\|\boldsymbol{h}\| = \|\boldsymbol{z}-\boldsymbol{x}\| \le \epsilon\|\boldsymbol{x}\|$ for some constant $0 < \epsilon < 1$, where the ball $\|\boldsymbol{z}-\boldsymbol{x}\| \le \epsilon\|\boldsymbol{x}\|$ is the so-called *basin of attraction*. Simple linear algebra along with the regularity condition in (37) leads to

$$\begin{aligned}
\mathrm{dist}^2\left(\boldsymbol{z} - \frac{\mu}{m}\nabla\ell_{\mathrm{tr}}(\boldsymbol{z}), \boldsymbol{x}\right) &= \left\|\boldsymbol{z} - \frac{\mu}{m}\nabla\ell_{\mathrm{tr}}(\boldsymbol{z}) - \boldsymbol{x}\right\|^2 \\
&= \|\boldsymbol{h}\|^2 - 2\mu\left\langle \boldsymbol{h}, \frac{1}{m}\nabla\ell_{\mathrm{tr}}(\boldsymbol{z})\right\rangle + \left\|\frac{\mu}{m}\nabla\ell_{\mathrm{tr}}(\boldsymbol{z})\right\|^2 \qquad (38)\\
&\le \|\boldsymbol{h}\|^2 - 2\mu\left(\frac{\mu}{2}\left\|\frac{1}{m}\nabla\ell_{\mathrm{tr}}(\boldsymbol{z})\right\|^2 + \frac{\lambda}{2}\|\boldsymbol{h}\|^2\right) + \left\|\frac{\mu}{m}\nabla\ell_{\mathrm{tr}}(\boldsymbol{z})\right\|^2 \\
&= (1-\lambda\mu)\|\boldsymbol{h}\|^2 = (1-\lambda\mu)\,\mathrm{dist}^2(\boldsymbol{z},\boldsymbol{x}) \qquad (39)
\end{aligned}$$

for all $\boldsymbol{z}$ obeying $\|\boldsymbol{h}\| \le \epsilon \|\boldsymbol{x}\|$. Clearly, if the $\mathrm{LRC}(\mu, \lambda, \epsilon)$ is proved for TAF, our goal (35) follows upon letting $\nu := \lambda\mu$.

1) Proof of the local regularity condition in (37)*:* By definition, justifying the local regularity condition in (37) entails controlling the norm of the truncated gradient $\frac{1}{m}\nabla \ell_{\mathrm{tr}}(\boldsymbol{z})$, i.e., bounding the last term in (38). Recall that

$$\frac{1}{m}\nabla \ell_{\mathrm{tr}}(\boldsymbol{z}) = \frac{1}{m}\sum_{i\in\mathcal{I}}\left(\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z} - \psi_i \frac{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}}{\left|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}\right|}\right)\boldsymbol{a}_i \triangleq \frac{1}{m}\boldsymbol{A}\boldsymbol{v} \tag{40}$$

where $\mathcal{I} := \{1 \le i \le m | |\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}| \ge |\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}|/(1+\gamma)\}$, and $\boldsymbol{v} := [v_1 \ \cdots \ v_m]^{\mathcal{T}} \in \mathbb{R}^m$ with $v_i := \frac{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}}{|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}|}\left(|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}| - \psi_i\right)\mathbb{1}_{\left\{|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}| \ge |\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}|/(1+\gamma)\right\}}$. Now, consider

$$|v_i|^2 = \left|\left(\left|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}\right| - \left|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}\right|\right)\mathbb{1}_{\left\{|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}| \ge |\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}|/(1+\gamma)\right\}}\right|^2 \le \left|\left|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}\right| - \left|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}\right|\right|^2 \le \left|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}\right|^2 = a_{i,1}^2\|\boldsymbol{h}\|^2, \tag{41}$$

where $\boldsymbol{h} = \boldsymbol{z} - \boldsymbol{x}$. Observe that $a_{i,1}^2$ obeys the *Chi-square* distribution with $k = 1$ degrees of freedom; yet due to our working assumption $\|\boldsymbol{a}_i\| \le \sqrt{2.3n}$, it has mean $\mathbb{E}[a_{i,1}^2] \le k = 1$. So fixing any $0 < \delta' < 1$ and applying the one-sided Bernstein-type inequality, the following holds with probability at least $1 - \mathrm{e}^{-m\delta'^2/2}$ [60, Proposition 5.16]

$$\|\boldsymbol{v}\|^2 = \sum_{i=1}^m v_i^2 \le \sum_{i=1}^m a_{i,1}^2\|\boldsymbol{h}\|^2 \le (1+\delta')m\|\boldsymbol{h}\|^2. \tag{42}$$

On the other hand, standard matrix concentration results confirm that the largest singular value of $\boldsymbol{A} = [\boldsymbol{a}_1 \ \cdots \ \boldsymbol{a}_m]^{\mathcal{T}}$ with i.i.d. Gaussian $\{\boldsymbol{a}_i\}$ satisfies $\sigma_1 := \|\boldsymbol{A}\| \le (1+\delta'')\sqrt{m}$ for some $\delta'' > 0$ with probability exceeding $1 - 2\mathrm{e}^{-c_0 m}$ as soon as $m \ge c_1 n$ for sufficiently large $c_1 > 0$, where $c_1 > 0$ is a universal constant depending on δ'' [60, Remark 5.25]. Putting together (40), (41), and (42) yields

$$\left\|\frac{1}{m}\nabla \ell_{\mathrm{tr}}(\boldsymbol{z})\right\| \le \frac{1}{m}\|\boldsymbol{A}\| \cdot \|\boldsymbol{v}\| \le (1+\delta')(1+\delta'')\|\boldsymbol{h}\| \le (1+\delta)^2\|\boldsymbol{h}\|, \quad \delta := \max\{\delta', \delta''\} \tag{43}$$

which holds with high probability. This condition essentially asserts that the truncated gradient of the objective function $\ell(\boldsymbol{z})$ or the search direction is well behaved (the function value does not vary too much).

Notice that to prove the LRC, it suffices to show that the truncated gradient $\frac{1}{m}\nabla \ell_{\mathrm{tr}}(\boldsymbol{z})$ ensures sufficient descent [40], i.e., it obeys a uniform lower bound along the search direction $\boldsymbol{h}$ taking the form

$$\left\langle \frac{1}{m}\nabla \ell_{\mathrm{tr}}(\boldsymbol{z}), \boldsymbol{h}\right\rangle \gtrsim \|\boldsymbol{h}\|^2 \tag{44}$$

which occupies the remaining of this section. Formally, this can be stated as follows.

Proposition 3. *Consider the noiseless measurements $\psi_i = |\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}|$ and fix any sufficiently small constant $\epsilon > 0$. There exist universal constants $c_0, c_1 > 0$ such that if $m > c_1 n$, then the following holds with probability exceeding $1 - 4\mathrm{e}^{-c_0 m}$*

$$\left\langle \boldsymbol{h}, \frac{1}{m}\nabla \ell_{\mathrm{tr}}(\boldsymbol{z})\right\rangle \ge 2\left(1 - \zeta_1 - \zeta_2 - 2\epsilon\right)\|\boldsymbol{h}\|^2 \tag{45}$$

for all $\boldsymbol{x}, \boldsymbol{z} \in \mathbb{R}^n$ such that $\|\boldsymbol{h}\| / \|\boldsymbol{x}\| \le \rho$ for $0 < \rho \le 1/10$ and any fixed $1/2 \le \gamma \le 4$.

Lemma 4. *Fix any $\gamma > 0$. For each $i \in [m]$, define the following events*

$$\mathcal{E}_i := \left\{ \frac{|\boldsymbol{a}_i^{\mathcal{T}} \boldsymbol{z}|}{|\boldsymbol{a}_i^{\mathcal{T}} \boldsymbol{x}|} \geq \frac{1}{1+\gamma} \right\}, \tag{46}$$

$$\mathcal{D}_i := \left\{ \frac{|\boldsymbol{a}_i^{\mathcal{T}} \boldsymbol{h}|}{|\boldsymbol{a}_i^{\mathcal{T}} \boldsymbol{x}|} \geq \frac{2+\gamma}{1+\gamma} \right\}, \tag{47}$$

$$\text{and} \quad \mathcal{K}_i := \left\{ \frac{\boldsymbol{a}_i^{\mathcal{T}} \boldsymbol{z}}{|\boldsymbol{a}_i^{\mathcal{T}} \boldsymbol{z}|} \neq \frac{\boldsymbol{a}_i^{\mathcal{T}} \boldsymbol{x}}{|\boldsymbol{a}_i^{\mathcal{T}} \boldsymbol{x}|} \right\} \tag{48}$$

where $\boldsymbol{h} = \boldsymbol{z} - \boldsymbol{x}$. Under the condition $\|\boldsymbol{h}\| / \|\boldsymbol{x}\| \leq \rho$, the following inclusion holds

$$\mathcal{E}_i \cap \mathcal{K}_i \subseteq \mathcal{D}_i. \tag{49}$$

Proof. From Fig. 1, it is clear that if $\boldsymbol{z} \in \xi_i^2$, then the sign of $\boldsymbol{a}_i^{\mathcal{T}} \boldsymbol{z}$ will be different than that of $\boldsymbol{a}_i^{\mathcal{T}} \boldsymbol{x}$. The region ξ_i^2, however, can be specified by the conditions that $\frac{\boldsymbol{a}_i^{\mathcal{T}} \boldsymbol{z}}{|\boldsymbol{a}_i^{\mathcal{T}} \boldsymbol{z}|} \neq \frac{\boldsymbol{a}_i^{\mathcal{T}} \boldsymbol{x}}{|\boldsymbol{a}_i^{\mathcal{T}} \boldsymbol{x}|}$ and $\frac{|\boldsymbol{a}_i^{\mathcal{T}} \boldsymbol{h}|}{|\boldsymbol{a}_i^{\mathcal{T}} \boldsymbol{x}|} \geq 1 + \frac{1}{1+\gamma} = \frac{2+\gamma}{1+\gamma}$. Under our initialization condition $\|\boldsymbol{h}\| / \|\boldsymbol{x}\| \leq \rho$, it is self-evident that $\mathcal{D}_i$ describes two spherical caps that contain ξ_i^2. Hence, it holds that $\mathcal{E}_i \cap \mathcal{K}_i = \xi_i^2 \subseteq \mathcal{D}_i$. □

Along the lines of (11), rewrite the truncated gradient as

$$\begin{aligned} \frac{1}{2m} \nabla \ell_{\text{tr}}(\boldsymbol{z}) &= \frac{1}{m} \sum_{i=1}^{m} \left(\boldsymbol{a}_i^{\mathcal{T}} \boldsymbol{z} - \left| \boldsymbol{a}_i^{\mathcal{T}} \boldsymbol{x} \right| \frac{\boldsymbol{a}_i^{\mathcal{T}} \boldsymbol{z}}{|\boldsymbol{a}_i^{\mathcal{T}} \boldsymbol{z}|} \right) \boldsymbol{a}_i \mathbb{1}_{\mathcal{E}_i} \\ &= \frac{1}{m} \sum_{i=1}^{m} \boldsymbol{a}_i \boldsymbol{a}_i^{\mathcal{T}} \boldsymbol{h} \mathbb{1}_{\mathcal{E}_i} - \frac{1}{m} \sum_{i=1}^{m} \left(\frac{\boldsymbol{a}_i^{\mathcal{T}} \boldsymbol{z}}{|\boldsymbol{a}_i^{\mathcal{T}} \boldsymbol{z}|} - \frac{\boldsymbol{a}_i^{\mathcal{T}} \boldsymbol{x}}{|\boldsymbol{a}_i^{\mathcal{T}} \boldsymbol{x}|} \right) \left| \boldsymbol{a}_i^{\mathcal{T}} \boldsymbol{x} \right| \boldsymbol{a}_i \mathbb{1}_{\mathcal{E}_i}. \end{aligned} \tag{50}$$

Using the definitions and properties in Lemma 4, one further arrives at

$$\begin{aligned} \left\langle \frac{1}{2m} \nabla \ell_{\text{tr}}(\boldsymbol{z}),\, \boldsymbol{h} \right\rangle &\geq \frac{1}{m} \sum_{i=1}^{m} \left(\boldsymbol{a}_i^{\mathcal{T}} \boldsymbol{h} \right)^2 \mathbb{1}_{\mathcal{E}_i} - \frac{1}{m} \sum_{i=1}^{m} \left| \boldsymbol{a}_i^{\mathcal{T}} \boldsymbol{x} \right| \left| \boldsymbol{a}_i^{\mathcal{T}} \boldsymbol{h} \right| \mathbb{1}_{\mathcal{E}_i \cap \mathcal{K}_i} \\ &\geq \frac{1}{m} \sum_{i=1}^{m} \left(\boldsymbol{a}_i^{\mathcal{T}} \boldsymbol{h} \right)^2 \mathbb{1}_{\mathcal{E}_i} - \frac{1}{m} \sum_{i=1}^{m} \left| \boldsymbol{a}_i^{\mathcal{T}} \boldsymbol{x} \right| \left| \boldsymbol{a}_i^{\mathcal{T}} \boldsymbol{h} \right| \mathbb{1}_{\mathcal{D}_i} \\ &\geq \frac{1}{m} \sum_{i=1}^{m} \left(\boldsymbol{a}_i^{\mathcal{T}} \boldsymbol{h} \right)^2 \mathbb{1}_{\mathcal{E}_i} - \frac{1+\gamma}{2+\gamma} \cdot \frac{1}{m} \sum_{i=1}^{m} \left(\boldsymbol{a}_i^{\mathcal{T}} \boldsymbol{h} \right)^2 \mathbb{1}_{\mathcal{D}_i} \end{aligned} \tag{51}$$

where the last inequality arises from the property $\left| \boldsymbol{a}_i^{\mathcal{T}} \boldsymbol{x} \right| \leq \frac{1+\gamma}{2+\gamma} \left| \boldsymbol{a}_i^{\mathcal{T}} \boldsymbol{h} \right|$ by the definition of $\mathcal{D}_i$.

Proving the regularity condition boils down to lower bounding the right-hand side of (51), specifically, to lower bounding the first term and to upper bounding the second one. Apparently, the first term approximately gives $\|\boldsymbol{h}\|^2$ by the SLLN as long as our truncation procedure does not eliminate too many generalized gradient components (i.e., summands in the first term). Regarding the second, one would expect its contribution to be small under our initialization condition in (26) and as the relative error $\|\boldsymbol{h}\| / \|\boldsymbol{x}\|$ decreases. Specifically, under our initialization, $\mathcal{D}_i$ is provably a rare event, thus eliminating the possibility of the second term exerting a noticeable influence on the first term. Rigorous analyses concerning the two terms are elaborated in Lemma 5 and Lemma 6, whose proofs can be found in Appendix D and Appendix E, respectively.

Lemma 5. *Fix $\gamma \geq 1/2$ and $\rho \leq 1/10$, and let $\mathcal{E}_i$ be defined in* (46)*. For independent random variables $W \sim \mathcal{N}(0,\,1)$ and $Z \sim \mathcal{N}(0,\,1)$, set*

$$\zeta_1 := 1 - \min\left\{ \mathbb{E}\left[\mathbb{1}_{\left\{ \left| \frac{1-\rho}{\rho} + \frac{W}{Z} \right| \geq \frac{\sqrt{1.01}}{\rho(1+\gamma)} \right\}} \right],\ \mathbb{E}\left[Z^2 \mathbb{1}_{\left\{ \left| \frac{1-\rho}{\rho} + \frac{W}{Z} \right| \geq \frac{\sqrt{1.01}}{\rho(1+\gamma)} \right\}} \right] \right\}. \tag{52}$$

Then for any $\epsilon > 0$ and any vector $\boldsymbol{h}$ obeying $\|\boldsymbol{h}\| / \|\boldsymbol{x}\| \leq \rho$, the following holds with probability exceeding $1 - 2\mathrm{e}^{-c_5 \epsilon^2 m}$

$$\frac{1}{m} \sum_{i=1}^{m} \left(\boldsymbol{a}_i^{\mathcal{T}} \boldsymbol{h} \right)^2 \mathbb{1}_{\mathcal{E}_i} \geq (1 - \zeta_1 - \epsilon) \left\| \boldsymbol{h} \right\|^2, \tag{53}$$

provided that $m > (c_6 \cdot \epsilon^{-2} \log \epsilon^{-1}) n$ for some universal constants $c_5,\, c_6 > 0$.

To have a sense of how large the quantities involved in (5) are, when $\gamma = 0.7$ and $\rho = 1/10$, it holds $\mathbb{E}\left[\mathbb{1}_{\left\{ \left| \frac{1-\rho}{\rho} + \frac{W}{Z} \right| \geq \frac{\sqrt{1.01}}{\rho(1+\gamma)} \right\}} \right] \approx 0.92$, and $\mathbb{E}\left[Z^2 \mathbb{1}_{\left\{ \left| \frac{1-\rho}{\rho} + \frac{W}{Z} \right| \geq \frac{\sqrt{1.01}}{\rho(1+\gamma)} \right\}} \right] \approx 0.99$, hence leading to $\zeta_1 \approx 0.08$.

Having derived a lower bound for the first term in the right-hand side of (51), it remains to deal with the second one.

Lemma 6. *Fix $1/2 \leq \gamma \leq 4$ and $\rho \leq 1/10$, and let $\mathcal{D}_i$ be defined in* (47)*. For any constant $\epsilon > 0$, there exists some universal constants $c_5,\, c_6 > 0$ such that*

$$\frac{1}{m} \sum_{i=1}^{m} \left(\boldsymbol{a}_i^{\mathcal{T}} \boldsymbol{h} \right)^2 \mathbb{1}_{\mathcal{D}_i} \leq (\zeta_2' + \epsilon) \left\| \boldsymbol{h} \right\|^2 \tag{54}$$

holds with probability at least $1 - 2\mathrm{e}^{-c_5 \epsilon^2 m}$ provided that $m/n > \left(c_6 \cdot \epsilon^{-2} \log \epsilon^{-1} \right)$ for some universal constants $c_5,\, c_6 > 0$, where $\zeta_2' := 1.5\sqrt{(1+\gamma)\rho}$.

Taking results in (51), (53), and (54) together, choosing m/n exceeding some sufficiently large constant such that $c_0 \leq c_5 \epsilon^2$, and denoting $\zeta_2 := \zeta_2'(1+\gamma)/(2+\gamma)$, then with probability exceeding $1 - 4\mathrm{e}^{-c_0 m}$, the following

$$\left\langle \boldsymbol{h},\, \frac{1}{2m} \nabla \ell_{\mathrm{tr}}(\boldsymbol{z}) \right\rangle \geq (1 - \zeta_1 - \zeta_2 - 2\epsilon) \left\| \boldsymbol{h} \right\|^2 \tag{55}$$

holds for all $\boldsymbol{x}$ and $\boldsymbol{z}$ such that $\|\boldsymbol{h}\| / \|\boldsymbol{x}\| \leq \rho$ for $0 < \rho \leq 1/10$ and any fixed $1/2 \leq \gamma \leq 4$. This combining with (37) and (39) proves Proposition 2 for appropriately chosen $\mu > 0$ and $\lambda > 0$.

To conclude this section, an estimate for the working step size is provided next. To be specific, plugging the results in (43) and (45) into (38) suggests that

$$\begin{aligned} \mathrm{dist}^2 \left(\boldsymbol{z} - \frac{\mu}{m} \nabla \ell_{\mathrm{tr}}(\boldsymbol{z}),\, \boldsymbol{x} \right) &= \left\| \boldsymbol{h} \right\|^2 - 2\mu \left\langle \boldsymbol{h},\, \frac{1}{m} \nabla \ell_{\mathrm{tr}}(\boldsymbol{z}) \right\rangle + \left\| \frac{\mu}{m} \nabla \ell_{\mathrm{tr}}(\boldsymbol{z}) \right\|^2 \qquad (56) \\ &\leq \left\{ 1 - \mu \left[4\left(1 - \xi_1 - \xi_2 - 2\epsilon\right) - \mu(1+\delta)^4 \right] \right\} \left\| \boldsymbol{h} \right\|^2 \\ &\overset{\triangle}{=} (1 - \nu) \left\| \boldsymbol{h} \right\|^2. \qquad (57) \end{aligned}$$

Taking ϵ and δ to be sufficiently small, one obtains the feasible range of the step size for TAF

$$\mu \leq \frac{4\left(0.99 - \xi_1 - \xi_2\right)}{1.02} \overset{\triangle}{=} \mu_0, \tag{58}$$

thus concluding the proof of Theorem 1.

VI. Concluding Remarks

This paper developed a linear-time algorithm termed TAF for solving generally unstructured systems of random quadratic equations. Our TAF algorithm builds on three key ingredients: a novel orthogonality-promoting initialization, along with a simple yet effective gradient truncation rule, as well as scalable gradient-like iterations. Numerical tests using synthetic data and real images corroborate the superior performance of TAF over state-of-the-art solvers of the same type. A few timely and pertinent future research directions are worth pointing out. First, in parallel with spectral initialization methods, the proposed orthogonality-promoting initialization can be applied for semidefinite optimization [36], [37], matrix completion [61], [40], as well as blind deconvolution [38], [39]. It is also interesting to investigate suitable gradient regularization rules in more general nonconvex optimization settings. Furthermore, extending the theory to the more challenging case where $\boldsymbol{a}_i$'s are generated from the coded diffraction pattern model [26] constitutes another meaningful direction.

APPENDIX

By homogeneity, it suffices to work with the case where $\|\boldsymbol{x}\| = 1$.

A. Proof of Lemma 1

It is easy to check that

$$\begin{aligned}\frac{1}{2}\left\|\boldsymbol{x}\boldsymbol{x}^{\mathcal{T}} - \tilde{\boldsymbol{z}}_0\boldsymbol{z}_0^{\mathcal{T}}\right\|_F^2 &= \frac{1}{2}\|\boldsymbol{x}\|^4 + \frac{1}{2}\|\tilde{\boldsymbol{z}}_0\|^4 - \left|\boldsymbol{x}^{\mathcal{T}}\tilde{\boldsymbol{z}}_0\right|^2 \\ &= 1 - \left|\boldsymbol{x}^{\mathcal{T}}\tilde{\boldsymbol{z}}_0\right|^2 \\ &= 1 - \cos^2\theta \end{aligned} \tag{59}$$

where $0 \le \theta \le \pi$ is the angle between the spaces spanned by $\boldsymbol{x}$ and $\tilde{\boldsymbol{z}}_0$. Then one can write

$$\boldsymbol{x} = \cos\theta\,\tilde{\boldsymbol{z}}_0 + \sin\theta\,\tilde{\boldsymbol{z}}_0^{\perp}, \tag{60}$$

where $\tilde{\boldsymbol{z}}_0^{\perp} \in \mathbb{R}^n$ is a unit vector that is orthogonal to $\tilde{\boldsymbol{z}}_0$ and has a nonnegative inner product with $\boldsymbol{x}$. Likewise, one can express

$$\boldsymbol{x}^{\perp} := -\sin\theta\,\tilde{\boldsymbol{z}}_0 + \cos\theta\,\tilde{\boldsymbol{z}}_0^{\perp}, \tag{61}$$

in which $\boldsymbol{x}^{\perp} \in \mathbb{R}^n$ is a unit vector orthogonal to $\boldsymbol{x}$.

Since $\tilde{\boldsymbol{z}}_0$ is the solution to the maximum eigenvalue problem

$$\tilde{\boldsymbol{z}}_0 := \arg\max_{\|\boldsymbol{z}\|=1} \quad \boldsymbol{z}^{\mathcal{T}}\overline{\boldsymbol{Y}}_0\boldsymbol{z} \tag{62}$$

for $\overline{\boldsymbol{Y}}_0 := \frac{1}{|\overline{\mathcal{I}}_0|}\overline{\boldsymbol{S}}_0^{\mathcal{T}}\overline{\boldsymbol{S}}_0$, it is the leading eigenvector of $\overline{\boldsymbol{Y}}_0$, i.e., $\overline{\boldsymbol{Y}}_0\tilde{\boldsymbol{z}}_0 = \lambda_1\tilde{\boldsymbol{z}}_0$, where $\lambda_1 > 0$ is the largest eigenvalue of $\overline{\boldsymbol{Y}}_0$. Premultiplying (60) and (61) by $\overline{\boldsymbol{S}}_0$ yields

$$\overline{\boldsymbol{S}}_0\boldsymbol{x} = \cos\theta\,\overline{\boldsymbol{S}}_0\tilde{\boldsymbol{z}}_0 + \sin\theta\,\overline{\boldsymbol{S}}_0\tilde{\boldsymbol{z}}_0^{\perp}, \tag{63a}$$

$$\overline{\boldsymbol{S}}_0\boldsymbol{x}^{\perp} = -\sin\theta\,\overline{\boldsymbol{S}}_0\tilde{\boldsymbol{z}}_0 + \cos\theta\,\overline{\boldsymbol{S}}_0\tilde{\boldsymbol{z}}_0^{\perp}. \tag{63b}$$

Pythagoras' relationship now gives

$$\left\|\overline{\boldsymbol{S}}_0\boldsymbol{x}\right\|^2 = \cos^2\theta\left\|\overline{\boldsymbol{S}}_0\tilde{\boldsymbol{z}}_0\right\|^2 + \sin^2\theta\left\|\overline{\boldsymbol{S}}_0\tilde{\boldsymbol{z}}_0^{\perp}\right\|^2, \tag{64a}$$

$$\left\|\overline{\boldsymbol{S}}_0\boldsymbol{x}^{\perp}\right\|^2 = \sin^2\theta\left\|\overline{\boldsymbol{S}}_0\tilde{\boldsymbol{z}}_0\right\|^2 + \cos^2\theta\left\|\overline{\boldsymbol{S}}_0\tilde{\boldsymbol{z}}_0^{\perp}\right\|^2, \tag{64b}$$

where the cross-terms vanish because $\tilde{\boldsymbol{z}}_0^{\mathcal{T}}\overline{\boldsymbol{S}}_0^{\mathcal{T}}\overline{\boldsymbol{S}}_0\tilde{\boldsymbol{z}}_0^{\perp} = |\overline{\mathcal{I}}_0|\tilde{\boldsymbol{z}}_0^{\mathcal{T}}\overline{\boldsymbol{Y}}_0\tilde{\boldsymbol{z}}_0^{\perp} = \lambda_1|\overline{\mathcal{I}}_0|\tilde{\boldsymbol{z}}_0^{\mathcal{T}}\tilde{\boldsymbol{z}}_0^{\perp} = 0$ following from the definition of $\tilde{\boldsymbol{z}}_0^{\perp}$.

We next construct the following expression

$$\begin{aligned}&\sin^2\theta\left\|\overline{\boldsymbol{S}}_0\boldsymbol{x}\right\|^2 - \left\|\overline{\boldsymbol{S}}_0\boldsymbol{x}^{\perp}\right\|^2 \\ &= \sin^2\theta\cos^2\theta\left\|\overline{\boldsymbol{S}}_0\tilde{\boldsymbol{z}}_0\right\|^2 + \sin^4\theta\left\|\overline{\boldsymbol{S}}_0\tilde{\boldsymbol{z}}_0^{\perp}\right\|^2 - \sin^2\theta\left\|\overline{\boldsymbol{S}}_0\tilde{\boldsymbol{z}}_0\right\|^2 - \cos^2\theta\left\|\overline{\boldsymbol{S}}_0\tilde{\boldsymbol{z}}_0^{\perp}\right\|^2 \\ &= \sin^2\theta\left(\cos^2\theta\left\|\overline{\boldsymbol{S}}_0\tilde{\boldsymbol{z}}_0\right\|^2 - \left\|\overline{\boldsymbol{S}}_0\tilde{\boldsymbol{z}}_0\right\|^2 + \sin^2\theta\left\|\overline{\boldsymbol{S}}_0\tilde{\boldsymbol{z}}_0^{\perp}\right\|^2\right) - \cos^2\theta\left\|\overline{\boldsymbol{S}}_0\tilde{\boldsymbol{z}}_0^{\perp}\right\|^2 \\ &= \sin^4\theta\left(\left\|\overline{\boldsymbol{S}}_0\tilde{\boldsymbol{z}}_0^{\perp}\right\|^2 - \left\|\overline{\boldsymbol{S}}_0\tilde{\boldsymbol{z}}_0\right\|^2\right) - \cos^2\theta\left\|\overline{\boldsymbol{S}}_0\tilde{\boldsymbol{z}}_0^{\perp}\right\|^2 \\ &\le 0\end{aligned}$$

where $\overline{\boldsymbol{S}}_0^{\mathcal{T}}\overline{\boldsymbol{S}}_0 \succeq \boldsymbol{0}$, so $\|\overline{\boldsymbol{S}}_0\tilde{\boldsymbol{z}}_0^{\perp}\|^2 - \|\overline{\boldsymbol{S}}_0\tilde{\boldsymbol{z}}_0\|^2 \leq 0$ holds for any unit vector $\tilde{\boldsymbol{z}}_0^{\perp} \in \mathbb{R}^n$ arising from the fact that $\tilde{\boldsymbol{z}}_0$ maximizes the term in (18), hence yielding

$$\sin^2\theta = 1 - \cos^2\theta \leq \frac{\left\|\overline{\boldsymbol{S}}_0\boldsymbol{x}^{\perp}\right\|^2}{\left\|\overline{\boldsymbol{S}}_0\boldsymbol{x}\right\|^2}. \tag{65}$$

Upon letting $\boldsymbol{u} = \boldsymbol{x}^{\perp}$, the last inequality taken together with (59) concludes the proof of (27).

B. Proof of Lemma 2

Recall that rows in $\overline{\boldsymbol{S}}_0 \in \mathbb{R}^{|\overline{\mathcal{I}}_0|\times n}$, hereafter denoted by $\boldsymbol{s}_i^{\mathcal{T}} \in \mathbb{R}^{1\times n}$, $\forall i \in [|\overline{\mathcal{I}}_0|]$, are drawn uniformly on the unit sphere. The uniformly spherical distribution is rotationally invariant, so it suffices to prove the results in the case where $\boldsymbol{x} = \boldsymbol{e}_1$ with $\boldsymbol{e}_1$ being the first canonical vector in $\mathbb{R}^n$. Indeed, any unit vector $\boldsymbol{x}$ can be expressed as $\boldsymbol{x} = \boldsymbol{U}\boldsymbol{e}_1$ for some orthogonal transformation $\boldsymbol{U} \in \mathbb{R}^{n\times n}$. To see this, consider the following [28]

$$|\langle \boldsymbol{s}_i, \boldsymbol{x}\rangle|^2 = |\langle \boldsymbol{s}_i, \boldsymbol{U}\boldsymbol{e}_1\rangle|^2 = \left|\left\langle \boldsymbol{U}^{\mathcal{T}}\boldsymbol{s}_i, \boldsymbol{e}_1\right\rangle\right|^2 \overset{d}{=} |\langle \boldsymbol{s}_i, \boldsymbol{e}_1\rangle|^2, \tag{66}$$

where $\overset{d}{=}$ means terms involved on both sides of the equality have the same distribution. Thus, the problem of finding any unit-normed $\boldsymbol{x}$ is equivalent to that of finding $\boldsymbol{e}_1$. Henceforth, we assume without any loss of generality that $\boldsymbol{x} = \boldsymbol{e}_1$.

Considering a unit vector $\boldsymbol{x}^{\perp}$ such that $\boldsymbol{x}^{\mathcal{T}}\boldsymbol{x}^{\perp} = \boldsymbol{e}_1^{\mathcal{T}}\boldsymbol{x}^{\perp} = 0$, there exists a unit vector $\boldsymbol{d} \in \mathbb{R}^{n-1}$ such that $\boldsymbol{x}^{\perp} = \begin{bmatrix}0 & \boldsymbol{d}^{\mathcal{T}}\end{bmatrix}^{\mathcal{T}}$. So it holds that

$$\left\|\overline{\boldsymbol{S}}_0\boldsymbol{x}^{\perp}\right\|^2 = \left\|\overline{\boldsymbol{S}}_0\begin{bmatrix}0 & \boldsymbol{d}^{\mathcal{T}}\end{bmatrix}^{\mathcal{T}}\right\|^2 = \left\|\boldsymbol{F}\boldsymbol{d}\right\|^2, \tag{67}$$

where $\boldsymbol{F} \in \mathbb{R}^{|\overline{\mathcal{I}}_0|\times(n-1)}$ is obtained through deleting the first column in $\overline{\boldsymbol{S}}_0$, denoted by $\overline{\boldsymbol{S}}_{0,1}$, i.e., $\overline{\boldsymbol{S}}_0 = \begin{bmatrix}\overline{\boldsymbol{S}}_{0,1} & \boldsymbol{F}\end{bmatrix}$. Letting $\boldsymbol{F} := \begin{bmatrix}\boldsymbol{f}_1 & \cdots & \boldsymbol{f}_{|\overline{\mathcal{I}}_0|}\end{bmatrix}^{\mathcal{T}}$, one can readily write $\boldsymbol{s}_i = \begin{bmatrix}s_{i,1} & \boldsymbol{f}_i^{\mathcal{T}}\end{bmatrix}^{\mathcal{T}}$, $\forall i \in [|\overline{\mathcal{I}}_0|] := \{1, \ldots, |\overline{\mathcal{I}}_0|\}$. Uniformly spherically distributed $\boldsymbol{s}_i \in \mathbb{R}^n$ has statistics $\mathbb{E}[\boldsymbol{s}_i] = \boldsymbol{0}$, and $\mathbb{E}[\boldsymbol{s}_i\boldsymbol{s}_i^{\mathcal{T}}] = \frac{1}{n}\boldsymbol{I}_n$ [62]. Leveraging the linearity of expectation operator, one arrives at

$$\mathbb{E}\left[\boldsymbol{s}_i\right] = \mathbb{E}\begin{bmatrix}s_{i,1}\\ \boldsymbol{f}_i\end{bmatrix} = \begin{bmatrix}\mathbb{E}[s_{i,1}]\\ \mathbb{E}[\boldsymbol{f}_i]\end{bmatrix} = \boldsymbol{0},\ \forall i \tag{68}$$

to yield

$$\mathbb{E}[\boldsymbol{f}_i] = \boldsymbol{0},\ \forall i. \tag{69}$$

A similar argument holds for the second-order moment

$$\mathbb{E}\left[\boldsymbol{s}_i\boldsymbol{s}_i^{\mathcal{T}}\right] = \begin{bmatrix}\mathbb{E}[s_{i,1}^2] & \mathbb{E}[s_{i,1}\boldsymbol{f}_i^{\mathcal{T}}]\\ \mathbb{E}[s_{i,1}\boldsymbol{f}_i] & \mathbb{E}[\boldsymbol{f}_i\boldsymbol{f}_i^{\mathcal{T}}]\end{bmatrix} = \frac{1}{n}\boldsymbol{I}_n,\ \forall i \tag{70}$$

leading to

$$\mathbb{E}\left[\boldsymbol{f}_i\boldsymbol{f}_i^{\mathcal{T}}\right] = \frac{1}{n}\boldsymbol{I}_{n-1},\ \forall i. \tag{71}$$

Recall that a random vector $\boldsymbol{z} \in \mathbb{R}^n$ is said to be *isotropic* if it has zero-mean and identity covariance matrix [60, Definition 5.19]. Then recognize, from (69) and (71), that a proper scaling of $\boldsymbol{f}_i$ renders $\sqrt{n}\boldsymbol{f}_i$ isotropic. Further, it is known that a spherical random vector is subgaussian, and its subgaussian norm is bounded by an absolute constant [60]. Indeed, this comes from the following geometric argument: using rotational invariance of the uniform spherical distribution $\mathcal{S}^{n-1}$ in $\mathbb{R}^n$, it holds that, given any $\epsilon \geq 0$,

the spherical cap $\{\boldsymbol{s}_i \in \mathcal{S}^{n-1} : s_{i,1} > \epsilon\}$ consists of at most $\mathrm{e}^{-\epsilon^2 n/2}$ proportion of the total area on the sphere. A similar argument carries over to $\boldsymbol{f}_i$, and thus, $\boldsymbol{f}_i$ is subgaussian as well.

Standard concentration inequalities on the sum of random positive semi-definite matrices composed of independent isotropic subgaussian rows [60, Remark 5.40] confirm that

$$\left\| \frac{1}{|\bar{\mathcal{I}}_0|} \left(\sqrt{n}\boldsymbol{F}\right)^{\mathcal{T}} \left(\sqrt{n}\boldsymbol{F}\right) - \boldsymbol{I}_{n-1} \right\| \leq \sigma \left\| \boldsymbol{I}_{n-1} \right\| \tag{72}$$

holds with probability at least $1 - 2\mathrm{e}^{-c_K n}$ as long as $|\bar{\mathcal{I}}_0|/n$ is sufficiently large, where σ is a numerical constant that can take arbitrarily small values and $c_K > 0$ is a universal constant. Without loss of generality, let us work with $\sigma := 0.01$ in (72), so for any unit vector $\boldsymbol{d} \in \mathbb{R}^{n-1}$, the following inequality holds with probability at least $1 - 2\mathrm{e}^{-c_K n}$,

$$\left| \frac{n}{|\bar{\mathcal{I}}_0|} \boldsymbol{d}^{\mathcal{T}} \boldsymbol{F}^{\mathcal{T}} \boldsymbol{F} \boldsymbol{d} - \boldsymbol{d}^{\mathcal{T}} \boldsymbol{d} \right| \leq 1.01 \boldsymbol{d}^{\mathcal{T}} \boldsymbol{d}, \tag{73}$$

or equivalently,

$$\|\boldsymbol{F}\boldsymbol{d}\|^2 = \left| \boldsymbol{d}^{\mathcal{T}} \boldsymbol{F}^{\mathcal{T}} \boldsymbol{F} \boldsymbol{d} \right| \leq 1.01 |\bar{\mathcal{I}}_0|/n. \tag{74}$$

Combining the last with (67), one readily concludes that

$$\left\| \bar{\boldsymbol{S}}_0 \boldsymbol{x}^{\perp} \right\|^2 \leq 1.01 |\bar{\mathcal{I}}_0|/n \tag{75}$$

holds with probability at least $1 - 2\mathrm{e}^{-c_K n}$, provided that $|\bar{\mathcal{I}}_0|/n$ exceeds some constant. Note that c_K depends on the maximum subgaussian norm of the rows of $\sqrt{n}\boldsymbol{F}$, and we assume without loss of generality $c_K \geq 1/2$. Hence, $\|\bar{\boldsymbol{S}}_0 \boldsymbol{u}\|^2$ in (27) is upper bounded simply by letting $\boldsymbol{u} = \boldsymbol{x}^{\perp}$ in (75).

C. Proof of Lemma 3

We next pursue a meaningful lower bound for $\|\bar{\boldsymbol{S}}_0 \boldsymbol{x}\|^2$ in (29). When $\boldsymbol{x} = \boldsymbol{e}_1$, one has $\|\bar{\boldsymbol{S}}_0 \boldsymbol{x}\|^2 = \|\bar{\boldsymbol{S}}_0 \boldsymbol{e}_1\|^2 = \sum_{i=1}^{|\bar{\mathcal{I}}_0|} s_{i,1}^2$. It is further worth mentioning that all squared entries of any spherical random vector $\boldsymbol{s}_i$ obey the *Beta* distribution with parameters $\alpha = \frac{1}{2}$, and $\beta = \frac{n-1}{2}$, i.e., $s_{i,j}^2 \sim \mathrm{Beta}\left(\frac{1}{2}, \frac{n-1}{2}\right)$, $\forall i, j$, [62, Lemma 2]. Although they have closed-form probability density functions (pdfs) that may facilitate deriving a wanted lower bound, we shall take another easier route detailed as follows. A simple yet useful inequality is established first.

Lemma 7. *Given m fractions obeying $1 > \frac{p_1}{q_1} \geq \frac{p_2}{q_2} \geq \cdots \geq \frac{p_m}{q_m} > 0$, in which p_i, $q_i > 0$, $\forall i \in [m]$, the following holds for all $1 \leq k \leq m$*

$$\sum_{i=1}^{k} \frac{p_i}{q_i} \geq \sum_{i=1}^{k} \frac{p_{[i]}}{q_{[1]}} \tag{76}$$

where $p_{[i]}$ denotes the i-th largest one among $\{p_i\}_{i=1}^m$, and hence, $q_{[1]}$ is the maximum in $\{q_i\}_{i=1}^m$.

Proof. For any $k \in [m]$, according to the definition of $q_{[i]}$, it holds that $p_{[1]} \geq p_{[2]} \geq \cdots \geq p_{[k]}$, so $\frac{p_{[1]}}{q_{[1]}} \geq \frac{p_{[2]}}{q_{[1]}} \geq \cdots \geq \frac{p_{[k]}}{q_{[1]}}$. Considering $q_{[1]} \geq q_i$, $\forall i \in [m]$, and letting $j_i \in [m]$ be the index such that $p_{j_i} = p_{[i]}$, then $\frac{p_{j_i}}{q_{j_i}} = \frac{p_{[i]}}{q_{j_i}} \geq \frac{p_{[i]}}{q_{[1]}}$ holds for any $i \in [k]$. Therefore, $\sum_{i=1}^k \frac{p_{j_i}}{q_{j_i}} = \sum_{i=1}^k \frac{p_{[i]}}{q_{j_i}} \geq \sum_{i=1}^k \frac{p_{[i]}}{q_{[1]}}$. Note that $\left\{\frac{p_{[i]}}{q_{j_i}}\right\}_{i=1}^k$ comprise a subset of terms in $\left\{\frac{p_i}{q_i}\right\}_{i=1}^m$. On the other hand, according to

our assumption, $\sum_{i=1}^{k}\frac{p_i}{q_i}$ is the largest among all sums of k summands; hence, $\sum_{i=1}^{k}\frac{p_i}{q_i}\geq\sum_{i=1}^{k}\frac{p_{[i]}}{q_{j_i}}$ yields $\sum_{i=1}^{k}\frac{p_i}{q_i}\geq\sum_{i=1}^{k}\frac{p_{[i]}}{q_{[1]}}$ concluding the proof. □

Without loss of generality and for simplicity of exposition, let us assume that indices of $\boldsymbol{a}_i$'s have been re-ordered such that

$$\frac{a_{1,1}^2}{\|\boldsymbol{a}_1\|^2}\geq\frac{a_{2,1}^2}{\|\boldsymbol{a}_2\|^2}\geq\cdots\geq\frac{a_{m,1}^2}{\|\boldsymbol{a}_m\|^2},\tag{77}$$

where $a_{i,1}$ denotes the first element of $\boldsymbol{a}_i$. Therefore, writing $\|\overline{\boldsymbol{S}}_0\boldsymbol{e}_1\|^2=\sum_{i=1}^{|\overline{\mathcal{I}}_0|}a_{i,1}^2/\|\boldsymbol{a}_i\|^2$, the next task amounts to finding the sum of the $|\overline{\mathcal{I}}_0|$ largest out of all m entities in (77). Applying the result (76) in Lemma 7 gives

$$\sum_{i=1}^{|\overline{\mathcal{I}}_0|}\frac{a_{i,1}^2}{\|\boldsymbol{a}_i\|^2}\geq\sum_{i=1}^{|\overline{\mathcal{I}}_0|}\frac{a_{[i],1}^2}{\max_{i\in[m]}\|\boldsymbol{a}_i\|^2},\tag{78}$$

in which $a_{[i],1}^2$ stands for the i-th largest entity in $\left\{a_{i,1}^2\right\}_{i=1}^m$.

Observe that for i.i.d. random vectors $\boldsymbol{a}_i\sim\mathcal{N}\big(\boldsymbol{0},\boldsymbol{I}_n\big)$, the property $\mathbb{P}(\|\boldsymbol{a}_i\|^2\geq 2.3n)\leq\mathrm{e}^{-n/2}$ holds for large enough n (e.g., $n\geq 20$), which can be understood upon substituting $\xi:=n/2$ into the following standard result [63, Lemma 1]

$$\mathbb{P}\left(\|\boldsymbol{a}_i\|^2-n\geq 2\sqrt{\xi}+2\xi\right)\leq\mathrm{e}^{-\xi}.\tag{79}$$

In addition, one readily concludes that $\mathbb{P}\left(\max_{i\in[m]}\|\boldsymbol{a}_i\|\leq\sqrt{2.3n}\right)\geq 1-m\mathrm{e}^{-n/2}$. We will henceforth build our subsequent proofs on this event without stating this explicitly each time encountering it. Therefore, (78) can be lower bounded by

$$\left\|\overline{\boldsymbol{S}}\boldsymbol{x}\right\|^2=\sum_{i=1}^{|\overline{\mathcal{I}}_0|}\frac{a_{i,1}^2}{\|\boldsymbol{a}_i\|^2}\geq\sum_{i=1}^{|\overline{\mathcal{I}}_0|}\frac{a_{[i],1}^2}{\max_{i\in[m]}\|\boldsymbol{a}_i\|^2}\geq\frac{1}{2.3n}\sum_{i=1}^{|\overline{\mathcal{I}}_0|}\left|a_{[i],1}\right|^2,\tag{80}$$

which holds with probability at least $1-m\mathrm{e}^{-n/2}$. The task left for bounding $\|\overline{\boldsymbol{S}}\boldsymbol{x}\|^2$ is to derive a meaningful lower bound for $\sum_{i=1}^{|\overline{\mathcal{I}}_0|}a_{[i],1}^2$. Roughly speaking, because the ratio $|\overline{\mathcal{I}}_0|/m$ is small, e.g., $|\overline{\mathcal{I}}_0|/m\leq 1/5$, a trivial result consists of bounding $(1/|\overline{\mathcal{I}}_0|)\sum_{i=1}^{|\overline{\mathcal{I}}_0|}a_{[i],1}^2$ by its sample average $(1/m)\sum_{i=1}^{m}a_{[i],1}^2$. The latter can be bounded using its ensemble mean, i.e., $\mathbb{E}[a_{i,1}^2]=1$, $\forall i\in[\overline{\mathcal{I}}_0]$, to yield $(1/m)\sum_{i=1}^{m}a_{[i],1}^2\geq(1-\epsilon)\mathbb{E}[a_{i,1}^2]=1-\epsilon$, which holds with high probability for some numerical constant $\epsilon>0$ [28, Lemma 3.1]. Therefore, one has a candidate lower bound $\sum_{i=1}^{|\overline{\mathcal{I}}_0|}a_{[i],1}^2\geq(1-\epsilon)|\overline{\mathcal{I}}_0|$. Nonetheless, this lower bound is in general too loose, and it contributes to a relatively large upper bound on the wanted term in (27).

To obtain an alternative bound, let us examine first the typical size of the maximum in $\left\{a_{i,1}^2\right\}_{i=1}^m$. Observe obviously that the modulus $|a_{i,1}|$ follows the half-normal distribution having the pdf $p(r)=\sqrt{2/\pi}\cdot\mathrm{e}^{-r^2/2}$, $r>0$, and it is easy to verify that

$$\mathbb{E}[|a_{i,1}|]=\sqrt{2/\pi}.\tag{81}$$

Then integrating the pdf from 0 to $+\infty$ yields the corresponding accumulative distribution function (cdf) expressible in terms of the error function $\mathbb{P}\left(|a_{i,1}|>\xi\right)=1-\mathrm{erf}\left(\xi/2\right)$, i.e., $\mathrm{erf}\left(\xi\right):=2/\sqrt{\pi}\cdot\int_0^{\xi}\mathrm{e}^{-r^2}\mathrm{d}r$. Appealing to a lower bound on the complimentary error function $\mathrm{erfc}\left(\xi\right):=1-\mathrm{erf}\left(\xi\right)$ from [64, Theorem

2], one establishes that $\mathbb{P}\left(|a_{i,1}|>\xi\right)=1-\mathrm{erf}\left(\xi/2\right)\geq(3/5)\mathrm{e}^{-\xi^2/2}$. Additionally, direct application of probability theory and Taylor expansion confirms that

$$\begin{aligned}\mathbb{P}\big(\max_{i\in[m]}|a_{i,1}|\geq\xi\big)&=1-\left[\mathbb{P}\left(|a_{i,1}|\leq\xi\right)\right]^m\\&\geq 1-\left(1-0.6\mathrm{e}^{-\xi^2/2}\right)^m\\&\geq 1-\mathrm{e}^{-0.6m\mathrm{e}^{-\xi^2/2}}.\end{aligned}\tag{82}$$

Choosing now $\xi:=\sqrt{2\log n}$ leads to

$$\mathbb{P}\big(\max_{i\in[m]}|a_{i,1}|\geq\sqrt{2\log n}\big)\geq 1-\mathrm{e}^{-0.6m/n}\geq 1-o(1)\tag{83}$$

which holds with the proviso that m/n is large enough, and the symbol $o(1)$ represents a small constant probability. Thus, provided that m/n exceeds some large constant, the event $\max_{i\in[m]}a_{i,1}^2\geq 2\log n$ occurs with high probability. Hence, one may expect a tighter lower bound than $(1-\epsilon_0)|\overline{\mathcal{I}}_0|$, which is on the same order of m under the assumption that $|\overline{\mathcal{I}}_0|/m$ is about a constant.

Although $a_{i,1}^2$ obeys the *Chi-square* distribution with $k=1$ degrees of freedom, its cdf is rather complicated and does not admit a nice closed-form expression. A small trick is hence taken in the sequel. Postulate without loss of generality that both m and $|\overline{\mathcal{I}}_0|$ are even. Grouping two consecutive $a_{[i],1}^2$'s together, introduce a new variable $\vartheta[i]:=a_{[2k-1],1}^2+a_{[2k],1}^2$, $\forall k\in[m/2]$, hence yielding a sequence of ordered numbers, i.e., $\vartheta_{[1]}\geq\vartheta_{[2]}\geq\cdots\geq\vartheta_{[m/2]}>0$. Then, one can equivalently write the wanted sum as

$$\sum_{i=1}^{|\overline{\mathcal{I}}_0|}a_{[i],1}^2=\sum_{i=1}^{|\overline{\mathcal{I}}_0|/2}\vartheta_{[i]}.\tag{84}$$

On the other hand, for i.i.d. standard normal random variables $\{a_{i,1}\}_{i=1}^m$, let us consider grouping randomly two of them and denote the corresponding sum of their squares by $\chi_k:=a_{k_i,1}^2+a_{k_j,1}^2$, where $k_i\neq k_j\in[m]$, and $k\in[m/2]$. It is self-evident that the χ_k's are identically distributed obeying the *Chi-square* distribution with $k=2$ degrees of freedom, having the pdf

$$p\left(r\right)=\frac{1}{2}\mathrm{e}^{-\frac{r}{2}},\quad r\geq 0,\tag{85}$$

and the following complementary cdf (ccdf)

$$\mathbb{P}(\chi_k\geq\xi):=\int_{\xi}^{\infty}\frac{1}{2}\mathrm{e}^{-\frac{r}{2}}\mathrm{d}r=\mathrm{e}^{-\frac{\xi}{2}},\quad\forall\xi\geq 0.\tag{86}$$

Ordering all χ_k's, summing the $|\overline{\mathcal{I}}_0|/2$ largest ones, and comparing the resultant sum with the one in (84) confirm that

$$\sum_{i=1}^{|\overline{\mathcal{I}}_0|/2}\chi_{[i]}\leq\sum_{i=1}^{|\overline{\mathcal{I}}_0|/2}\vartheta_{[i]}=\sum_{i=1}^{|\overline{\mathcal{I}}_0|}a_{[i],1}^2,\quad\forall|\overline{\mathcal{I}}_0|\in[m].\tag{87}$$

Upon setting $\mathbb{P}(\chi_k\geq\xi)=|\overline{\mathcal{I}}_0|/m$, one obtains an estimate of $\chi_{|\overline{\mathcal{I}}_0|/2}$, the $(|\overline{\mathcal{I}}_0|/2)$-th largest value in $\{\chi_k\}_{k=1}^{m/2}$ as follows

$$\hat{\chi}_{|\overline{\mathcal{I}}_0|/2}:=2\log\left(m/|\overline{\mathcal{I}}_0|\right).\tag{88}$$

Furthermore, applying the Hoeffding-type inequality [60, Proposition 5.10] and leveraging the convexity of the ccdf in (86), one readily establishes that

$$\mathbb{P}\Big(\hat{\chi}_{|\bar{\mathcal{I}}_0|/2} - \chi_{|\bar{\mathcal{I}}_0|/2} > \xi\Big) \leq \mathrm{e}^{-\frac{1}{4}m\xi^2\mathrm{e}^{-\xi}(|\bar{\mathcal{I}}_0|/m)^2}, \quad \forall \xi > 0. \tag{89}$$

Taking without loss of generality $\xi := 0.05\hat{\chi}_{|\bar{\mathcal{I}}_0|/2} = 0.1\log\big(m/|\bar{\mathcal{I}}_0|\big)$ gives

$$\mathbb{P}\Big(\chi_{|\bar{\mathcal{I}}_0|/2} < 0.95\hat{\chi}_{|\bar{\mathcal{I}}_0|/2}\Big) \leq \mathrm{e}^{-c_0 m} \tag{90}$$

for some universal constants c_0, $c_\chi > 0$, and sufficiently large n such that $|\bar{\mathcal{I}}_0|/m \gtrsim c_\chi > 0$. The remaining part in this section assumes that this event occurs.

Choosing $\xi := 4\log n$ and substituting this into the ccdf in (86) leads to

$$\mathbb{P}\left(\chi \leq 4\log n\right) = 1 - 1/n^2. \tag{91}$$

Notice that each summand in $\sum_{i=1}^{|\bar{\mathcal{I}}_0|/2}\chi_{[i]} \geq \sum_{i=1}^{m/2}\chi_i\mathbb{1}_{\tilde{\mathcal{E}}_i}$ is Chi-square distributed, and hence could be unbounded, so we choose to work with the truncation $\sum_{i=1}^{m/2}\chi_i\mathbb{1}_{\tilde{\mathcal{E}}_i}$, where the $\mathbb{1}_{\tilde{\mathcal{E}}_i}$'s are independent copies of $\mathbb{1}_{\tilde{\mathcal{E}}}$, and $\mathbb{1}_{\tilde{\mathcal{E}}}$ denotes the indicator function for the ensuing events

$$\tilde{\mathcal{E}} := \Big\{\chi \geq \hat{\chi}_{|\bar{\mathcal{I}}_0|/2}\Big\} \cap \left\{\chi \leq 4\log n\right\}. \tag{92}$$

Apparently, it holds that $\sum_{i=1}^{|\bar{\mathcal{I}}_0|/2}\chi_{[i]} \geq \sum_{i=1}^{m/2}\chi_i\mathbb{1}_{\tilde{\mathcal{E}}_i}$. One further establishes that

$$\begin{aligned}
\mathbb{E}\left[\chi_i\mathbb{1}_{\tilde{\mathcal{E}}_i}\right] &:= \int_{\hat{\chi}_{|\bar{\mathcal{I}}_0|/2}}^{4\log n}\frac{1}{2}r\mathrm{e}^{-r/2}\mathrm{d}r \\
&= \Big(\hat{\chi}_{|\bar{\mathcal{I}}_0|/2} + 2\Big)\,\mathrm{e}^{-\hat{\chi}_{|\bar{\mathcal{I}}_0|/2}/2} - (4\log n + 2)\,\mathrm{e}^{-2\log n} \\
&= \frac{2|\bar{\mathcal{I}}_0|}{m}\Big[1 + \log\big(m/|\bar{\mathcal{I}}_0|\big)\Big] - \frac{(4\log n + 2)}{n^2}.
\end{aligned} \tag{93}$$

The task of bounding $\sum_{i=1}^{|\bar{\mathcal{I}}_0|}a_{[i],1}^2$ in (87) now boils down to bounding $\sum_{i=1}^{m/2}\chi_i\mathbb{1}_{\tilde{\mathcal{E}}_i}$ from its expectation in (93). A convenient way to accomplish this is using the Bernstein inequality [60, Proposition 5.16], that deals with bounded random variables. That also justifies the reason of introducing the upper-bound truncation on χ in (92). Specifically, let us define

$$\vartheta_i := \chi_i\mathbb{1}_{\tilde{\mathcal{E}}_i} - \mathbb{E}\left[\chi_i\mathbb{1}_{\tilde{\mathcal{E}}_i}\right], \quad \forall i \in [m/2]. \tag{94}$$

Thus, $\{\vartheta_i\}_{i=1}^{m/2}$ are i.i.d. centered and bounded random variables following from the mean-subtraction and the upper-bound truncation. Further, according to the ccdf (86) and the definition of sub-exponential random variables [60, Definition 5.13], the terms $\{\vartheta_i\}_{i=1}^{m/2}$ are sub-exponential. Then, the following

$$\left|\sum_{i=1}^{m/2}\vartheta_i\right| \geq \tau \tag{95}$$

holds with probability at least $1-2\mathrm{e}^{-c_s\min\left(\tau/K_s,\tau^2/K_s^2\right)}$, in which $c_s>0$ is a universal constant, and $K_s:=\max_{i\in[m/2]}\|\vartheta_i\|_{\psi_1}$ represents the maximum subexponential norm of the ϑ_i's. Indeed, K_s can be found as follows [60, Definition 5.13]

$$\begin{aligned}K_s&:=\sup_{p\geq 1}p^{-1}\left(\mathbb{E}\left[|\vartheta_i|^p\right]\right)^{1/p}\\&\leq\left(4\log n-2\log\left(m/|\overline{\mathcal{I}}_0|\right)\right)\left[|\overline{\mathcal{I}}_0|/m-1/n^2\right]\\&\leq\frac{2|\overline{\mathcal{I}}_0|}{m}\log\left(n^2|\overline{\mathcal{I}}_0|/m\right)\\&\leq\frac{4|\overline{\mathcal{I}}_0|}{m}\log n.\end{aligned}\tag{96}$$

Choosing $\tau:=8|\overline{\mathcal{I}}_0|/(c_sm)\cdot\log^2 n$ in (95) yields

$$\begin{aligned}\sum_{i=1}^{m/2}\chi_i\mathbb{1}_{\tilde{\mathcal{E}}_i}&\geq|\overline{\mathcal{I}}_0|\left[1+\log\left(m/|\overline{\mathcal{I}}_0|\right)\right]-8|\overline{\mathcal{I}}_0|/(c_sm)\cdot\log^2 n-m\left(2\log n+1\right)/n^2\\&\geq(1-\epsilon_s)|\overline{\mathcal{I}}_0|\left[1+\log\left(m/|\overline{\mathcal{I}}_0|\right)\right]\end{aligned}\tag{97}$$

for some small constant $\epsilon_s>0$, which holds with probability at least $1-m\mathrm{e}^{-n/2}-\mathrm{e}^{-c_0m}-3/n^2$ as long as m/n exceeds some numerical constant and n is sufficiently large. Therefore, combining (80), (87), and (97), one concludes that the following holds with high probability

$$\left\|\overline{\boldsymbol{S}}_0\boldsymbol{x}\right\|^2=\sum_{i=1}^{|\overline{\mathcal{I}}_0|}\frac{a_{i,1}^2}{\|\boldsymbol{a}_i\|^2}\geq(1-\epsilon_s)\frac{|\overline{\mathcal{I}}_0|}{2.3n}\left[1+\log\left(m/|\overline{\mathcal{I}}_0|\right)\right].\tag{98}$$

Taking $\epsilon_s:=0.01$ without loss of generality concludes the proof of Lemma 3.

D. Proof of Lemma 5

Let us first prove the argument for a fixed pair $\boldsymbol{h}$ and $\boldsymbol{x}$, so $\boldsymbol{h}$ and $\boldsymbol{z}$ are independent of $\{\boldsymbol{a}_i\}_{i=1}^m$, and then apply a covering argument. To start, introduce a Lipschitz-continuous counterpart for the discontinuous indicator function [6, A.2]

$$\chi_E(\theta):=\begin{cases}1, & |\theta|\geq\frac{\sqrt{1.01}}{1+\gamma},\\ 100(1+\gamma)^2\theta^2-100, & \frac{1}{1+\gamma}\leq|\theta|<\frac{\sqrt{1.01}}{1+\gamma},\\ 0, & |\theta|<\frac{1}{1+\gamma}\end{cases}\tag{99}$$

with Lipschitz constant $\mathcal{O}(1)$. Recall that $\mathcal{E}_i:=\left\{\left|\frac{\boldsymbol{a}_i^\mathcal{T}\boldsymbol{z}}{\boldsymbol{a}_i^\mathcal{T}\boldsymbol{x}}\right|\geq\frac{1}{1+\gamma}\right\}$, so it holds that $0\leq\chi_E\left(\left|\frac{\boldsymbol{a}_i^\mathcal{T}\boldsymbol{z}}{\boldsymbol{a}_i^\mathcal{T}\boldsymbol{x}}\right|\right)\leq\mathbb{1}_{\mathcal{E}_i}$ for any $\boldsymbol{x}\in\mathbb{R}^n$ and $\boldsymbol{h}\in\mathbb{R}^n$, thus yielding

$$\frac{1}{m}\sum_{i=1}^m\left(\boldsymbol{a}_i^\mathcal{T}\boldsymbol{h}\right)^2\mathbb{1}_{\mathcal{E}_i}\geq\frac{1}{m}\sum_{i=1}^m\left(\boldsymbol{a}_i^\mathcal{T}\boldsymbol{h}\right)^2\chi_E\left(\left|\frac{\boldsymbol{a}_i^\mathcal{T}\boldsymbol{z}}{\boldsymbol{a}_i^\mathcal{T}\boldsymbol{x}}\right|\right)=\frac{1}{m}\sum_{i=1}^m\left(\boldsymbol{a}_i^\mathcal{T}\boldsymbol{h}\right)^2\chi_E\left(\left|1+\frac{\boldsymbol{a}_i^\mathcal{T}\boldsymbol{h}}{\boldsymbol{a}_i^\mathcal{T}\boldsymbol{x}}\right|\right).\tag{100}$$

By homogeneity and rotational invariance property of normal distributions, it suffices to prove the case where $\boldsymbol{x}=\boldsymbol{e}_1$ and $\|\boldsymbol{h}\|/\|\boldsymbol{x}\|=\|\boldsymbol{h}\|\leq\rho$. According to (100), lower bounding the first term in (51) can be achieved by lower bounding $\sum_{i=1}^m(\boldsymbol{a}_i^\mathcal{T}\boldsymbol{h})^2\chi_E\left(\left|1+\frac{\boldsymbol{a}_i^\mathcal{T}\boldsymbol{h}}{\boldsymbol{a}_i^\mathcal{T}\boldsymbol{x}}\right|\right)$ instead. To that end, let us find the

mean of $\left(\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}\right)^2\chi_E\left(\left|1+\frac{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}}{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}}\right|\right)$. Note that $\left(\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}\right)^2$ and $\chi_E\left(\left|1+\frac{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}}{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}}\right|\right)$ are dependent. Introduce an orthonormal matrix $\boldsymbol{U}_{\boldsymbol{h}}$ that contains $\boldsymbol{h}^{\mathcal{T}}/\|\boldsymbol{h}\|$ as its first row, i.e.,

$$\boldsymbol{U}_{\boldsymbol{h}} := \left[\begin{array}{c} \boldsymbol{h}^{\mathcal{T}}/\|\boldsymbol{h}\| \\ \widetilde{\boldsymbol{U}}_{\boldsymbol{h}} \end{array}\right] \tag{101}$$

for some orthogonal matrix $\widetilde{\boldsymbol{U}}_{\boldsymbol{h}} \in \mathbb{R}^{(n-1)\times n}$ such that $\boldsymbol{U}_{\boldsymbol{h}}$ is orthonormal. Moreover, define $\tilde{\boldsymbol{h}} := \boldsymbol{U}_{\boldsymbol{h}}\boldsymbol{h}$, and $\tilde{\boldsymbol{a}}_i := \boldsymbol{U}_{\boldsymbol{h}}\boldsymbol{a}_i$; and let $\tilde{a}_{i,1}$ and $\tilde{\boldsymbol{a}}_{i,\setminus 1}$ denote the first entry and the remaining entries in vector $\tilde{\boldsymbol{a}}_i$; likewise for vector $\tilde{\boldsymbol{h}}$. Then, for any $\boldsymbol{h}$ such that $\|\boldsymbol{h}\| \le \rho$, the next holds

$$\begin{aligned}
\mathbb{E}\left[\left(\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}\right)^2\chi_E\left(\left|1+\frac{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}}{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}}\right|\right)\right] &= \mathbb{E}\left[\left(\tilde{a}_{i,1}\tilde{h}_1\right)^2\chi_E\left(\left|1+\frac{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}}{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}}\right|\right)\right] + \mathbb{E}\left[\left(\tilde{\boldsymbol{a}}_{i,\setminus 1}^{\mathcal{T}}\tilde{\boldsymbol{h}}_{\setminus 1}\right)^2\chi_E\left(\left|1+\frac{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}}{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}}\right|\right)\right] \\
&= \tilde{h}_1^2\mathbb{E}\left[\tilde{a}_{i,1}^2\chi_E\left(\left|1+\frac{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}}{a_{i,1}}\right|\right)\right] + \mathbb{E}\left[\left(\tilde{\boldsymbol{a}}_{i,\setminus 1}^{\mathcal{T}}\tilde{\boldsymbol{h}}_{\setminus 1}\right)^2\right]\mathbb{E}\left[\chi_E\left(\left|1+\frac{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}}{a_{i,1}}\right|\right)\right] \\
&= \tilde{h}_1^2\mathbb{E}\left[\tilde{a}_{i,1}^2\chi_E\left(\left|1+\frac{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}}{a_{i,1}}\right|\right)\right] + \left\|\tilde{\boldsymbol{h}}_{\setminus 1}\right\|^2\mathbb{E}\left[\chi_E\left(\left|1+\frac{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}}{a_{i,1}}\right|\right)\right] \\
&\ge \left(\tilde{h}_1^2 + \left\|\tilde{\boldsymbol{h}}_{\setminus 1}\right\|^2\right)\min\left\{\mathbb{E}\left[a_{i,1}^2\chi_E\left(\left|1+h_1+\frac{\boldsymbol{a}_{i,\setminus 1}^{\mathcal{T}}\boldsymbol{h}_{\setminus 1}}{a_{i,1}}\right|\right)\right],\right. \\
&\qquad \left.\mathbb{E}\left[\chi_E\left(\left|1+h_1+\frac{\boldsymbol{a}_{i,\setminus 1}^{\mathcal{T}}\boldsymbol{h}_{\setminus 1}}{a_{i,1}}\right|\right)\right]\right\} \\
&\ge \|\boldsymbol{h}\|^2\min\left\{\mathbb{E}\left[a_{i,1}^2\chi_E\left(\left|1-\rho+\frac{a_{i,2}}{a_{i,1}}\rho\right|\right)\right], \mathbb{E}\left[\chi_E\left(1-\rho+\frac{a_{i,2}}{a_{i,1}}\rho\right)\right]\right\} \\
&= (1-\zeta_1)\|\boldsymbol{h}\|^2
\end{aligned} \tag{102}$$

where the second equality follows from the independence between $\tilde{\boldsymbol{a}}_{i,\setminus 1}^{\mathcal{T}}\tilde{\boldsymbol{h}}_{\setminus 1}$ and $\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}$, the second inequality holds for $\rho \le 1/10$ and $\gamma > 1/2$, and the last equality comes from the definition of ζ_1 in (94). Notice that $\varrho := (\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h})^2\chi_E\left(\left|1+\frac{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}}{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}}\right|\right) \le (\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h})^2 \overset{d}{=} \|\boldsymbol{h}\|^2 a_{i,1}^2$ is a subexponential variable, and thus its subexponential norm $\|\varrho\|_{\psi_1} := \sup_{p\ge 1}\left[\mathbb{E}(|\varrho|^p)\right]^{1/p}$ is finite.

Direct application of the Berstein-type inequality [60, Proposition 5.16] confirms that for any $\epsilon > 0$, the following

$$\begin{aligned}
\frac{1}{m}\sum_{i=1}^{m}\left(\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}\right)^2\chi_E\left(\left|1+\frac{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}}{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}}\right|\right) &\ge \mathbb{E}\left[\left(\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}\right)^2\chi_E\left(\left|1+\frac{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}}{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}}\right|\right)\right] - \epsilon\|\boldsymbol{h}\|^2 \\
&\ge (1-\zeta_1-\epsilon)\|\boldsymbol{h}\|^2
\end{aligned} \tag{103}$$

holds with probability at least $1-\mathrm{e}^{-c_5 m\epsilon^2}$ for some numerical constant $c_5 > 0$ provided that $\epsilon \le \|\varrho\|_{\psi_1}$ by assumption.

To obtain uniform control over all vectors $\boldsymbol{z}$ and $\boldsymbol{x}$ such that $\|\boldsymbol{z}-\boldsymbol{x}\| \le \rho$, the net covering argument is applied [60, Definition 5.1]. Let $\mathcal{S}_\epsilon$ be an ϵ-net of the unit sphere, $\mathcal{L}_\epsilon$ be an ϵ-net of $[0,\,\rho]$, and define

$$\mathcal{N}_\epsilon := \left\{(\boldsymbol{z},\,\boldsymbol{h},\,t) : (\boldsymbol{z}_0,\,\boldsymbol{h}_0,\,t_0) \in \mathcal{S}_\epsilon \times \mathcal{S}_\epsilon \times \mathcal{L}_\epsilon\right\}. \tag{104}$$

Since the cardinality $|\mathcal{S}_\epsilon| \le (1+2/\epsilon)^n$ [60, Lemma 5.2], then

$$|\mathcal{N}_\epsilon| \le (1+2/\epsilon)^{2n}\,\rho/\epsilon \le (1+2/\epsilon)^{2n+1} \tag{105}$$

due to the fact that $\rho/\epsilon < 2/\epsilon < 1 + 2/\epsilon$ for $0 \le \rho < 1$.

Consider now any $(\boldsymbol{z},\, \boldsymbol{h},\, t)$ obeying $\|\boldsymbol{h}\| = t \le \rho$. There exists a pair $(\boldsymbol{z}_0,\, \boldsymbol{h}_0,\, t_0) \in \mathcal{N}_\epsilon$ such that $\|\boldsymbol{z} - \boldsymbol{z}_0\|$, $\|\boldsymbol{h} - \boldsymbol{h}_0\|$, and $|t - t_0|$ are each at most ϵ. Taking the union bound yields

$$\begin{aligned}\frac{1}{m}\sum_{i=1}^{m}\left(\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}_0\right)^2\chi_E\left(\left|1+\frac{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}_0}{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}}\right|\right) &\ge \frac{1}{m}\sum_{i=1}^{m}\left(\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}_0\right)^2\chi_E\left(\left|1-t_0+\frac{a_{i,2}}{a_{i,1}}t_0\right|\right)\\ &\ge (1-\zeta_1-\epsilon)\,\|\boldsymbol{h}_0\|^2, \quad \forall\,(\boldsymbol{z}_0,\, \boldsymbol{h}_0,\, t_0) \in \mathcal{N}_\epsilon \end{aligned} \tag{106}$$

with probability at least $1-(1+2/\epsilon)^{2n+1}\,\mathrm{e}^{-c_5\epsilon^2 m} \ge 1-\mathrm{e}^{-c_0 m}$, which follows by choosing m such that $m \ge \left(c_6\cdot\epsilon^{-2}\log\epsilon^{-1}\right)n$ for some constant $c_6 > 0$.

Recall that $\chi_E(\tau)$ is Lipschitz continuous, thus

$$\begin{aligned}&\left|\frac{1}{m}\sum_{i=1}^{m}\left\{\left(\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}\right)^2\chi_E\left(\left|1+\frac{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}}{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}}\right|\right)-\left(\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}_0\right)^2\chi_E\left(\left|1+\frac{\boldsymbol{a}^{\mathcal{T}}\boldsymbol{h}_0}{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}}\right|\right)\right\}\right|\\ &\quad\lesssim \frac{1}{m}\sum_{i=1}^{m}\left|\left(\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}\right)^2-\left(\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}_0\right)^2\right|\\ &\quad= \frac{1}{m}\sum_{i=1}^{m}\left|\boldsymbol{a}_i^{\mathcal{T}}\left(\boldsymbol{h}\boldsymbol{h}^{\mathcal{T}}-\boldsymbol{h}_0\boldsymbol{h}_0^{\mathcal{T}}\right)\boldsymbol{a}_i\right|\\ &\quad\lesssim c_7\sum_{i=1}^{m}\left|\boldsymbol{h}\boldsymbol{h}^{\mathcal{T}}-\boldsymbol{h}_0\boldsymbol{h}_0^{\mathcal{T}}\right|\\ &\quad\le 2.5c_7\,\|\boldsymbol{h}-\boldsymbol{h}_0\|\,\|\boldsymbol{h}\|\\ &\quad\le 2.5c_7\rho\epsilon \end{aligned} \tag{107}$$

for some numerical constant c_7 and provided that $\epsilon < 1/2$ and $m \ge \left(c_6\cdot\epsilon^{-2}\log\epsilon^{-1}\right)n$, where the first inequality arises from the Lipschitz property of $\chi_E(\tau)$, the second uses the results in Lemma 1 in [6], and the third from Lemma 2 in [6].

Putting all results together confirms that with probability exceeding $1-2\mathrm{e}^{-c_0 m}$, the following holds

$$\frac{1}{m}\sum_{i=1}^{m}\left(\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}\right)^2\chi_E\left(\left|1+\frac{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}}{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}}\right|\right) \ge \left[1-\zeta_1-(1+2.5c_7\rho)\,\epsilon\right]\|\boldsymbol{h}\|^2 \tag{108}$$

for all vectors $\|\boldsymbol{h}\| \,/\, \|\boldsymbol{x}\| \le \rho$, concluding the proof.

E. Proof of Lemma 6

Similar to the proof in Section D, it is convenient to work with the following auxiliary function instead of the discontinuous indicator function

$$\chi_D(\theta) := \begin{cases} 1, & |\theta| \ge \frac{2+\gamma}{1+\gamma} \\ -100\left(\frac{1+\gamma}{2+\gamma}\right)^2\theta^2+100, & \sqrt{0.99}\cdot\frac{2+\gamma}{1+\gamma} \le |\theta| < \frac{2+\gamma}{1+\gamma} \\ 0, & |\theta| < \sqrt{0.99}\cdot\frac{2+\gamma}{1+\gamma} \end{cases} \tag{109}$$

which is Lipschitz continuous in θ with Lipschitz constant $\mathcal{O}(1)$. For $\mathcal{D}_i := \left\{\left|\frac{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}}{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}}\right| \geq \frac{2+\gamma}{1+\gamma}\right\}$, it holds that $0 \leq \mathbb{1}_{\mathcal{D}_i} \leq \chi_D\left(\left|\frac{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}}{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}}\right|\right)$ for any $\boldsymbol{x} \in \mathbb{R}^n$ and $\boldsymbol{h} \in \mathbb{R}^n$. Assume without loss of generality $\boldsymbol{x} = \boldsymbol{e}_1$. Then for $1/2 \leq \gamma \leq 4$ and $\rho \leq 1/10$, it holds that

$$\begin{aligned}
\frac{1}{m}\sum_{i=1}^m \mathbb{1}_{\left\{\frac{|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}|}{|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}|} \geq \frac{2+\gamma}{1+\gamma}\right\}} &\leq \frac{1}{m}\sum_{i=1}^m \chi_D\left(\left|\frac{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{z}}{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}}\right|\right) = \frac{1}{m}\sum_{i=1}^m \chi_D\left(\left|1+\frac{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}}{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}}\right|\right) = \frac{1}{m}\sum_{i=1}^m \chi_D\left(\left|1+\frac{\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}}{a_{i,1}}\right|\right) \\
&= \frac{1}{m}\sum_{i=1}^m \chi_D\left(\left|1+h_1+\frac{a_{i,2}}{a_{i,1}}\left\|\boldsymbol{h}_{\setminus 1}\right\|\right|\right) \\
&\overset{\text{(i)}}{\leq} \frac{1}{m}\sum_{i=1}^m \mathbb{1}_{\left\{\left|1+h_1+\frac{a_{i,2}}{a_{i,1}}\|\boldsymbol{h}_{\setminus 1}\|\right| \geq \sqrt{0.99}\cdot\frac{2+\gamma}{1+\gamma}\right\}} \\
&\overset{\text{(ii)}}{\leq} \frac{1}{m}\sum_{i=1}^m \mathbb{1}_{\left\{\left|\frac{1+\rho}{\rho}+\frac{a_{i,2}}{a_{i,1}}\right| \geq \sqrt{0.99}\cdot\frac{2+\gamma}{\rho(1+\gamma)}\right\}}
\end{aligned} \tag{110}$$

where (i) arises from the definition of χ_D, and (ii) follows upon noticing that $a_{i,2}/a_{i,1}$ obeys the standard Cauchy distribution, i.e., $a_{i,2}/a_{i,1} \sim \mathrm{Cauchy}(0,1)$ [65], and particularly, transformation properties of Cauchy distributions assert that $(1+\rho)/\rho + a_{i,2}/a_{i,1} \sim \mathrm{Cauchy}((1+\rho)/\rho, 1)$ [66]. Recall that the cdf of a Cauchy distributed random variable $z \sim \mathrm{Cauchy}\left(\mu_0, \alpha\right)$ is given by [65]

$$F(z;\mu_0,\alpha) = \frac{1}{\pi}\arctan\left(\frac{z-\mu_0}{\alpha}\right) + \frac{1}{2}. \tag{111}$$

Define for notational brevity $z := a_{i,2}/a_{i,1}$, $\mu_0 := (1+\rho)/\rho$, and $z_0 := \sqrt{0.99}\cdot\frac{2+\gamma}{\rho(1+\gamma)}$ to yield

$$\begin{aligned}
\mathbb{E}\left[\mathbb{1}_{\{|\mu_0+z|\geq z_0\}}\right] &= 1 - \left[F\left(z_0;\mu_0,1\right) - F\left(-z_0;\mu_0,1\right)\right] \\
&= \frac{1}{\pi}\arctan\left(\frac{2\sqrt{0.99}\rho(2+\gamma)/(1+\gamma)}{0.99(2+\gamma)^2/(1+\gamma)^2 - (1+2\rho+2\rho^2)}\right) \\
&\overset{\text{(i)}}{\leq} \frac{1}{\pi}\cdot\frac{2\sqrt{0.99}(2+\gamma)/(1+\gamma)}{0.99(2+\gamma)^2/(1+\gamma)^2 - (1+2\rho+2\rho^2)}\rho \\
&\overset{\text{(ii)}}{\leq} \frac{\sqrt{0.99}(1+\gamma)}{0.42\pi}\rho
\end{aligned} \tag{112}$$

provided that γ and ρ are chosen such that $0.99(2+\gamma)^2/(1+\gamma)^2 > (1+2\rho+2\rho^2)$, which holds true for $1/2 \leq \gamma \leq 4$ and $\rho \leq 1/10$. In deriving (i), the inequality $\arctan(z) \leq z$ for any $z > 0$ is employed. The bound in (ii) is rather loose, yet it suffices for our purpose. Note that $\frac{\sqrt{0.99}(1+\gamma)}{0.42\pi}$ is $\mathcal{O}(1)$, so the probability in (112) is on the order of ρ, which can be made arbitrarily small as demonstrated by our analysis in Section V-A. Specifically, when taking values $\gamma = 0.7$ and $\rho = 1/10$, it holds $\mathbb{1}_{\{|\mu_0+z|\geq z_0\}} \leq 0.13$. Apparently, $\mathbb{1}_{\{|\mu_0+z|\geq z_0\}}$ is bounded; and it is known that all bounded random variables are subexponential. Thus, upon applying the Bernstein-type inequality [60, Corollary 5.17], the next holds with probability at least $1 - \mathrm{e}^{-c_5 m\epsilon^2}$ for some numerical constant $c_5 > 0$ and any sufficiently small $\epsilon > 0$

$$\begin{aligned}
\frac{1}{m}\sum_{i=1}^m \mathbb{1}_{\left\{\frac{|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}|}{|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}|} \geq \frac{2+\gamma}{1+\gamma}\right\}} &\leq \frac{1}{m}\sum_{i=1}^m \mathbb{1}_{\left\{\left|\frac{1+\rho}{\rho}+\frac{a_{i,2}}{a_{i,1}}\right| \geq \sqrt{0.99}\cdot\frac{2+\gamma}{\rho(1+\gamma)}\right\}} \leq \mathbb{E}\left[\mathbb{1}_{\left\{\left|\frac{1+\rho}{\rho}+\frac{a_{i,2}}{a_{i,1}}\right| \geq \sqrt{0.99}\cdot\frac{2+\gamma}{\rho(1+\gamma)}\right\}}\right] + \epsilon \\
&\leq (1+\epsilon)\frac{\sqrt{0.99}(1+\gamma)}{0.42\pi}\rho.
\end{aligned} \tag{113}$$

On the other hand, one can easily establish that the following holds true for all $\boldsymbol{h}$

$$\mathbb{E}\left[\left(\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}\right)^4\right] = \mathbb{E}\left[a_{i,1}^4\right]\|\boldsymbol{h}\|^4 = 3\|\boldsymbol{h}\|^4 \tag{114}$$

which has also been established in Lemma 1 [6] and Lemma 6.1 [35]. Further recalling our working assumption $\|\boldsymbol{a}_i\| \le 2.3n$, then random variables $\left(\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}\right)^4$ are bounded, and thus they are subexponential [60]. Appealing again to the Bernstein-type inequality for subexponential random variables and provided that $m/n > c_6 \cdot \epsilon^{-2}\log\epsilon^{-1}$ for some numerical constant $c_6 > 0$, then

$$\frac{1}{m}\sum_{i=1}^{m}\left(\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}\right)^4 \le 3(1+\epsilon)\|\boldsymbol{h}\|^4 \tag{115}$$

which holds with probability exceeding $1-\mathrm{e}^{-c_5 m\epsilon^2}$ for some universal constant $c_5 > 0$ and any sufficiently small $\epsilon > 0$.

Collecting together results in (113) and (115) and leveraging the Cauchy-Schwartz inequality, one establishes that for any $\rho \le 1/10$ and $1/2 \le \gamma \le 4$, the following

$$\begin{aligned}
\frac{1}{m}\sum_{i=1}^{m}\left(\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}\right)^2\mathbb{1}_{\mathcal{D}_i} &\le \sqrt{\frac{1}{m}\sum_{i=1}^{m}\left(\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}\right)^4}\sqrt{\frac{1}{m}\sum_{i=1}^{m}\mathbb{1}_{\left\{\frac{|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{h}|}{|\boldsymbol{a}_i^{\mathcal{T}}\boldsymbol{x}|}\ge\frac{2+\gamma}{1+\gamma}\right\}}} \\
&\le \sqrt{3(1+\epsilon)\|\boldsymbol{h}\|^4}\sqrt{(1+\epsilon)\frac{\sqrt{0.99}(1+\gamma)}{0.42\pi}\rho} \\
&= 1.5(1+\epsilon)\sqrt{1+\gamma}\sqrt{\rho}\|\boldsymbol{h}\|^2 \\
&\le (\zeta_2'+\epsilon)\|\boldsymbol{h}\|^2
\end{aligned} \tag{116}$$

with $\zeta_2' := 1.5\sqrt{(1+\gamma)\rho}$, which holds with probability at least $1-2\mathrm{e}^{-c_0 m}$. The latter arises if choosing $c_0 \le c_5\epsilon^2$ in $1-2\mathrm{e}^{-c_5 m\epsilon^2}$, which can be accomplished by taking m/n sufficiently large.

Acknowledgments

The authors would like to thank Mahdi Soltanolkotabi, Yuxin Chen, Kejun Huang, and Ju Sun for helpful discussions.

REFERENCES

[1] R. Balan, P. Casazza, and D. Edidin, "On signal reconstruction without phase," *Appl. Comput. Harmon. Anal.*, vol. 20, no. 3, pp. 345–356, May 2006.
[2] A. Conca, D. Edidin, M. Hering, and C. Vinzant, "An algebraic characterization of injectivity in phase retrieval," *Appl. Comput. Harmon. Anal.*, vol. 38, no. 2, pp. 346–356, Mar. 2015.
[3] B. G. Bodmann and N. Hammen, "Stable phase retrieval with low-redundancy frames," *Adv. Comput. Math.*, vol. 41, no. 2, pp. 317–331, Apr. 2015.
[4] P. M. Pardalos and S. A. Vavasis, "Quadratic programming with one negative eigenvalue is NP-hard," *J. Global Optim.*, vol. 1, no. 1, pp. 15–22, 1991.
[5] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*. SIAM, 2001, vol. 2.
[6] Y. Chen and E. J. Candès, "Solving random quadratic systems of equations is nearly as easy as solving linear systems," in *Adv. Neural Inf. Process. Syst.*, 2015, pp. 739–747.
[7] J. R. Fienup, "Reconstruction of an object from the modulus of its Fourier transform," *Opt. letters*, vol. 3, no. 1, pp. 27–29, July 1978.
[8] E. J. Candès, Y. C. Eldar, T. Strohmer, and V. Voroninski, "Phase retrieval via matrix completion," *SIAM Rev.*, vol. 57, no. 2, pp. 225–251, May 2015.
[9] K. Jaganathan, Y. C. Eldar, and B. Hassibi, "Phase retrieval: An overview of recent developments," *arXiv:1510.07713*, 2015.

[10] J. Miao, P. Charalambous, J. Kirz, and D. Sayre, "Extending the methodology of X-ray crystallography to allow imaging of micrometre-sized non-crystalline specimens," *Nature*, vol. 400, no. 6742, pp. 342–344, July 1999.

[11] R. P. Millane, "Phase retrieval in crystallography and optics," *J. Opt. Soc. Am. A*, vol. 7, no. 3, pp. 394–411, 1990.

[12] L. Bian, J. Suo, G. Zheng, K. Guo, F. Chen, and Q. Dai, "Fourier ptychographic reconstruction using Wirtinger flow optimization," *Opt. Express*, vol. 23, no. 4, pp. 4856–4866, 2015.

[13] A. Chai, M. Moscoso, and G. Papanicolaou, "Array imaging using intensity-only measurements," *Inverse Probl.*, vol. 27, no. 1, p. 015005, Dec. 2011.

[14] S. Marchesini, Y.-C. Tu, and H.-T. Wu, "Alternating projection, ptychographic imaging and phase synchronization," *Appl. Comput. Harmon. Anal.*, June 2015, to appear.

[15] C. Fienup and J. Dainty, "Phase retrieval and image reconstruction for astronomy," *Image Recovery: Theory and Application*, pp. 231–275, 1987.

[16] J. Miao, I. 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, May 2008.

[17] 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, Aug. 1991.

[18] C. J. Hillar and L.-H. Lim, "Most tensor problems are NP-hard," *J. ACM*, vol. 60, no. 6, p. 45, 2013.

[19] K. G. Murty and S. N. Kabadi, "Some NP-complete problems in quadratic and nonlinear programming," *Math. Prog.*, vol. 39, no. 2, pp. 117–129, 1987.

[20] 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 Proc. Mag.*, vol. 32, no. 3, pp. 87–109, May 2015.

[21] E. Hofstetter, "Construction of time-limited functions with specified autocorrelation functions," *IEEE Trans. Inf. Theory*, vol. 10, no. 2, pp. 119–126, Apr. 1964.

[22] Y. Shechtman, A. Beck, and Y. C. Eldar, "GESPAR: Efficient phase retrieval of sparse signals," vol. 62, no. 4, pp. 928–938, Feb. 2014.

[23] K. Jaganathan, S. Oymak, and B. Hassibi, "Sparse phase retrieval: Uniqueness guarantees and recovery algorithms," *arXiv:1311.2745*, 2013.

[24] Y. C. Eldar and S. Mendelson, "Phase retrieval: Stability and recovery guarantees," *Appl. Comput. Harmon. Anal.*, vol. 36, no. 3, pp. 473–494, May 2014.

[25] Y. C. Eldar, P. Sidorenko, D. G. Mixon, S. Barel, and O. Cohen, "Sparse phase retrieval from short-time Fourier measurements," *IEEE Signal Process. Lett.*, vol. 22, no. 5, pp. 638–642, May 2015.

[26] E. J. Candès, X. Li, and M. Soltanolkotabi, "Phase retrieval from coded diffraction patterns," *Appl. Comput. Harmon. Anal.*, vol. 39, no. 2, pp. 277–299, Sep. 2015.

[27] P. Netrapalli, P. Jain, and S. Sanghavi, "Phase retrieval using alternating minimization," in *Adv. Neural Inf. Process. Syst.*, 2013, pp. 2796–2804.

[28] E. J. Candès, T. Strohmer, and V. Voroninski, "PhaseLift: Exact and stable signal recovery from magnitude measurements via convex programming," *Appl. Comput. Harmon. Anal.*, vol. 66, no. 8, pp. 1241–1274, 2013.

[29] 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, Apr. 2015.

[30] H. Zhang, Y. Chi, and Y. Liang, "Provable non-convex phase retrieval with outliers: Median truncated Wirtinger flow," *arXiv:1603.03805*, 2016.

[31] R. W. Gerchberg and W. O. Saxton, "A practical algorithm for the determination of phase from image and diffraction," *Optik*, vol. 35, pp. 237–246, Nov. 1972.

[32] J. R. Fienup, "Phase retrieval algorithms: A comparison," *Appl. Opt.*, vol. 21, no. 15, pp. 2758–2769, Aug. 1982.

[33] M. Soltanolkotabi, "Algorithms and theory for clustering and nonconvex quadratic programming," Ph.D. dissertation, Stanford University, 2014.

[34] K. Wei, "Solving systems of phaseless equations via Kaczmarz methods: A proof of concept study," *Inverse Probl.*, vol. 31, no. 12, p. 125008, 2015.

[35] J. Sun, Q. Qu, and J. Wright, "A geometric analysis of phase retrieval," *arXiv:1602.06664*, 2016.

[36] S. Tu, R. Boczar, M. Soltanolkotabi, and B. Recht, "Low-rank solutions of linear matrix equations via Procrustes flow," *arXiv:1507.03566*, 2015.

[37] S. Sanghavi, R. Ward, and C. D. White, "The local convexity of solving systems of quadratic equations," *Results Math.*, pp. 1–40, June 2016.

[38] A. G. Marques, G. Mateos, and Y. C. Eldar, "SIGIBE: Solving random bilinear equations via gradient descent with spectral initialization," in *Proc. of European Signal Process. Conf.*, Budapest, Hungary, Aug. 2016.

[39] X. Li, S. Ling, T. Strohmer, and K. Wei, "Rapid, robust, and reliable blind deconvolution via nonconvex optimization," *arXiv:1606.04933*, 2016.

[40] R. Sun and Z.-Q. Luo, "Guaranteed matrix completion via nonconvex factorization," in *IEEE 56th Annual Symposium on Foundations of Computer Science*, 2015, pp. 270–289.

[41] N. Z. Shor, "Quadratic optimization problems," *USSR Technical Cybernetics (Moscow)*, vol. 1, pp. 102–106, 1987.

[42] I. Waldspurger, A. dAspremont, and S. Mallat, "Phase recovery, maxcut and complex semidefinite programming," *Math. Prog.*, vol. 149, no. 1-2, pp. 47–81, 2015.

[43] K. Huang, Y. C. Eldar, and N. D. Sidiropoulos, "Phase retrieval from 1D Fourier measurements: Convexity, uniqueness, and algorithms," *arXiv:1603.05215*, 2016.

[44] C. Qian, X. Fu, N. D. Sidiropoulos, L. Huang, and J. Xie, "Inexact alternating optimization for phase retrieval in the presence of outliers," *arXiv:1605.00973v1*, 2016.

[45] Y. C. Eldar and S. Mendelson, "Phase retrieval: Stability and recovery guarantees," *Appl. Comput. Harmon. Anal.*, vol. 36, no. 3, pp. 473–494, May 2014.

[46] E. J. Candès and X. Li, "Solving quadratic equations via PhaseLift when there are about as many equations as unknowns," *Found. Comput. Math.*, vol. 14, no. 5, pp. 1017–1026, 2014.

[47] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, vol. 3, no. 27, pp. 20–34, May 2010.

[48] R. H. Keshavan, A. Montanari, and S. Oh, "Matrix completion from a few entries," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2980–2998, June 2010.

[49] G. Wang, D. Berberidis, V. Kekatos, and G. B. Giannakis, "Online reconstruction from big data via compressive censoring," in *IEEE Global Conf. Signal and Inf. Process.*, Atlanta, GA, 2014, pp. 326–330.

[50] D. K. Berberidis, V. Kekatos, G. Wang, and G. B. Giannakis, "Adaptive censoring for large-scale regressions," in *IEEE Intl. Conf. Acoustics, Speech and Signal Process.*, South Brisbane, QLD, Australia, 2015, pp. 5475–5479.

[51] L.-H. Yeh, J. Dong, J. Zhong, L. Tian, M. Chen, G. Tang, M. Soltanolkotabi, and L. Waller, "Experimental robustness of Fourier ptychography phase retrieval algorithms," *Opt. Express*, vol. 23, no. 26, pp. 33 214–33 240, Dec. 2015.

[52] T. Cai, J. Fan, and T. Jiang, "Distributions of angles in random packing on spheres," *J. Mach. Learn. Res.*, vol. 14, no. 1, pp. 1837–1864, Jan. 2013.

[53] N. Z. Shor, "A class of almost-differentiable functions and a minimization method for functions of this class," *Cybern. Syst. Anal.*, vol. 8, no. 4, pp. 599–606, July 1972.

[54] R. Rockafellar and R. J.-B. Wets, *Variational Analysis*. Berlin-Heidelberg: Springer Verlag, 1998.

[55] N. Z. Shor, K. C. Kiwiel, and A. Ruszcayǹski, *Minimization Methods for Non-differentiable Functions*. Springer-Verlag New York, Inc., 1985.

[56] F. H. Clarke, *Optimization and Nonsmooth Analysis*. SIAM, 1990, vol. 5.

[57] ——, "Generalized gradients and applications," *T. Am. Math. Soc.*, vol. 205, pp. 247–262, 1975.

[58] P. Chen, A. Fannjiang, and G.-R. Liu, "Phase retrieval with one or two diffraction patterns by alternating projections of the null vector," *arXiv:1510.07379v2*, 2015.

[59] P. Chen and F. A., "Fourier phase retrieval with a single mask by Douglas-Rachford algorithm," *arXiv:1509.00888*, 2015.

[60] R. Vershynin, "Introduction to the non-asymptotic analysis of random matrices," *arXiv:1011.3027*, 2010.

[61] R. H. Keshavan, A. Montanari, and S. Oh, "Matrix completion from a few entries," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2980–2998, Jun. 2010.

[62] S. Cambanis, S. Huang, and G. Simons, "On the theory of elliptically contoured distributions," *J. Multivar. Anal.*, vol. 11, no. 3, pp. 368–385, Sep. 1981.

[63] B. Laurent and P. Massart, "Adaptive estimation of a quadratic functional by model selection," *Ann. Stat.*, vol. 28, no. 5, pp. 1302–1338, 2000.

[64] S.-H. Chang, P. C. Cosman, and L. B. Milstein, "Chernoff-type bounds for the Gaussian error function," *IEEE Trans. Commun.*, vol. 59, no. 11, pp. 2939–2944, July 2011.

[65] T. S. Ferguson, "A representation of the symmetric bivariate Cauchy distribution," *Ann. Math. Stat.*, vol. 33, no. 4, pp. 1256–1266, 1962.

[66] H. Y. Lee, G. J. Parka, and H. M. Kim, "A clarification of the Cauchy distribution," *Commun. Stat. Appl. Methods*, vol. 21, no. 2, pp. 183–191, Mar. 2014.