

# Harnessing the Power of Sample Abundance: Theoretical Guarantees and Algorithms for Accelerated One-Bit Sensing

Arian Eamaz, Student Member, IEEE, Farhang Yeganegi, Deanna Needell,  
Member, IEEE, and  
Mojtaba Soltanalian, Senior Member, IEEE

## Abstract

One-bit quantization with time-varying sampling thresholds (also known as random dithering) has recently found significant utilization potential in statistical signal processing applications due to its relatively low power consumption and low implementation cost. In addition to such advantages, an attractive feature of one-bit analog-to-digital converters (ADCs) is their superior sampling rates as compared to their conventional multi-bit counterparts. This characteristic endows one-bit signal processing frameworks with what one may refer to as sample abundance. We show that sample abundance plays a pivotal role in many signal recovery and optimization problems that are formulated as (possibly non-convex) quadratic programs with linear feasibility constraints. Of particular interest to our work are low-rank matrix recovery and compressed sensing applications that take advantage of one-bit quantization. We demonstrate that the sample abundance paradigm allows for the transformation of such problems to merely linear feasibility problems by forming large-scale overdetermined linear systems—thus removing the need for

This work relates in part to Department of Navy Award N00014-22-1-2666 issued by the Office of Naval Research. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Office of Naval Research. A conference precursor to this work was presented at the 2023 IEEE International Symposium on Information Theory (ISIT). The first two authors contributed equally to this work.

A. Eamaz, F. Yeganegi, and M. Soltanalian are with the Department of Electrical and Computer Engineering, University of Illinois Chicago, Chicago, IL 60607, USA (Corresponding author: Arian Eamaz).

D. Needell is with the Department of Mathematics, University of California Los Angeles, Los Angeles, CA 90095, USA.

handling costly optimization constraints and objectives. To make the proposed computational cost savings achievable, we offer enhanced randomized Kaczmarz algorithms to solve these highly overdetermined feasibility problems and provide theoretical guarantees in terms of their convergence, sample size requirements, and overall performance. Several numerical results are presented to illustrate the effectiveness of the proposed methodologies.

## Index Terms

Convex-relaxed problems, compressed sensing, low-rank matrix sensing, one-bit quantization, one-bit ADCs, randomized Kaczmarz algorithm, statistical signal processing, time-varying sampling thresholds.

## I. Introduction

We consider an optimization problem of the form:

$$\underset{\mathbf{X} \in \Omega_c}{\text{minimize}} \quad f(\mathbf{X}) \quad \text{subject to} \quad \mathcal{A}(\mathbf{X}) = \mathbf{y}, \quad (1)$$

where  $f(\cdot)$  is a cost function,  $\Omega_c$  is a feasible set,  $\mathbf{X} \in \mathbb{C}^{n_1 \times n_2}$  is the matrix of unknowns,  $\mathbf{y} \in \mathbb{R}^n$  is the measurement vector, and  $\mathcal{A}$  is a linear transformation mapping  $\mathbb{C}^{n_1 \times n_2}$  into  $\mathbb{R}^n$ . This problem emerges in a wide variety of applications, particularly as the relaxed variant of some well-known NP-hard problems. Although many problems can be expressed in the form in (??), the applications we will focus on in this paper include some specific problems of interest which can take advantage of low-resolution (and particularly one-bit) sampling and processing:

- The task of recovering a low-rank matrix from its linear measurements plays a central role in computational science. The problem occurs in many areas of applied mathematics, such as signal processing [?, ?, ?, ?, ?, ?, ?], machine learning [?, ?, ?, ?, ?, ?], and computer vision [?]. In this scenario, the cost function of (??),  $f(\cdot)$ , is typically the nuclear norm or the Frobenius norm, and the constraint set  $\Omega_c$  would be an amplitude restriction on the elements of matrix  $\mathbf{X}$ ; see [?, ?].
- Compressed sensing (CS) offers a framework for the simultaneous sensing and compression of finite-dimensional vectors by relying on linear dimensionality reduction. Through a CS formulation, sparse signals may be recovered from highly incomplete measurements [?, ?, ?]. The problem (1) can be adopted in the CS context when  $f(\mathbf{X}) = \|\text{vec}(\mathbf{X})\|_1$ .

This paper will primarily concentrate on relaxation problems in which the objective function  $f(\cdot)$  is relaxed to address the NP-hard cost caused due to the associated feasibility constraints. Such relaxation techniques prove significantly helpful in various applications, such as nuclear norm minimization for low-rank matrix recovery and  $\ell_1$  minimization for compressed sensing problems.

Sampling the signals of interest at high data rates with high-resolution ADCs would dramatically increase the overall manufacturing cost and power consumption of such ADCs. In multi-bit sampling scenarios, a very large number of quantization levels is necessary in order to represent the original continuous signal in with high accuracy, which in turn leads to a considerable reduction in sampling rate [?, ?]. This attribute of multi-bit sampling is the key reason for the general emergence of underdetermined systems  $n_1 n_2 \geq n$  in (??) [?, ?, ?]. An alternative solution to such challenges is to deploy one-bit quantization which is an extreme sampling scenario, where the signals are merely compared with given threshold levels at the ADCs, producing sign data ( $\pm 1$ ). This enables signal processing equipments to sample at a very high rate, with a considerably lower cost and energy consumption, compared to their counterparts which employ multi-bit ADCs [?, ?, ?, ?]. Several applications abound of one-bit ADCs, such as multiple-input multiple-output wireless communications [?, ?, ?], channel estimation [?, ?], and array signal processing [?]. Extensive research on one-bit Sigma-Delta quantization explores the impact of adaptive threshold selection on achieving rapidly decaying reconstruction errors with respect to the number of measurements or the oversampling rate, depending on the scenario. For more detailed insights, readers can refer to relevant works such as [?, ?, ?]. Various methods for one-bit quantization, including our own, offer distinct trade-offs between computational and implementation complexity and reconstruction accuracy. Take the Sigma-Delta approach as an example, which necessitates memory elements to store state-variables associated with thresholds. This sequential quantization of incoming measurements demands significant memory and computational power, causing delays in acquiring measurements while the thresholds are being updated. Therefore, sensors used to collect measurements must possess substantial computational capabilities to accommodate such requirements [?, ?].

The classical one-bit sampling problem involves reconstructing a signal by comparing it with a fixed threshold, often set at zero. This method has limitations in accurately estimating signal parameters, especially when the input signal  $\mathbf{x}$  is converted into one-bit data since

the power information is lost. This occurs because the signs of  $\mathbf{x}$  and  $\eta\mathbf{x}$  are the same for  $\eta > 0$ . The effectiveness of incorporating random thresholds (dithers) within the framework of one-bit quantization has been extensively established in various contexts [?, ?]. Extensive research has been dedicated to investigating the impact of thresholding (dithering) on the quantization process and its effect on quantization error [?, ?, ?, ?, ?, ?]. In the evaluation of quantizing systems, particularly in digital signal processing applications, the mean squared error (MSE) serves as a relevant parameter for assessing performance. Dithering provides the flexibility to control system performance by enabling a trade-off between accuracy and resolution. An intriguing property of dithering is its ability to reduce the overall average quantization error [?].

Various dithering schemes have been studied, including uniform dither, discrete dither, and Gaussian dither. Research findings suggest that a carefully chosen dither signal can significantly improve the resolution and performance of digital instrumentation. While discrete dithering has been examined, it generally does not perform as effectively as uniform or Gaussian dithering in terms of reducing quantization error [?]. Uniformly distributed dither sequences have been selected due to their effectiveness in minimizing the average quantization error under appropriate conditions [?]. Gaussian distributed dither signals, on the other hand, are favored for their ability to provide good resolution [?, ?]. Furthermore, studies have shown that while Uniform dithering theoretically offers infinite resolution, its performance can rapidly deteriorate when uncertainties in dither amplitude are present. In contrast, Gaussian dithering has proven to be more robust in practical scenarios, exhibiting fewer sensitivity issues [?, ?]. After a thorough analysis in [?], it becomes evident that when considering both resolution and accuracy in quantizers, the choice between Gaussian and uniformly distributed dither signals has minimal impact on the performance of the quantizing scheme.

The implementation of a dithered generator, specifically for generating Gaussian dithering in quantization systems within an ADC system, was described in [?]. According to their work, the dither generator produces Gaussian dithering by generating a random analog noise signal. One method of achieving this is by utilizing a low-cost thermal noise diode, which introduces analog Gaussian dither. The intrinsic quantum mechanical properties of electron-hole pairing in such devices generate a truly random noise signal. In some cases, the power levels of the noise generated by certain dither generators may be relatively small. Therefore,

optional processing circuitry can be employed alongside the dither generator, which may include components such as gain control circuitry for the analog noise signal, low-cost operational amplifiers, or similar elements, as required. Additionally, the implementation of multiple random dithering was demonstrated in [?] for a 12-bit, 18-GS/s ADC.

Recent works, such as [?, ?, ?, ?, ?], have proposed using time-varying thresholds to improve the estimation of signal characteristics. For instance, [?] considered Gaussian dithering for recovering covariance from one-bit measurements, while [?] extended this method for better estimation of signal autocorrelation through the "modified arcsine law." In non-stationary scenarios, the authors in [?] utilized time-varying sampling thresholds with the modified arcsine law. In [?], the authors provide a comprehensive investigation of estimating the covariance matrix from one-bit measurements using Uniform dithering. Dithered One-bit quantization has also been applied to various problems, such as sparse parameter estimation [?, ?, ?], phase retrieval [?], and sampling theory [?] as discussed in contemporary literature. The study presented in [?] demonstrates that one-bit signal parameter estimation with time-varying thresholds yields notably superior results compared to using constant thresholds. Similarly, constant thresholds, just like zero thresholds, have the potential to overlook certain signal information. While Uniform dithering has been extensively studied and proven to offer theoretical guarantees and good performance for bounded measurements, recent studies [?, ?] have shown that Gaussian dithering outperforms Uniform dithering in signal reconstruction. Specifically, in the context of Cramér-Rao bound (CRB) analysis, Gaussian dithering exhibits better performance for input signals with bell-shaped distributions.

Randomized iterative algorithms have become a widely employed technique for dimension reduction, encompassing various iterative methods for solving linear systems and their variations, as well as extensions to non-linear optimization problems [?]. These algorithms include well-known methods such as randomized Kaczmarz algorithm (RKA) [?, ?], coordinate descent [?, ?], and variants of the Newton method in convex optimization [?], among others. In this approach, instead of utilizing all available information at each iteration, a sketched subset of the information is used, selected through a sketching process. By multiplying the data with a random matrix, randomized sketching effectively reduces the size of the problem while preserving its essential characteristics. Common examples of sketching matrices include Gaussian matrices with independent entries and other matrices that satisfy the Johnson-Lindenstrauss property. Another example is block-identity matrices,

which effectively subsample the data matrix. The Kaczmarz method [?] is an iterative projection algorithm for solving linear systems of equations and inequalities. It is usually applied to highly overdetermined systems because of its simplicity. Each iteration projects onto the solution space corresponding to one row in the linear system, in a sequential regimen. The method has been applied to various applications in image reconstruction, digital signal processing, and computer tomography [?, ?, ?]. Many variants of this iterative method and their convergence rates have been proposed and studied in recent decades for both consistent and inconsistent systems including the randomized Kaczmarz algorithm, the randomized block Kaczmarz algorithm and most recently, the sampling Kaczmarz-Motzkin method [?, ?, ?, ?, ?].

#### A. Prior Art

Our work aligns seamlessly with the recent advancements in the exploration of coarse quantization for parameter estimation from random measurements. The emergence of one-bit compressed sensing (one-bit CS) has led to a plethora of papers that have greatly enhanced our understanding of employing dithering for one-bit signal reconstruction.

The domain of one-bit low-rank matrix sensing has garnered considerable interest, yet it remains relatively unexplored. Several noteworthy papers, such as [?, ?], have shed light on this intriguing subject. In [?], the authors delve into the theoretical guarantees of one-bit sensing from low-rank matrices, employing two distinct algorithms: one based on hard singular value thresholding and the other utilizing semidefinite programming. The initial investigation in this study considers a scenario without thresholds and subsequently progresses to the Gaussian sampling matrix and Gaussian dithering scenario, where both adaptive and non-adaptive thresholds are employed. However, the recovery algorithms and theoretical guarantees presented in this study require the availability of side information regarding the signal of interest, specifically an upper bound on the Frobenius norm of input signal. This requirement contradicts the essence of employing dithering in one-bit sensing, as the literature suggests that the primary advantage of dithering lies in the ability to recover the signal magnitude without additional information. In the notable work of [?], the authors explore one-bit sensing for both low-rank and bispase matrices. Notably, they achieve a significant milestone by deriving tailored theoretical guarantees for these types of matrices, providing a novel contribution to the field. Additionally, they investigate the

impact of coarse quantization in thresholdless scenarios, focusing on recovering only the signal direction rather than its exact magnitude. It is worth mentioning that both of these influential papers exclusively focus on scenarios involving Gaussian sampling matrices.

One-bit CS was first introduced in [?]. Later, in [?], this method was extensively investigated, and they presented a lower bound on the reconstruction error, alongside proposing heuristic algorithms for recovering the underlying signals. Extensive research has been dedicated to the problem of one-bit CS without dithering, resulting in numerous proposed approaches for its solution. Notably, the binary iterative hard thresholding (BIHT) and its normalized version, normalized BIHT (NBIHT), have emerged as prominent methods in this field [?, ?, ?, ?, ?, ?, ?, ?]. Support recovery in one-bit CS has been extensively studied in the literature, resulting in a wealth of research that has yielded improved bounds for performance guarantees [?, ?, ?]. In some earlier works, the ditherless setup was taken into account with a further assumption on the norm of the signal,  $\|\mathbf{x}\|_2 = 1$ , leading to the derivation of various theoretical guarantees for reconstruction algorithms [?, ?, ?, ?, ?]. For instance, the authors of [?] presented the required number of measurements for robust sparse signal reconstruction within a ball around the original signal, specifically focusing on Gaussian sampling matrices. However, recent studies have shown that under appropriate conditions, complete signal reconstruction can be accomplished by employing nonzero random thresholds [?, ?, ?, ?]. The authors of [?] introduced two algorithms to address the one-bit CS problem with nonzero thresholds under the second norm restriction. The first method utilizes augmented convex programming to recover the signal  $\mathbf{x}$  from affine one-bit measurements with assumed structural constraint  $\|\mathbf{x}\|_1/\|\mathbf{x}\|_2 \leq \sqrt{s}$ . Note that the error guarantees associated with this method (augmented convex programming) were derived in the case of the Gaussian dither. The second method firstly estimates the norm of the unknown signal from the one-bit measurements by empirical cumulative distribution function, leveraging the Gaussianity of the sampling matrix. Specifically, this method partitions the measurements into two sets and utilizes a constant threshold in one-bit data acquisition, thereby estimating the second norm of the signal  $\|\mathbf{x}\|_2$ . Simultaneously, this method employs a zero threshold for the remaining one-bit data acquisition part, allowing for the estimation of the direction of the signal. In [?], the authors addressed the problem of one-bit CS with adaptively designed thresholds and Gaussian thresholds, while considering the second norm restriction,  $\|\mathbf{x}\|_2 \leq R$ . They provided theoretical guarantees specifically for Gaussian sampling matrices.

The problem was approached through (i) the second order cone programming, and (ii) hard thresholding based order one recovery scheme (BIHT algorithm). In [?], the authors proposed the second order cone program for the one-bit CS problem with the Gaussian circulant matrix as the sensing matrix. They also established the theoretical guarantees for their approach by leveraging the  $\ell_1/\ell_2$ -restricted isometry property (RIP) for the sampling matrix. A recent study, as demonstrated in [?], has revealed that one-bit CS is feasible for certain non-Gaussian sampling matrices, while considering the norm restriction  $\|\mathbf{x}\|_2 \leq 1$ , and employing Uniform thresholds. This finding expands the understanding of one-bit CS beyond Gaussian sampling matrices, opening up new possibilities and avenues for research in this field. In [?], the authors tackled the problem of parameter estimation with coarse quantization using Uniform dithering by formulating it as a generalized LASSO problem. This approach allowed the authors to leverage the existing theoretical guarantees available in the literature for generalized LASSO. They achieved this by constructing a quadratic objective that bridges the gap between full precision measurements and quantized data (nonlinear measurements) through a tuning parameter selected based on the quantization process. In the work presented in [?], the authors built upon the foundations laid in [?] and extended its application to various parameter estimation problems. They derive an upper bound for the performance guarantee, which they show to be a nearly optimal minimax rate in the sub-Gaussian regime. Their results encompass scenarios involving both sub-Gaussian and heavy-tailed data, using Uniform dithering and bounded measurements. When faced with situations where the measurements are not bounded, they employ truncation on the measurements to leverage the guarantees obtained from the assumption of bounded measurements with Uniform dithering. Also, authors of [?] comprehensively studied the one-bit CS problem with sub-sampled circulant sensing matrices and Uniform thresholds. Another recent study [?] presented that a scalar quantization scheme combined with a uniformly distributed dithering signal promises pushing much of the theory beyond Gaussian measurements. They established theoretical guarantees for the sampling matrices that satisfy the RIP under certain conditions.

Our work was primarily motivated by two key references, [?] and [?]. Similar to these works, we explore the theoretical guarantees of one-bit sensing without being confined to a specific problem formulation. In this paper, we try to broaden the theoretical guarantees beyond Uniform dithering. Specifically, we seek to include various sampling matrices,



including deterministic ones like the Discrete Cosine Transform (DCT). Additionally, we delve into the scenario of sample abundance, a common situation in one-bit sensing, where we reformulate the relaxed optimization problem (??) into a linear feasibility problem. Strong linear feasibility solvers only provide a solution that satisfies all inequalities. Nevertheless, we require a solution that not only satisfies the given inequalities but also fulfills the constraints of the original problem (??). Using our theorems, we aim to address the question of how many samples are necessary for a linear inequality system solver to obtain a solution within the ball with a specific radius, around the original signal. Moreover, we endeavor to harness the advantages of randomized algorithms for one-bit sensing, a novel endeavor where we explore their theoretical guarantees and convergence analysis in this context for the first time. It is important to note that all the theoretical guarantees derived in this paper are uniform reconstruction guarantees. This means that they hold true for all desired signals in the given space, ensuring consistent and reliable reconstruction across the entire signal set.

## B. Contributions of the Paper

This paper principally contributes to the following areas:

- 1) RKA-based recovery for the dithered one-bit sensing. In this paper, we consider the deployment of one-bit sampling with time-varying thresholds, leading to an increased sample size and a highly overdetermined system as a result. The proposed One-bit aided Randomized Kaczmarz Algorithm, which we refer to as ORKA, can find the desired signal  $\mathbf{X}_\star$  in (??) by (i) generating abundant one-bit measurements, in order to define a large scale overdetermined system where a finite volume feasible set is created for (??), and (ii) solving this obtained linear feasibility problem by leveraging one of the efficient solver families of overdetermined systems, Kaczmarz algorithms. In [?], we showed that the sheer number of measurements acquired in one bit sampling facilitates recovering the signal of interest in a less costly manner by making costly constraints such as semidefiniteness and rank redundant. Then, a simple RKA was utilized to solve the obtained linear feasibility problem. This idea is generalized in this paper to (??) where we generate the abundant samples and eventually introduce a one-bit linear feasibility region named the one-bit polyhedron. In other words, by using this technique, we make (??) into a large-scale overdetermined system, an ideal application setting for Kaczmarz algorithms. To solve our highly overdetermined system,

we propose two novel variants of RKA which will be compared with the existing RKA variants. We conduct a thorough investigation into the robustness of RKA when dealing with a noisy linear inequality system. Our findings demonstrate that RKA for the linear inequality system, remains robust in the presence of noise, and we are able to obtain the upper recovery bound under this condition. Furthermore, an algorithm is proposed based on our model to design the time-varying sampling thresholds.

2) Theoretical Guarantees of dithered one-bit sensing. In order to determine the necessary number of one-bit samples for achieving perfect signal reconstruction, we introduce the concept of the finite volume property (FVP). The FVP explores an upper bound on the probability of creating a finite volume enclosed by hyperplanes around the original signal. By leveraging the FVP, we can obtain the required number of samples needed for one-bit sensing to ensure that the solution lies within a ball around the original signal, thereby enabling perfect reconstruction. This required number of samples is investigated for various scenarios, including low-rank matrices and sparse signals, considering different sampling matrices and dithering schemes. Additionally, the convergence rate of the proposed algorithms is derived for both noiseless and noisy scenarios.

3) Projected Kaczmarz algorithm-based recovery. Various sampling schemes present trade-offs among the number of samples, computational cost of the reconstruction, and storage requirements. While the abundance of samples is desirable in scenarios involving dithered one-bit sensing, as it proves effective with low-cost ADCs, practical applications such as matrix completion and phase retrieval from coded diffraction patterns [?] impose restrictions on the number of available measurements. In such cases, it is not feasible to increase the number of measurements or obtain informative samples even with the presence of multiple comparisons in the dithered sensing process. In these restricted scenarios, the only viable approach is to develop algorithms that require fewer samples, albeit at the cost of increased computational complexity. This can be achieved by incorporating convex projectors at each iteration, which helps mitigate the limitations imposed by the restricted number of measurements. While this may result in higher computational costs, it becomes necessary to handle the measurement constraints prevalent in these applications. In our pursuit of a storage-friendly algorithm for the matrix sensing problem, we strive to avoid the burden of recording the full matrix per iteration. To achieve this, we leverage the power of two well-studied techniques: sketch-and-project and low-rank matrix factorization. These approaches

are skillfully integrated with our proposed randomized algorithms.

4) One-bit low-rank matrix sensing with sample abundance and randomized algorithms. In what we believe is a pioneering contribution to the field, we unveil a novel approach that combines low-rank matrix factorization with RKA to address the challenging problem of one-bit low-rank matrix sensing. In traditional matrix sensing, a large number of measurements are typically required for perfect recovery due to the inherent dimensionality of the matrix space. However, our proposed randomized algorithms offer a more efficient solution by leveraging only a subset of the measurements in each iteration. A key advantage of our proposed algorithms is that they do not rely on any additional side information, such as known upper bounds on the Frobenius norm. This makes them more practical and versatile for real-world applications. Through comprehensive numerical evaluations, we demonstrate the superior performance of our proposed algorithms compared to state-of-the-art approaches.

5) One-bit CS with sample abundance and randomized algorithm. We utilize the proposed randomized algorithms to tackle the one-bit CS problem, incorporating multiple time-varying dithering. Initially, we apply the randomized algorithms under the sample abundance condition, where the number of measurements is not a limiting factor. Subsequently, we employ projected algorithms to address scenarios afflicted by the curse of dimensionality. One of the key advantages of employing randomized algorithms in one-bit CS, as compared to methods like BIHT [?, ?], is that instead of utilizing all one-bit data in each iteration, we leverage sketched one-bit data, representing only a subset of the complete data. This approach leads to a highly efficient algorithm. We will provide numerical evidence demonstrating the superior performance of the proposed algorithms in both ditherless and dithering scenarios.

### C. Organization and Notations

Section II is dedicated to a review of proximal methods which have been utilized to tackle (??) by projecting the final solution on the desired feasible set. In Section III, we will introduce our algorithm to solve (??), ORKA, which tackles the problem as a large-scale overdetermined system and finds the original signal in the one-bit polyhedron by an accelerated Kaczmarz approach. Moreover, two new variants of the Kaczmarz algorithms are proposed that enhance the convergence rate and the computational complexity of these solvers. To investigate the convergence rate of ORKA, we will first introduce the FVP

theorem in Section IV based on the general Hoeffding's bound. In Section V, we delve into the robustness of the proposed algorithms in the presence of noise. We thoroughly discuss the impact of noise on the performance of the algorithms and address the need for novel algorithms specifically designed for noisy scenarios. Section VI discusses an iterative algorithm to achieve optimized time-varying sampling threshold sequences which benefit the signal recovery process with enhanced accuracy. As a representative application, in Section VII, ORKA and other proposed algorithms will be applied in the context of low-rank matrix recovery in the form of a nuclear norm minimization problem. In Section VIII, we apply the ORKA and other proposed algorithms within the context of one-bit CS. Dedicated to numerical evaluations, Section IX focuses on investigating the performance of the proposed algorithms in the tasks of one-bit low-rank matrix sensing and one-bit CS. We consider scenarios involving both noisy and noiseless conditions. Through comprehensive comparisons with state-of-the-art methods, we aim to demonstrate the superior performance and efficiency of our proposed algorithms. Finally, Section X concludes the paper.

**Notation:** We use bold lowercase letters for vectors and bold uppercase letters for matrices.  $\mathbb{C}$  and  $\mathbb{R}$  represent the set of complex and real numbers, respectively.  $(\cdot)^\top$  and  $(\cdot)^H$  denote the vector/matrix transpose, and the Hermitian transpose, respectively.  $\mathbf{I}_N \in \mathbb{R}^{N \times N}$  is the identity matrix of size  $N$ .  $\text{Tr}(\cdot)$  denotes the trace of the matrix argument.  $\langle \mathbf{B}_1, \mathbf{B}_2 \rangle = \text{Tr}(\mathbf{B}_1^H \mathbf{B}_2)$  is the standard inner product between two matrices. The nuclear norm of a matrix  $\mathbf{B} \in \mathbb{C}^{N_1 \times N_2}$  is denoted  $\|\mathbf{B}\|_* = \sum_{i=1}^M \sigma_i$  where  $M$  and  $\{\sigma_i\}$  are the rank and singular values of  $\mathbf{B}$ , respectively. The Frobenius norm of a matrix  $\mathbf{B}$  is defined as  $\|\mathbf{B}\|_F = \sqrt{\sum_{r=1}^{N_1} \sum_{s=1}^{N_2} |b_{rs}|^2}$  where  $\{b_{rs}\}$  are elements of  $\mathbf{B}$ . The  $\ell_k$ -norm of a vector  $\mathbf{b}$  is defined as  $\|\mathbf{b}\|_k^k = \sum_i |b_i|^k$ . The Hadamard (element-wise) product of two matrices  $\mathbf{B}_1$  and  $\mathbf{B}_2$  is denoted as  $\mathbf{B}_1 \odot \mathbf{B}_2$ . Additionally, the Kronecker product is denoted as  $\mathbf{B}_1 \otimes \mathbf{B}_2$ . The vectorized form of a matrix  $\mathbf{B}$  is written as  $\text{vec}(\mathbf{B})$ .  $\mathbf{1}_s$  is the  $s$ -dimensional all-one vector. Given a scalar  $x$ , we define  $(x)^+$  as  $\max\{x, 0\}$ . The set  $[n]$  is defined as  $[n] = \{1, \dots, n\}$ .  $\text{Diag}\{\mathbf{b}\}$  denotes a diagonal matrix with  $\{b_i\}$  as its diagonal elements. A ball with radius  $r$  centered at a point  $\mathbf{y} \in \mathbb{R}^n$  is defined as  $\mathcal{B}_r(\mathbf{y}) = \{\mathbf{y}_1 \in \mathbb{R}^n \mid \|\mathbf{y} - \mathbf{y}_1\|_2 \leq r\}$ . The function  $\text{sgn}(\cdot)$  yields the sign of its argument. The function  $\log(\cdot)$  denotes the natural logarithm, unless its base is otherwise stated. The notation  $x \sim \mathcal{U}(a, b)$  means a random variable drawn from the uniform distribution over the interval  $[a, b]$  and  $x \sim \mathcal{N}(\mu, \sigma^2)$  represents the normal distribution with mean  $\mu$  and

variance  $\sigma^2$ . The sub-Gaussian norm of a random variable  $X$  characterized by

$$\|X\|_{\psi_2} = \inf \left\{ t > 0 : \mathbb{E} \left\{ e^{X^2/t^2} \right\} \leq 2 \right\}. \quad (2)$$

## II. Projections On Convex Sets: Dealing With Costly Constraints

To tackle (??), many non-convex and local optimization algorithms have been developed over the years. Nevertheless, in recent decades, convex programming formulations via relaxation have come to the fore to approximate global solutions. Within the convex framework, various iterative methods have been proposed to tackle the problem with a Lagrangian formulation such as Uzawa's algorithm and the proximal forward-backward splitting method (PFBS) [?, ?, ?]. Moreover, to keep the problem solution inside the constraint set  $\Omega_c$ , the orthogonal projection  $P_{\Omega_c}$  is applied to solutions at each iteration. The Lagrangian for (??) is written as  $\mathcal{L}(\mathbf{X}, \boldsymbol{\lambda}) = f(\mathbf{X}) + \langle \boldsymbol{\lambda}, \mathbf{y} - \mathcal{A}(\mathbf{X}) \rangle$ , where  $\boldsymbol{\lambda} \in \mathbb{R}^n$  [?]. Uzawa's algorithm aims to find a saddle point  $(\mathbf{X}_*, \boldsymbol{\lambda}_*)$ , where  $\sup_{\boldsymbol{\lambda}} \inf_{\mathbf{X}} \mathcal{L}(\mathbf{X}, \boldsymbol{\lambda}) = \inf_{\mathbf{X}} \sup_{\boldsymbol{\lambda}} \mathcal{L}(\mathbf{X}, \boldsymbol{\lambda})$ , with the iterative procedure:

$$\begin{cases} \mathbf{X}_k = P_{\Omega_c}(\mathcal{A}^*(\boldsymbol{\lambda}_{k-1})), \\ \boldsymbol{\lambda}_k = \boldsymbol{\lambda}_{k-1} + \alpha_k(\mathbf{y} - \mathcal{A}(\mathbf{X}_k)), \end{cases} \quad (3)$$

where  $\alpha_k$  is the step size and  $\mathcal{A}^*$  is the adjoint of  $\mathcal{A}$ . Since every linear equation can be reformulated in standard form, we recast  $\mathcal{A}(\mathbf{X}) = \mathbf{y}$  as  $\mathbf{A}\mathbf{x} = \mathbf{y}$ , where  $\mathbf{A} \in \mathbb{C}^{n \times n_1 n_2}$  is a matrix version of the operator  $\mathcal{A}$ , and  $\mathbf{x} = \text{vec}(\mathbf{X})$  [?]. The optimization problem (??) is equivalently given by [?, ?]

$$\underset{\mathbf{X} \in \Omega_c}{\text{minimize}} \quad g(\mathbf{X}) = \frac{1}{2} \|\mathbf{y} - \mathbf{A} \text{vec}(\mathbf{X})\|_2^2 + \lambda f(\mathbf{X}) \quad (4)$$

To solve this problem, instead of using proximal methods, a projected gradient method such as Nesterov iterative approach may be utilized, i.e.,  $\mathbf{X}_k = P_{\Omega_c}(\mathbf{X}_{k-1} - \alpha_k \nabla g(\mathbf{X}_{k-1}))$ .

Famous examples for  $P_{\Omega_c}$ , are the singular value thresholding operator (SVT) and the semi-definite orthogonal projector, respectively. SVT is useful when  $f(\mathbf{X}) = \|\mathbf{X}\|_*$ , mathematically defined as:  $P_{\Omega_c} = \mathbf{U} \text{Diag} \{(\sigma_k - \delta)^+\} \mathbf{V}^\top$  [?], where  $\mathbf{U}$  and  $\mathbf{V}$  are unitary matrices from singular value decomposition (SVD), and  $\{\sigma_k\}$  are the singular values. The approximate solution should be projected onto a feasible convex set at each iteration via recovering all singular values and eigenvalues and comparing their smaller elements with a threshold, which is quite expensive [?].

An interesting alternative to enforcing the feasible set  $\Omega_c$  in (??) emerges when one increases the number of samples  $n$ , and solves the overdetermined linear system of equations with  $n \geq n_1 n_2$ . In this sample abundance regimen, the linear constraint  $\mathcal{A}(\mathbf{X}) = \mathbf{y}$  may actually yield the optimum inside  $\Omega_c$ . As a result of increasing the number of samples, it is possible that the intersection of these hyperplanes will achieve the desired point without the need to consider costly constraints. However, this idea may face practical limitations in the case of multi-bit quantization systems since ADCs capable of ultra-high rate sampling are difficult and expensive to produce. Moreover, one cannot necessarily expect these constraints to intersect with  $\Omega_c$  in such a way to form a finite-volume space before the optimum is obtained [?, ?].

In the next section, by deploying the idea of one-bit quantization with time-varying thresholds, linear equality constraints are superseded by a massive array of linear inequalities in forming the feasible polyhedron. Therefore, by increasing the number of samples, a finite-volume space may be created inside  $\Omega_c$  with shrinking size; making projections on  $\Omega_c$  redundant. From a practical point of view, one-bit sampling is done efficiently at a very high rate with a significantly lower cost compared to its high-resolution counterpart. It has been examined in [?] that even though only partial information is made available to one-bit signal processing algorithms, they can achieve acceptable recovery performance with less complexity compared to the high-resolution scenario. Thus, it is both practical and necessary to study the ground-breaking opportunities that emerge in the context of the wide array of problems formulated as (??) due to the availability of a large number of one-bit samples.

### III. ORKA: Towards Circumventing Costly Constraints

In this section, we first formulate the one-bit quantization with multiple time-varying thresholds. In Sections ?? and ??, we present a summarized review of RKA and Sampling Kaczmarz-Motzkin algorithm (SKM), respectively. Then in Section ??, we propose a novel Kaczmarz method variant formulated based on the SKM and a preconditioning approach. One-bit sampling via time-varying thresholds will be combined with the proposed randomized Kaczmarz method to create highly overdetermined linear inequalities. This paves the way for the recovery of the desired signal  $\mathbf{X}_\star$  in (??) without solving the original optimization problem; merely by tacking accounts of its linear constraints. Due to the

block structure of the linear feasibility in ORKA, we will propose a block-based Kaczmarz algorithm accordingly in Section ??.

#### A. One-Bit Quantization With Multiple Time-Varying Thresholds

Let  $y_k = y(kT)$  denote the uniform samples of signal  $y(t)$  with the sampling rate  $1/T$ . In practice, the discrete-time samples occupy pre-determined quantized values. We denote the quantization operation on  $y_k$  by the function  $Q(\cdot)$ . This yields the scalar quantized signal as  $r_k = Q(y_k)$ . In one-bit quantization, compared to zero or constant thresholds, time-varying sampling thresholds yield a better recovery performance [?, ?]. These thresholds may be chosen from any distribution. In the case of one-bit quantization with such time-varying sampling thresholds, we have  $r_k = \text{sgn}(y_k - \tau_k)$ . The information gathered through the one-bit sampling with time-varying thresholds presented here may be formulated in terms of an overdetermined linear system of inequalities. We have  $r_k = +1$  when  $y_k > \tau_k$  and  $r_k = -1$  when  $y_k < \tau_k$ . Therefore, one can formulate the geometric location of the signal as  $r_k(y_k - \tau_k) \geq 0$ . Collecting all the elements in the vectors as  $\mathbf{y} = [y_k] \in \mathbb{R}^n$  and  $\mathbf{r} = [r_k] \in \mathbb{R}^n$ , we have  $\mathbf{r} \odot (\mathbf{y} - \boldsymbol{\tau}) \succeq \mathbf{0}$ , or equivalently

$$\boldsymbol{\Omega} \mathbf{y} \succeq \mathbf{r} \odot \boldsymbol{\tau}, \quad (5)$$

where  $\boldsymbol{\Omega} \triangleq \text{diag}\{\mathbf{r}\}$ . Denote the time-varying sampling threshold in  $\ell$ -th signal sequence by  $\boldsymbol{\tau}^{(\ell)}$ , where  $\ell \in [m]$ . It follows from (??) that

$$\boldsymbol{\Omega}^{(\ell)} \mathbf{y} \succeq \mathbf{r}^{(\ell)} \odot \boldsymbol{\tau}^{(\ell)}, \quad \ell \in [m], \quad (6)$$

where  $\boldsymbol{\Omega}^{(\ell)} = \text{diag}(\mathbf{r}^{(\ell)})$ . Denote the concatenation of all  $m$  sign matrices as

$$\tilde{\boldsymbol{\Omega}} = \left[ \begin{array}{c|c|c} \boldsymbol{\Omega}^{(1)} & \dots & \boldsymbol{\Omega}^{(m)} \end{array} \right]^\top, \quad \tilde{\boldsymbol{\Omega}} \in \mathbb{R}^{mn \times n}. \quad (7)$$

Rewrite the  $m$  linear inequalities in (??) as

$$\tilde{\boldsymbol{\Omega}} \mathbf{y} \succeq \text{vec}(\mathbf{R}) \odot \text{vec}(\mathbf{\Gamma}), \quad (8)$$

where  $\mathbf{R}$  and  $\mathbf{\Gamma}$  are matrices, whose columns are the sequences  $\{\mathbf{r}^{(\ell)}\}_{\ell=1}^m$  and  $\{\boldsymbol{\tau}^{(\ell)}\}_{\ell=1}^m$ , respectively.

Assuming a large number of samples — a common situation in one-bit sampling scenarios — hereafter we treat (??) as an overdetermined linear system of inequalities associated with the one-bit sensing scheme. The inequality (??) can be recast as a polyhedron,

$$\mathcal{P}_{\mathbf{y}} = \left\{ \mathbf{y} \mid \tilde{\mathbf{\Omega}}\mathbf{y} \succeq \text{vec}(\mathbf{R}) \odot \text{vec}(\mathbf{\Gamma}) \right\}, \quad (9)$$

which we refer to as the one-bit polyhedron. Generally, it can be assumed that the signal  $\mathbf{x} \in \mathbb{R}^d$  is observed linearly through the sampling matrix  $\mathbf{A} \in \mathbb{R}^{n \times d}$ , creating the measurements as  $\mathbf{y} = \mathbf{A}\mathbf{x}$ . Based on (??), the one-bit polyhedron for this type of problem is given by

$$\mathcal{P}_{\mathbf{x}} = \left\{ \mathbf{x} \mid \mathbf{P}\mathbf{x} \succeq \text{vec}(\mathbf{R}) \odot \text{vec}(\mathbf{\Gamma}) \right\}, \quad (10)$$

where  $\mathbf{P} = \tilde{\mathbf{\Omega}}\mathbf{A}$  or equivalently

$$\mathbf{P} = \left[ \mathbf{A}^\top \mathbf{\Omega}^{(1)} \mid \dots \mid \mathbf{A}^\top \mathbf{\Omega}^{(m)} \right]^\top, \quad \mathbf{P} \in \mathbb{R}^{mn \times d}. \quad (11)$$

By taking advantage of one-bit sampling in the asymptotic sample size scenario (sample abundance), the space constrained by the one-bit polyhedron  $\mathcal{P}_{\mathbf{x}}$  shrinks to become contained within the feasible set  $\Omega_c$ . Note that this shrinking space always contains the global minima, with a volume that is diminished with an increased sample size. To find a solution inside the one-bit polyhedron, the ORKA employs the variants of RKA introduced in the following subsections.

## B. Randomized Kaczmarz Algorithm

The RKA is a sub-conjugate gradient method to solve a linear feasibility problem, i.e,  $\mathbf{C}\mathbf{x} \succeq \mathbf{b}$  where  $\mathbf{C}$  is a  $m \times n$  matrix with  $m > n$  [?, ?]. Conjugate-gradient methods immediately turn the mentioned inequality to an equality in the following form  $(\mathbf{b} - \mathbf{C}\mathbf{x})^+ = 0$ , and then, approach the solution by the same process as used for systems of equations. The projection coefficient  $\beta_i$  of the RKA is [?, ?]

$$\beta_i = \begin{cases} (b_j - \mathbf{c}_j \mathbf{x}_i)^+ & (j \in \mathcal{I}_{\geq}), \\ b_j - \mathbf{c}_j \mathbf{x}_i & (j \in \mathcal{I}_{=}). \end{cases} \quad (12)$$

where the disjoint index sets  $\mathcal{I}_{\geq}$  and  $\mathcal{I}_{=}$  partition  $\mathcal{J}$  and  $\{\mathbf{c}_j\}$  are the rows of  $\mathbf{C}$ . Also, the unknown column vector  $\mathbf{x}$  is iteratively updated as

$$\mathbf{x}_{i+1} = \mathbf{x}_i + \frac{\beta_i}{\|\mathbf{c}_j\|_2^2} \mathbf{c}_j^H, \quad (13)$$



where, at each iteration  $i$ , the index  $j$  is drawn from the set  $\mathcal{J}$  independently at random following the distribution  $\Pr\{j = k\} = \frac{\|\mathbf{c}_k\|_2^2}{\|\mathbf{C}\|_F^2}$ . Assuming that the linear system is consistent with nonempty feasible set  $\mathcal{P}_{\mathbf{x}}$  created by the intersection of hyperplanes around the desired point  $\mathbf{x}_*$ , RKA converges linearly in expectation to the solution  $\hat{\mathbf{x}} \in \mathcal{P}_{\mathbf{x}}$  [?, ?]:

$$\mathbb{E} \{h(\mathbf{x}_i, \hat{\mathbf{x}})\} \leq (1 - q_{\text{RKA}})^i h(\mathbf{x}_0, \hat{\mathbf{x}}), \quad (14)$$

where  $h(\mathbf{x}_i, \hat{\mathbf{x}}) = \|\mathbf{x}_i - \hat{\mathbf{x}}\|_2^2$ , is the euclidean distance between two points in the space,  $i$  is the number of required iterations for RKA, and  $q_{\text{RKA}} \in (0, 1)$  is given by  $q_{\text{RKA}} = \frac{1}{\kappa^2(\mathbf{C})}$ , with  $\kappa(\mathbf{C}) = \|\mathbf{C}\|_F \|\mathbf{C}^\dagger\|_2$  denoting the scaled condition number of a matrix  $\mathbf{C}$ .

### C. Sampling Kaczmarz Motzkin Method

The SKM combines the ideas of both the RKA and the Motzkin method. The generalized convergence analysis of the SKM with sketch matrix which has been formulated based on the convergence analysis of RKA, and sampling Motzkin method for solving linear feasibility problem has been comprehensively explored in [?]. The central contribution of SKM lies in its innovative way of projection plane selection. The hyperplane selection is done as follows: At iteration  $i$ , the SKM algorithm selects a collection of  $\gamma$  (denoted by the set  $\mathcal{T}_i$ ) rows, uniformly at random out of  $m$  rows of the constraint matrix  $\mathbf{C}$ . Then, out of these  $\gamma$  rows, the row  $j^*$  with the maximum positive residual is selected; i.e.

$$j^* = \operatorname{argmax} \{ (b_j - \mathbf{c}_j \mathbf{x}_i)^+ \}, \quad j \in \mathcal{T}_i. \quad (15)$$

Finally, the solution is updated as [?, ?]  $\mathbf{x}_{i+1} = \mathbf{x}_i + \lambda_i \frac{\beta_i}{\|\mathbf{c}_{j^*}\|_2^2} \mathbf{c}_{j^*}^H$ , where  $\lambda_i$  is a relaxation parameter which for consistent systems must satisfy  $0 \leq \lim_{i \rightarrow \infty} \inf \lambda_i \leq \lim_{i \rightarrow \infty} \sup \lambda_i < 2$  [?], to ensure convergence.

### D. Preconditioned SKM

Assume  $\mathcal{P}_{\mathbf{x}}$  as the space created by the intersection of hyperplanes in a linear feasibility problem. According to the convergence rate of RKA, reducing the value of the scaled condition number  $\kappa(\mathbf{C})$  or equivalently increasing the value of  $q_{\text{RKA}}$  in (??) leads to an accelerated convergence of the RKA to  $\mathcal{P}_{\mathbf{x}}$ . As a result, the upper bound of the recovery error  $\mathbb{E} \{ \|\mathbf{x}_i - \hat{\mathbf{x}}\|_2^2 \}$  with  $\hat{\mathbf{x}} \in \mathcal{P}_{\mathbf{x}}$ , decreases, as well. From another perspective, this property (lower value of  $\kappa(\mathbf{C})$ ) provides the RKA or its variant, SKM, enjoying a lower number of

iterations required to reach a specific recovery error bound, usually considered to be the algorithm's termination criterion. Consequently, assuming  $I$  as a number of iterations, the computational cost of RKA which behaves as  $\mathcal{O}(In)$ , is diminished as well. The following theorem states how one can achieve the infimum of the scaled condition number  $\kappa(\mathbf{C})$ .

Theorem 1. The infimum scaled condition number of a matrix  $\mathbf{C} \in \mathbb{R}^{m \times n}$  is given by

$$\inf_{\mathbf{C}} \kappa(\mathbf{C}) = \sqrt{n}, \quad (16)$$

which is achieved if and only if  $\mathbf{C}$  is of the form  $\mathbf{C} = \alpha \mathbf{U}$ , where  $\mathbf{U}$  is an orthonormal matrix and  $\alpha \in \mathbb{R}$  is a scalar.

Proof. The condition number of the matrix  $\mathbf{C}$  is defined as  $\varrho(\mathbf{C}) = \frac{\sigma_{\max}}{\sigma_{\min}}$ , where  $\sigma_{\max}$  and  $\sigma_{\min}$  are the maximum and minimum singular values of the matrix  $\mathbf{C}$ , respectively [?]. The scaled condition number can be written as  $\kappa(\mathbf{C}) = \frac{\|\mathbf{C}\|_F}{\sigma_{\min}}$ . Therefore, the scaled condition number  $\kappa(\mathbf{C})$  has the following relation with the condition number  $\varrho(\mathbf{C})$ ,

$$\kappa(\mathbf{C}) = \frac{\|\mathbf{C}\|_F}{\sigma_{\max}} \varrho(\mathbf{C}). \quad (17)$$

Based on the relation between the norm-2 and the Frobenius norm of a matrix  $\mathbf{C}$  [?],  $\|\mathbf{C}\|_F \leq \sqrt{n} \|\mathbf{C}\|_2$  or equivalently,  $\frac{\|\mathbf{C}\|_F}{\sigma_{\min}} \leq \sqrt{n} \frac{\|\mathbf{C}\|_2}{\sigma_{\min}}$ , the condition number  $\varrho(\mathbf{C})$  can be considered to be an upper bound for the scaled condition number  $\kappa(\mathbf{C})$  as follows,  $\kappa(\mathbf{C}) \leq \sqrt{n} \varrho(\mathbf{C})$ . Thus, lowering  $\varrho(\mathbf{C})$  generally decreases the scaled condition number  $\kappa(\mathbf{C})$ . Additionally, the lowest possible value of  $\varrho$  is 1 which is achieved for scaled unitary matrices  $\mathbf{U}$  as if we let  $\mathbf{S} = \alpha \mathbf{U}$ , and  $\mathbf{O} = \mathbf{S}^\top \mathbf{S} = \alpha^2 \mathbf{I}_n$ , then,  $\sigma_{\min} = \sigma_{\max} = \alpha$ , and  $\varrho = 1$ . Vice versa, if  $\varrho = 1$ , it means  $\sigma_{\min} = \sigma_{\max}$  which leads to a diagonal matrix  $\mathbf{O} = \alpha^2 \mathbf{I}_n$ . It is straightforward to verify that the decomposition of  $\mathbf{O}$  results in  $\mathbf{S}$  which is a scaled-version of an orthonormal matrix. As a result, the lowest achievable upper bound for the scaled condition number  $\kappa(\mathbf{C})$  is obtained as  $\kappa(\mathbf{C}) \leq \sqrt{n}$ , and according to (??),  $\kappa(\mathbf{C}) = \frac{\alpha \|\mathbf{U}\|_F}{\alpha} = \sqrt{n}$ .  $\square$

Following Theorem ??, it would be enough to make the matrix  $\mathbf{C}$  unitary by a process which is referred to as the preconditioning method. In preconditioning, the linear feasibility is rewritten as  $\mathbf{CMz} \succeq \mathbf{b}$ , where  $\mathbf{M}$  is the preconditioner and  $\mathbf{x} = \mathbf{Mz}$ . The straightforward way to approach this task is to use QR decomposition where the constraint matrix is decomposed as  $\mathbf{C} = \mathbf{Q}_c \mathbf{R}_c$ , where  $\mathbf{Q}_c \in \mathbb{R}^{m \times n}$  is a unitary matrix, and  $\mathbf{R}_c \in \mathbb{R}^{n \times n}$  is an

---

**Algorithm 1** PrSKM Algorithm
 

---

Input: A matrix  $\mathbf{C}$  and the measurement vector  $\mathbf{b}$ .

Output: A solution  $\bar{\mathbf{x}}$  in a nonempty feasible set of  $\mathbf{C}\mathbf{x} \succeq \mathbf{b}$ .

- 1:  $[\mathbf{Q}_c, \mathbf{R}_c] = \text{QR}(\mathbf{C}) \triangleright \text{QR}(\cdot)$  computes the QR-decomposition of a matrix  $\mathbf{C}$ .
  - 2:  $\mathbf{Q}_c \mathbf{z} \succeq \mathbf{b} \triangleright$  The new problem that we should solve respect to  $\mathbf{z}$ .
  - 3: Choose a sample set of  $\gamma$  constraints (denoted as  $\mathcal{T}_i$ ) uniformly at random from the rows of  $\mathbf{Q}_c$ .
  - 4:  $j^* \leftarrow \arg\max \{ (b_j - \mathbf{q}_j \mathbf{z}_i)^+ \}, j \in \mathcal{T}_i \triangleright \mathbf{q}_j$  is the  $j$ -th row of  $\mathbf{Q}_c$ .
  - 5:  $\mathbf{z}_{i+1} \leftarrow \mathbf{z}_i + \lambda_i \frac{(b_{j^*} - \mathbf{q}_{j^*} \mathbf{z}_i)^+}{\|\mathbf{q}_{j^*}\|_2^2} \mathbf{q}_{j^*}^H$ .
  - 6: Repeat steps (4)-(6) until convergence and obtain  $\bar{\mathbf{z}}$ .
  - 7:  $\mathbf{R}_c \bar{\mathbf{x}} = \bar{\mathbf{z}} \triangleright$  Obtain  $\bar{\mathbf{x}}$  via the Gaussian elimination algorithm.
  - 8: return  $\bar{\mathbf{x}}$
- 

upper triangular matrix, leading to<sup>1</sup>  $\mathbf{Q}_c = \mathbf{C}\mathbf{R}_c^{-1}$ . Thus, based on Theorem ??, a good choice for the preconditioner is  $\mathbf{M} = \mathbf{R}_c^{-1}$ . To find a point  $\bar{\mathbf{z}}$  in a nonempty feasible set, the SKM method described in Section ?? is employed. Finally, one may approach the solution of the original linear feasibility  $\mathbf{C}\mathbf{x} \succeq \mathbf{b}$  by computing  $\bar{\mathbf{x}} = \mathbf{R}_c^{-1}\bar{\mathbf{z}}$ . We refer to this method Preconditioned SKM (PrSKM) which is summarized in Algorithm ?. As shown in Theorem ??, the scaled condition number of the matrix  $\mathbf{Q}_c$  is  $\kappa(\mathbf{Q}_c) = \sqrt{n}$ . Therefore, step 5 of Algorithm ?? converges linearly in expectation to a nonempty feasible set of  $\mathbf{Q}_c \mathbf{z} \succeq \mathbf{b}$  as follows,

$$\mathbb{E} \{ \hbar(\mathbf{x}_i, \hat{\mathbf{x}}) \} \leq \left( 1 - \frac{1}{n} \right)^i \hbar(\mathbf{x}_0, \hat{\mathbf{x}}). \quad (18)$$

Note that to run the Algorithm ?? for solving the one-bit polyhedron (??), it is enough to set  $\mathbf{C} = \mathbf{P}$  and  $\mathbf{b} = \text{vec}(\mathbf{R}) \odot \text{vec}(\mathbf{\Gamma})$ . Later, in Theorem ??, we will show that it is only required to set  $\mathbf{C} = \mathbf{A}$  in Algorithm ?? which is more computationally and storagelley efficient compared to that of setting  $\mathbf{C} = \mathbf{P}$ .

<sup>1</sup>For a matrix  $\mathbf{C} \in \mathbb{R}^{m \times n}$ , since we have assumed  $m > n$ , we can obtain a unitary matrix  $\mathbf{Q}_c \in \mathbb{R}^{m \times n}$  and an upper triangular matrix  $\mathbf{R}_c \in \mathbb{R}^{n \times n}$  such that the QR-decomposition holds; i.e.  $\mathbf{C} = \mathbf{Q}_c \mathbf{R}_c$ .

### E. Storage-Friendly PrSKM

The PrSKM algorithm uses QR decomposition as a preconditioning process, which can be computationally challenging for high-dimensional matrices. To address this issue, the literature has proposed a technique called sketch-and-precondition, as outlined in [?]. This approach is designed to mitigate the curse of dimensionality associated with QR-decomposition-based preconditioning.

To explain how this method works, consider for simplicity an overdetermined linear system, i.e.,  $\mathbf{U}\mathbf{x} = \mathbf{b}$  with  $\mathbf{U} \in \mathbb{R}^{m \times n}$ ,  $m \gg n$ . Define a Gaussian matrix  $\mathbf{N} \in \mathbb{R}^{m \times s}$ . Sketch-and-precondition computes a sketch  $\mathbf{N}^\top \mathbf{U}$  using a random test matrix  $\mathbf{N}$  with sketch size  $s = \mathcal{O}(n/\epsilon^2) \ll m$  where  $\epsilon \in (0, 1)$ . Next, it performs a QR-decomposition  $\mathbf{N}^\top \mathbf{U} = \mathbf{Q}_p \mathbf{R}_p$  of the resulting  $s \times n$  matrix and then uses  $\mathbf{R}_p^{-1}$  as a preconditioner for  $\mathbf{U}$ . This approach significantly reduces the computational cost of QR-decomposition from  $\mathcal{O}(mn^2)$  to  $\mathcal{O}(sn^2)$ , regardless of the number of rows. This reduction in computational cost is particularly advantageous for highly-overdetermined linear systems, demonstrating the significance of this approach. For more information, we encourage a reader to see [?, ?]. This paper introduces the first instance of utilizing the sketch-and-precondition technique on a linear inequality system within the literature. To explicitly state the algorithmic implementation of storage-friendly PrSKM, the first two steps of Algorithm ?? must be replaced by the following steps:

- 1) Generate a Gaussian test matrix  $\mathbf{N} \in \mathbb{R}^{m \times s}$  where  $s$  follows  $\mathcal{O}(n/\epsilon^2)$  with  $\epsilon \in (0, 1)$ , sketch the matrix  $\mathbf{C} \in \mathbb{R}^{m \times n}$  from  $\mathbf{C}\mathbf{x} \succeq \mathbf{b}$  as  $\mathbf{S} = \mathbf{N}^\top \mathbf{C}$ .
- 2)  $[\mathbf{Q}_s, \mathbf{R}_s] = \text{QR}(\mathbf{S}) \triangleright \text{QR}(\cdot)$  computes the QR-decomposition of the sketched matrix  $\mathbf{S}$ , where  $\mathbf{Q}_s \in \mathbb{R}^{s \times n}$  and  $\mathbf{R}_s \in \mathbb{R}^{n \times n}$ .
- 3)  $\mathbf{M} = \mathbf{R}_s^{-1} \triangleright \mathbf{M}$  is the preconditioner.
- 4)  $\mathbf{CM}\mathbf{z} \succeq \mathbf{b} \triangleright$  New problem that we should solve respect to  $\mathbf{z}$ .

We provide the convergence of storage-friendly PrSKM to the nonempty feasible set of  $\mathbf{CM}\mathbf{z} \succeq \mathbf{b}$  in the following theorem:

**Theorem 2.** Suppose  $\mathbf{N} \in \mathbb{R}^{m \times s}$  is a Gaussian matrix with the sketch size  $s = \mathcal{O}(n/\epsilon^2)$  with  $\epsilon = \frac{1}{2}$ . Consider the sketch matrix as  $\mathbf{S} = \mathbf{N}^\top \mathbf{C}$  where  $\mathbf{C} \in \mathbb{R}^{m \times n}$ ,  $m \gg n$ . Assume  $[\mathbf{Q}_s, \mathbf{R}_s] = \text{QR}(\mathbf{S})$  where  $\text{QR}(\cdot)$  denotes the QR-decomposition operator. Then, the storage-friendly PrSKM algorithm converges linearly in expectation to the nonempty feasible set

of  $\mathbf{C}\mathbf{R}_s^{-1}\mathbf{z} \succeq \mathbf{b}$  as follows:

$$\mathbb{E} \{h(\mathbf{x}_i, \hat{\mathbf{x}})\} \leq \left(1 - \frac{1}{3n}\right)^i h(\mathbf{x}_0, \hat{\mathbf{x}}). \quad (19)$$

The proof of Theorem ?? is presented in Appendix ??.

## F. Block SKM

In two previous sections, we have proposed PrSKM and storage-friendly PrSKM algorithms in order to solve the one-bit polyhedron (??) in an asymptotic sample size scenario. It is worth noting that both methods are row-based approaches, where at each iteration, the row index of the matrix  $\mathbf{P}$  is chosen independently at random. However, the matrix  $\mathbf{P}$  in (??) has a block structure as formulated in (??). This fact motivates us to investigate the block-based RKA methods to find the desired signal in the one-bit polyhedron  $\mathcal{P}_{\mathbf{x}}$  for further efficiency enhancement. Our proposed algorithm for block systems, Block SKM, is motivated by (i) random selection of one block at each iteration, (ii) choosing a subset of rows using the idea of Motzkin sampling, and (iii) updating the solution using the randomized block Kaczmarz method [?, ?], which takes advantage of the efficient matrix-vector multiplication, thus giving the method a significant reduction in computational cost [?]. Algorithm ?? shows the implementation of our proposed Block SKM method to solve the linear feasibility  $\mathbf{B}\mathbf{x} \succeq \mathbf{b}$  with  $\mathbf{B} = \left[ \mathbf{B}_1^\top \mid \dots \mid \mathbf{B}_m^\top \right]^\top$  and  $\mathbf{b} = \left[ \mathbf{b}_1^\top \mid \dots \mid \mathbf{b}_m^\top \right]^\top$  where  $\mathbf{B}_j \in \mathbb{R}^{n \times d}$  and  $\mathbf{b}_j \in \mathbb{R}^n$  for all  $j \in \{1, \dots, m\}$ . Note that in step 4 of Algorithm ??, the reason behind choosing  $k' < d$  is due to the computation of  $(\mathbf{B}'_j \mathbf{B}'_j{}^\top)^{-1}$  in the next step (step 5). For  $k' > d$ , the matrix  $\mathbf{B}'_j \mathbf{B}'_j{}^\top$  is rank-deficient and its inverse is not available. The Block SKM algorithm can be considered to be a special case of the more general sketch-and-project method with a sparse block sketch matrix as defined in [?]. The convergence of Block SKM algorithm with the sparse Gaussian sketch in the case of  $k' = 1$  is presented in the following lemma:

**Lemma 1.** The Block SKM algorithm with the sparse Gaussian sketch in the case of  $k' = 1$ , converges linearly in expectation to the nonempty feasible set of  $\mathbf{B}\mathbf{x} \succeq \mathbf{b}$ ,  $\mathbf{B} \in \mathbb{R}^{mn \times d}$ , as follows,

$$\mathbb{E} \{h(\mathbf{x}_i, \hat{\mathbf{x}})\} \leq \left(1 - \frac{c\sigma_{\min}^2(\hat{\mathbf{B}}) \log n}{\|\hat{\mathbf{B}}\|_{\text{F}}^2}\right)^K h(\mathbf{x}_0, \hat{\mathbf{x}}), \quad (20)$$

---

**Algorithm 2** Block SKM Algorithm
 

---

- Input: A block matrix  $\mathbf{B} = \left[ \mathbf{B}_1^\top \mid \cdots \mid \mathbf{B}_m^\top \right]^\top$ , and the measurement vector  $\mathbf{b} = \left[ \mathbf{b}_1^\top \mid \cdots \mid \mathbf{b}_m^\top \right]^\top$ .
- Output: A solution  $\bar{\mathbf{x}}$  in a nonempty feasible set of  $\mathbf{B}\mathbf{x} \succeq \mathbf{b}$ .
- 1: Choose a block  $\mathbf{B}_j$  uniformly at random with the probability  $\Pr\{j = k\} = \frac{\|\mathbf{B}_k\|_F^2}{\|\mathbf{B}\|_F^2}$ .
  - 2:  $\mathbf{e} \leftarrow \mathbf{B}_j\mathbf{x} - \mathbf{b}_j$ .
  - 3:  $\mathbf{e}' \leftarrow \text{sort}(\mathbf{e}) \triangleright \text{sort}(\cdot)$  is the operator that sorts the elements of the vector  $\mathbf{e}$  from  $e_{\max}$  (the maximum element of  $\mathbf{e}$ ) to  $e_{\min}$  (the minimum element of  $\mathbf{e}$ ). This step is inspired by the idea of the Motzkin sampling, presented in Section ??, to have an accelerated convergence.
  - 4:  $[\mathbf{B}'_j, \mathbf{b}'_j] = \text{select}_{k'}(\mathbf{B}_j, \mathbf{b}_j, \mathbf{e}') \triangleright \text{select}_{k'}(\cdot)$  is the operator which selects the first  $k' < d$  element of  $\mathbf{e}'$  and construct the sub-problem  $\mathbf{B}'_j\mathbf{x} \succeq \mathbf{b}'_j$ , where  $\mathbf{B}'_j \in \mathbb{R}^{k' \times d}$  and  $\mathbf{b}'_j \in \mathbb{R}^{k'}$ .
  - 5:  $\mathbf{B}_j^{\dagger} \leftarrow \mathbf{B}_j^\top (\mathbf{B}_j \mathbf{B}_j^\top)^{-1} \triangleright \mathbf{B}_j^{\dagger}$  is the Moore-Penrose pseudoinverse of  $\mathbf{B}_j$ .
  - 6:  $\mathbf{x}_{i+1} \leftarrow \mathbf{x}_i + \lambda_i \mathbf{B}_j^{\dagger} (\mathbf{b}'_j - \mathbf{B}'_j\mathbf{x})^+$ .
  - 7: Repeat steps (1)-(6) until convergence and obtain  $\bar{\mathbf{x}}$ .
  - 8: return  $\bar{\mathbf{x}}$
- 

where  $c$  is a positive constant value,  $\hat{\mathbf{B}}$  is the  $n \times d$  submatrix of  $\mathbf{B}$  (one of the candidates of  $\mathbf{B}_j$  in Algorithm ??), and  $K$  is the number of update process for the linear inequalities.

The proof of Lemma ?? is provided in Appendix ??. In the rest of the paper, we will derive all theoretical guarantees for the RKA. Similar guarantees can also be derived for the variants of the RKA; i.e. SKM, Block SKM and etc. Throughout the paper, we represent each iterate of the update process for RKA, PrSKM, and Block SKM, aimed at solving the feasibility problem  $\mathbf{B}\mathbf{x} \succeq \mathbf{b}$ , as  $\text{KA}_r(\mathbf{x})$ ,  $\text{KA}_p(\mathbf{x})$ , and  $\text{KA}_b(\mathbf{x})$ , respectively. To examine the performance of the Block SKM, we will compare it with the PrSKM, SKM and RKA.

### G. Comparing RKA, SKM, PrSKM and Block SKM

In this section, we numerically compare the RKA, SKM, PrSKM, and Block SKM in linear system of inequalities. Accordingly, we utilize ORKA to make a linear equation  $\mathbf{B}\mathbf{x} = \mathbf{y}$  a linear inequality system, where the number of time-varying sampling threshold sequences is  $m = 40$ ,  $\mathbf{B} \in \mathbb{R}^{100 \times 10}$ ,  $\mathbf{x} \in \mathbb{R}^{10}$ , and  $\mathbf{y} \in \mathbb{R}^{100}$ . Each row of  $\mathbf{B}$  is generated as  $\mathbf{b}_j \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_{10})$ .

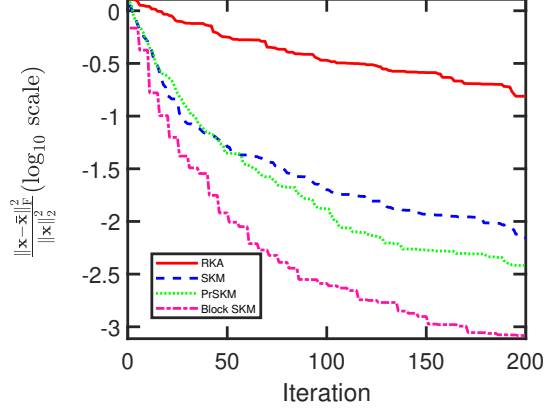


Figure 1. Comparing the recovery performance of the two proposed Kaczmarz algorithms, namely the PrSKM and the Block SKM, with that of SKM and RKA for a linear inequality system.

Also, the desired signal  $\mathbf{x}$  is generated as  $\mathbf{x} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_{10})$ . All time-varying sampling threshold sequences are generated according to  $\{\boldsymbol{\tau}^{(\ell)} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_{10})\}_{\ell=1}^m$ . The performance of the RKA, SKM, PrSKM, and Block SKM is illustrated in Fig. ?? . Similar to the linear feasibility of equalities, it can be seen that the Block SKM has a better accuracy in recovering the desired signal  $\mathbf{x}$  in the one-bit polyhedron (??) compared to the other three approaches. The NMSE results in Fig. ?? are averaged over 1000 experiments.

#### IV. Probabilistic Effect of Sample Abundance In One-Bit Sensing

In Section ??, we will introduce the concept of FVP and subsequently obtain the required number of one-bit samples  $m'$  to accurately capture the solution in the case of sample abundance, one-bit CS, and one-bit low-rank matrix recovery. In Section ??, we will provide the convergence of ORKA based on the theoretical results obtained in Section ??.

##### A. Finite Volume Property

Define the distance between the original signal  $\mathbf{x}_\star$  and the  $j$ -th hyperplane presented in (??) as

$$d_j^{(\ell)}(\mathbf{x}_\star, \tau_j^{(\ell)}) = \left| r_j^{(\ell)} \mathbf{a}_j \mathbf{x}_\star - r_j^{(\ell)} \tau_j^{(\ell)} \right|, \quad j \in [n], \quad \ell \in [m], \quad (21)$$

It is essential to clarify that in our analysis, we adopt the worst-case scenario for the distance between the desired point and the solutions. This approach considers the solution to lie precisely on the hyperplane. However, it is crucial to note that in reality, the solution to a

linear inequality system can exist anywhere between the desired point and the hyperplane. It is easy to observe that by reducing the distances between  $\mathbf{x}_\star$  and the constraint-associated hyperplanes in (??), the possibility of capturing the original signal  $\mathbf{x}_\star$  is increased. For a specific sample size  $m'$ , when  $\text{vol}(\mathcal{P}_\mathbf{x})$  is reduced, the average of  $\left\{d_j^{(\ell)}(\mathbf{x}_\star, \tau_j^{(\ell)})\right\}_{j,\ell=1}^{m'}$  is diminished as well. This average of distances can be written as:

$$T_{\text{ave}} = \frac{1}{m'} \sum_{j,\ell=1}^{m'} d_j^{(\ell)}(\mathbf{x}_\star, \tau_j^{(\ell)}). \quad (22)$$

The one-bit phase retrieval problem, as investigated in [?], derived a general Hoeffding's bound [?, Theorem 2.6.2] to quantify the likelihood of achieving a finite volume and determine the necessary number of samples for one-bit signal reconstruction. In Theorem ??, we utilize this result to address the problem of one-bit sensing. Specifically, we consider the distance between the original signal  $\mathbf{x}_\star$  and the  $j$ -th hyperplane within the polyhedron defined in (??), as described in (??).

**Theorem 3 (Finite volume property (FVP)).** Assume the distances  $\left\{d_j^{(\ell)}(\mathbf{x}_\star, \tau_j^{(\ell)})\right\}_{j,\ell=1}^{m'}$  defined in (??) between the desired point  $\mathbf{x}_\star$  and the hyperplanes of the one-bit polyhedron (??) are independent random variables with,  $\mathbb{E}\{T_{\text{ave}}\} = \mu$  and  $\left\|d_j^{(\ell)}(\mathbf{x}_\star, \tau_j^{(\ell)})\right\|_{q_2}^2 \leq K$ . Then, based on the general Hoeffding's inequality, the probability of the finite volume created by hyperplanes lying within a ball  $\mathcal{B}_\rho(\mathbf{x}_\star)$  centered at the original signal, with a radius of  $\rho$ , is bounded by:

$$\Pr(T_{\text{ave}} \geq C\rho) \leq e^{\frac{-c_1(C\rho-\mu)^2}{K}m'}, \quad (23)$$

where  $C$  and  $c_1$  are positive numbers.

Theorem ?? provides the probability of the finite volume, created by hyperplanes, being contained within the ball around the original signal. The positive number  $C$  serves to consider the distances beyond the radius  $\rho$  of the ball. These distances correspond to ineffective hyperplanes that are incapable of forming a finite volume around the original signal. Mathematically, based on the general Hoeffding's inequality presented in (??), it follows  $\delta = C\rho - \mu$  for a positive constant  $\delta$ . We utilize this constant,  $\delta$ , in all the theoretical guarantees that will be obtained in this paper. In the remainder of the paper, our objective is to derive theoretical guarantees that enable us to achieve the uniform perfect reconstruction, as defined in the following definition:



Definition 1 (Uniform perfect reconstruction). Uniform perfect reconstruction of a signal from one-bit polyhedron is achieved when all possible recovery solutions for  $\bar{\mathbf{x}} \in \mathcal{P}_{\mathbf{x}}$  with

$$\text{sgn}(\mathbf{a}_j \mathbf{x}_\star - \tau_j^{(\ell)}) = \text{sgn}(\mathbf{a}_j \bar{\mathbf{x}} - \tau_j^{(\ell)}), \quad j \in [n], \quad \ell \in [m], \quad (24)$$

satisfy  $\bar{\mathbf{x}} \in \mathcal{B}_\rho(\mathbf{x}_\star)$ , for all  $\mathbf{x}_\star$  in the space.

Note that the sign preservation mentioned in our work, as stated in Definition 1, is equivalent to the solution provided by a linear feasibility solver, denoted as  $\bar{\mathbf{x}}$ , satisfying all the given inequalities, meaning  $\bar{\mathbf{x}} \in \mathcal{P}_{\mathbf{x}}$ . To fulfill Definition 1, it appears sufficient to acquire the number of samples that ensures the creation of a finite volume of intersections of hyperplanes with the maximum radius  $\rho$ . This central idea underlies all of our theorems. In the following theorem, we establish the minimum number of one-bit samples  $m'$  required in the sample abundance scenario to achieve an accurate recovery:

Theorem 4. Assume a  $n \times d$  sampling matrix  $\mathbf{A}$  such that each distance defined in (??) is a sub-Gaussian random variable, and the one-bit sampling matrix  $\mathbf{P} = \tilde{\mathbf{\Omega}}\mathbf{A} \in \mathbb{R}^{m' \times d}$ , where  $m' = mn$  denotes the total number of one-bit samples, and  $m$  represents the number of time-varying sampling threshold sequences. Denote  $\delta$  and  $C_1$  as positive constants. If the number of one-bit samples obeys

$$m' \geq C_1 \delta^{-2} \left( \frac{3d}{\rho} + \log \left( \frac{1}{\eta} \right) \right), \quad (25)$$

then with a probability of at least  $1 - \eta$ , we achieve the uniform perfect reconstruction with  $\bar{\mathbf{x}} \in \mathcal{B}_\rho(\mathbf{x}_\star)$ .

The proof of Theorem ?? is presented in Appendix ??. Following the discussion provided in [?] regarding the benefit of uniform dithering in the case of bounded measurements, the Corollary ?? presents the result of Theorem ?? in the case of the DCT sensing matrix  $\mathbf{A}$  and uniform dither:

Corollary 1. Consider a  $n \times d$  DCT matrix  $\mathbf{A}$ , and the time-varying thresholds which are generated according to  $\left\{ \tau_j^{(\ell)} \sim \mathcal{U}(-\tilde{b}, \tilde{b}) \right\}_{j,\ell=1}^{m'}$  such that  $\tilde{b} > 0$ . Denote  $f_j$  as the  $j$ -th DCT coefficient of  $\mathbf{x}_\star$ . If we have

$$m' \geq \frac{1}{2} \delta^{-2} \left( \tilde{b} + \sqrt{2} \right)^2 \left( \frac{3d}{\rho} + \log \left( \frac{1}{\eta} \right) \right), \quad (26)$$

then with a probability of at least  $1 - \eta$ , we achieve the uniform perfect reconstruction with  $\bar{\mathbf{x}} \in \mathcal{B}_\rho(\mathbf{x}_\star)$ , and  $\delta = C\rho - \mu$  for a positive value  $C$  with

$$\begin{aligned} \mu = \frac{1}{2\tilde{b}n} \sum_{j=1}^n I(-\tilde{b} \leq f_j \leq \tilde{b}) (f_j^2 + \tilde{b}^2) \\ + I(f_j \geq \tilde{b})(2\tilde{b}f_j) + I(f_j \leq -\tilde{b})(-2\tilde{b}f_j). \end{aligned} \quad (27)$$

The proof of Corollary ?? is presented in Appendix ?. In the following corollary, we present the result of Theorem ?? for the Gaussian sampling matrix with Gaussian dithering scheme:

Corollary 2. Consider a  $n \times d$  sampling matrix  $\mathbf{A} = [a_{ij}]$  with  $a_{ij}$  independently drawn from the standard Gaussian distribution,  $\mathcal{N}(0, 1)$ , and the time-varying thresholds which are generated according to  $\left\{ \tau_j^{(\ell)} \sim \mathcal{N}(0, \sigma^2) \right\}_{j,\ell=1}^{m'}$ . If we have

$$m' \geq C_1 \delta^{-2} (\sigma^2 + 1) \left( \frac{3d}{\rho} + \log \left( \frac{1}{\eta} \right) \right), \quad (28)$$

then with a probability of at least  $1 - \eta$ , we achieve the uniform perfect reconstruction with  $\bar{\mathbf{x}} \in \mathcal{B}_\rho(\mathbf{x}_\star)$ , and  $\delta = C\rho - \mu$  for a positive value  $C$  with  $\mu \leq \sqrt{\frac{2}{\pi}(\sigma^2 + 1)}$ .

The proof of Corollary ?? is presented in Appendix ?. It is important to note that in the absence of dithering, creating a finite volume around the desired solution becomes impossible, rendering the FVP ineffective. Furthermore, in such scenarios, certain information, such as the signal amplitude, may be lost. However, in specific circumstances where the signal is  $s$ -sparse and confined within a unit ball, some sampling theorems have been derived in the literature [?, Theorem 2].

The existing literature provides numerous theoretical derivations for the case where the input signal possesses an  $s$ -sparse structure, primarily focusing on the zero threshold scenario (e.g., [?, Theorem 2]). However, in order to determine the minimum number of samples required for achieving perfect reconstruction of sparse signals, we need to introduce the distance within the  $s$ -dimensional space. Let us consider an  $s$ -support column-submatrix of the sampling matrix, denoted as  $\mathbf{A}_s$ . The rows of this submatrix are represented by  $\mathbf{a}_j^{(s)}$ . Additionally, we have an  $s$ -subvector of  $\mathbf{x}_\star$ , denoted as  $\mathbf{x}_\star^{(s)}$ . The distances between the  $s$ -sparse original signal and the hyperplanes are given by

$$d_j^{(\ell)s}(\mathbf{x}_\star^{(s)}, \tau_j^{(\ell)}) = \left| r_j^{(\ell)} \mathbf{a}_j^{(s)} \mathbf{x}_\star^{(s)} - r_j^{(\ell)} \tau_j^{(\ell)} \right|, \quad s \in [d]. \quad (29)$$

It is important to note that for a  $s$ -sparse signal, we have a total of  $\binom{d}{s}$  possible choices for the distances. In other words, in the case of a  $s$ -support problem, we can select  $\binom{d}{s}$  different submatrices  $\mathbf{A}_s$  from the sampling matrix. By utilizing the general Hoeffding's inequality along with the union bound for these  $\binom{d}{s}$  possibilities, the following theorem provides the necessary number of one-bit samples to achieve uniform perfect reconstruction of a  $s$ -sparse signal:

**Theorem 5.** Assume a  $n \times d$  sampling matrix  $\mathbf{A}$  such that each distance defined in (??) is a sub-Gaussian random variable. Assume that the desired signal is  $s$ -sparse and  $\rho$  is the radius of a ball centered at the original signal  $\mathbf{x}_\star \in \mathbb{R}^d$ . Consider a fixed  $0 < \eta < 1$  and  $\delta_s$  to be a positive constant value. Then the minimum number of one-bit samples  $m'$  required to accurately capture a  $s$ -sparse solution  $\bar{\mathbf{x}}$ , satisfies

$$m' \geq \delta_s \left( \log \left( \frac{1}{\eta} \right) + s \left( \log \left( \frac{ed}{s} \right) + \frac{3}{\rho} \right) \right), \quad (30)$$

with a probability higher than  $1 - \eta$ .

The proof of Theorem ?? is provided in Appendix ??. In the following corollary, we provide the result of Theorem ?? in the case of Gaussian sensing matrix with a Gaussian dithering scheme:

**Corollary 3.** Assume a  $n \times d$  sampling matrix  $\mathbf{A} = [a_{ij}]$  with  $a_{ij}$  independently drawn from the standard Gaussian distribution,  $\mathcal{N}(0, 1)$ , and the time-varying thresholds which are generated according to  $\left\{ \tau_j^{(\ell)} \sim \mathcal{N}(0, \sigma^2) \right\}_{j,\ell=1}^{m'}$ . Define the set  $\mathcal{H}_{d,s}$  as

$$\mathcal{H}_{d,s} = \left\{ \mathbf{x} \in \mathbb{R}^d \mid \|\mathbf{x}\|_1 \leq \sqrt{s}, \|\mathbf{x}\|_2 \leq 1 \right\}, \quad (31)$$

with  $s \ll d$ . Denote  $\delta = C\rho - \sqrt{\frac{2}{\pi}(\sigma^2 + 1)}$  for a positive value  $C$  and  $0 < \rho < 1$ . For a positive constant  $C_1$ , if  $m' \geq C_1 \delta^{-2} (\sigma^2 + 1) \left( \log \left( \frac{1}{\eta} \right) + s \left( \log \left( \frac{ed}{s} \right) + \frac{3}{\rho} \right) \right)$ , then with probability of at least  $1 - \eta$ , and for all  $\mathbf{x}_\star, \bar{\mathbf{x}} \in \mathcal{H}_{d,s}$  we achieve the uniform perfect reconstruction with  $\bar{\mathbf{x}} \in \mathcal{B}_\rho(\mathbf{x}_\star)$ .

The proof of Corollary ?? follows directly from the proof of Theorem ?? and the proof of Corollary ??. To investigate the necessary number of one-bit samples for  $r$ -rank matrices, one can identify a value  $k \in [K]$  from a set of  $K$  possibilities, such that  $\mathbf{X}_k$  lies within a ball centered at the original signal with a radius of  $\rho$ . Assume  $n_1 = n_2$ , previous studies, such

as [?, Lemma 3.1], have demonstrated that  $K$  satisfies the inequality  $K \leq (1 + \frac{6}{\rho})^{(2n_1+1)r}$ . By combining the general Hoeffding's inequality with the union bound applied to all  $K$  possibilities, we can establish the following theorem:

**Theorem 6.** Assume a  $n \times n_1^2$  sampling matrix  $\mathbf{A}$  such that each distance defined in (??) is a sub-Gaussian random variable. Assume that the desired signal is  $r$ -rank matrix and  $\rho$  is the radius of a ball centered at the original signal  $\mathbf{X}_\star \in \mathbb{R}^{n_1 \times n_1}$ . Consider a fixed  $0 < \eta < 1$  and  $\delta_r$  to be a positive constant value. Then the minimum number of one-bit samples  $m'$  required to precisely capture a  $r$ -rank solution, satisfies

$$m' \geq \delta_r \left( \frac{18n_1r}{\rho} + \log \left( \frac{1}{\eta} \right) \right), \quad (32)$$

with a probability higher than  $1 - \eta$ .

The proof of Theorem ?? is provided in Appendix ??. In the following corollary, we provide the result of Theorem ?? in the case of Gaussian sensing matrix with a Gaussian dithering scheme:

**Corollary 4.** Assume a  $n \times n_1^2$  sampling matrix  $\mathbf{A} = [a_{ij}]$  with  $a_{ij}$  independently drawn from the standard Gaussian distribution,  $\mathcal{N}(0, 1)$ , and the time-varying thresholds which are generated according to  $\left\{ \tau_j^{(\ell)} \sim \mathcal{N}(0, \sigma^2) \right\}_{j,\ell=1}^{m'}$ . Define the set  $\mathcal{K}_{n_1,r}$  as

$$\mathcal{K}_{n_1,r} = \left\{ \mathbf{X} \in \mathbb{R}^{n_1 \times n_1} \mid \text{rank}(\mathbf{X}) \leq r, \|\mathbf{X}\|_F \leq 1 \right\}. \quad (33)$$

Denote  $\delta = C\rho - \sqrt{\frac{2}{\pi}(\sigma^2 + 1)}$  for a positive value  $C$  and  $0 < \rho < 1$ . For a positive constant  $C_1$ , if  $m' \geq C_1\delta^{-2}(\sigma^2 + 1) \left( \frac{18n_1r}{\rho r} + \log \left( \frac{1}{\eta} \right) \right)$ , then with a probability of at least  $1 - \eta$ , and for all  $\mathbf{X}_\star, \bar{\mathbf{X}} \in \mathcal{K}_{n_1,r}$ , we achieve the uniform perfect reconstruction with  $\bar{\mathbf{X}} \in \mathcal{B}_\rho(\text{vec}(\mathbf{X}_\star))$ .

The proof of Corollary ?? follows directly from the proof of Theorem ?? and the proof of Corollary ??. As per the aforementioned theorems, the specific structure of the signal, such as sparsity or low-rank, has a direct impact on the probability of creating a finite volume around the original signal using one-bit hyperplanes. This structural characteristic manifests itself in the required number of samples for achieving accurate reconstruction. Note that Theorems ??, ??, and ?? present uniform reconstruction results, indicating that with high probability, all vectors can be reconstructed. This differs from a nonuniform result where each vector is individually reconstructed with high probability.

To have a robust recovery performance in one-bit signal sensing, it is necessary to design the time-varying thresholds such that the dynamic range (DR) of such thresholds covers the DR of the high-resolution measurements. In the following proposition, we focus our analysis to the sub-Gaussian sensing matrix with both Gaussian and Uniform dithering:

Lemma 2. Assume a  $n \times d$  sub-Gaussian sensing matrix  $\mathbf{A}$  with the rows  $\{\mathbf{a}_j\}_{j=1}^n$  and the time-varying thresholds which are generated according to  $\{\tau_j^g \sim \mathcal{N}(0, \sigma^2)\}_{j=1}^n$  and  $\{\tau_j^u \sim \mathcal{U}(-\tilde{b}, \tilde{b})\}_{j=1}^n$ . The high-resolution measurements are represented as  $\mathbf{y} = \mathbf{A}\mathbf{x}$ . Define  $K = \max_j \|\mathbf{a}_j\|_{\psi_2}$ . If  $\|\mathbf{x}\|_2 \leq R$ , then setting  $\sigma = \mathcal{O}\left(\sqrt{\frac{3}{8}}KR\right)$  in the Gaussian dither and  $\tilde{b} = \mathcal{O}(\log(2)KR)$  in the Uniform dither guarantee the cover of  $\text{DR}_{\mathbf{y}}$ .

The proof of Lemma ?? is presented in Appendix ??. In the case of the Gaussian sensing matrix  $\mathbf{A} = [a_{ij}]$  with  $a_{ij}$  independently drawn from the standard Gaussian distribution,  $\mathcal{N}(0, 1)$ , the result of Lemma ?? is simplified to  $\sigma = \mathcal{O}(R)$  in the Gaussian dither and  $\tilde{b} = \mathcal{O}\left(R \log(2) \sqrt{\frac{8}{3}}\right)$  in the Uniform dither. For the example provided in Corollary ?? regarding the DCT sampling matrix and uniform dithering, we should have  $\tilde{b} = \mathcal{O}(\sqrt{2})$  to cover the DR of high-resolution measurements (DCT coefficients). The reason behind this can be seen in the proof of Corollary ??.

## B. Recovery Error Upper Bound for ORKA

As the scaled condition number is the central parameter governing the recovery error of the RKA or its variants, in the following theorem we will evaluate the scaled condition number of the matrix  $\mathbf{P}$  defined in (??) to unveil the connection between the convergence bounds of the RKA (or its variants) and ORKA.

Theorem 7. Consider the one-bit polyhedron (??) associated with the linear system of equations  $\mathbf{A}\mathbf{x} = \mathbf{y}$  with  $\mathbf{A} \in \mathbb{R}^{n \times d}$  and  $m$  denotes the number of time-varying sampling threshold sequences. Then, the scaled condition number of the matrix  $\mathbf{P}$  is equal to that of  $\mathbf{A}$ ,

$$\kappa(\mathbf{P}) = \kappa(\mathbf{A}). \quad (34)$$

The proof of Theorem ?? is provided in Appendix ??. Considering  $\mathbf{A} = \mathbf{I}$  corresponds to the one-bit sampled signal sensing problem with  $\mathbf{P} = \tilde{\mathbf{\Omega}}$  as formulated in (??). In the

following corollary, we present the scaled condition number of the matrix  $\tilde{\Omega}$  in the light of Theorem ??:

Corollary 5. For  $\mathbf{A} = \mathbf{I}$ , corresponding to the one-bit sampled signal reconstruction problem formulated in (??), the scaled condition number of  $\tilde{\Omega} = \left[ \Omega^{(1)} \mid \dots \mid \Omega^{(m)} \right]^\top$ ,  $\tilde{\Omega} \in \mathbb{R}^{mn \times n}$  is  $\kappa(\Omega) = \sqrt{n}$ , which is the infimum of the scaled condition number as was shown in Theorem ??.

Based on Theorem ?? and the convergence bound of the RKA, it is concluded that ORKA converges to the feasible set of  $\mathbf{P}\mathbf{x} \succeq \text{vec}(\mathbf{R}) \odot \text{vec}(\mathbf{\Gamma})$  with the rate that depends on the sampling matrix  $\mathbf{A}$ . Therefore, the convergence bound of ORKA is independent of the number of time-varying sampling threshold sequences  $m$  which means it cannot take into account the effect of increasing the number of time-varying threshold sequences. Our objective is to recover the signal  $\bar{\mathbf{x}}$  within the polyhedron defined in (??), aiming for its proximity to the original signal  $\mathbf{x}_\star$ . From a mathematical perspective, we seek a signal  $\bar{\mathbf{x}}$  that lies within a ball space centered at the desired signal, characterized by a small radius  $\rho$ . In other words, we aim for  $\bar{\mathbf{x}}$  to belong to the ball space  $\mathcal{B}_\rho(\mathbf{x}_\star)$ . To achieve that, we are required to increase the number of samples in the polyhedron (??), as typically provided via one-bit sensing, to make a non-linear constraint redundant in the original problem of interest. In such a case, the offered convergence rate for the RKA appears to be insufficient since we must have enough number of samples to fulfill costly constraints (or equivalently, we must have enough number of samples so that the finite-volume space created by the intersection of hyperplanes of (??) be inside the desired signal's ball space  $\mathcal{B}_\rho(\mathbf{x}_\star)$ ). Define  $\text{vol}(\mathcal{P}_\mathbf{x})$  as the volume space created by the intersection of hyperplanes in (??). In the following proposition, we present the convergence rate of ORKA:

Proposition 1 (Convergence rate of ORKA). Consider the one-bit polyhedron  $\mathcal{P}_\mathbf{x}$  obtained in (??) associated with the linear system of equations  $\mathbf{A}\mathbf{x} = \mathbf{y}$  with  $\mathbf{A} \in \mathbb{R}^{n \times d}$ . Consider a ball centered at the original signal  $\mathcal{B}_\rho(\mathbf{x}_\star)$  and  $\hat{\mathbf{x}} \in \mathcal{P}_\mathbf{x}$ , a convergence rate for ORKA may be formulated as:

$$\mathbb{E} \{ \bar{h}(\mathbf{x}_i, \mathbf{x}_\star) \} \leq \left( 1 - \frac{1}{\kappa^2(\mathbf{A})} \right)^i \bar{h}(\mathbf{x}_0, \hat{\mathbf{x}}) + \rho^2, \quad (35)$$

with a probability higher than  $1 - e^{-\frac{c_1(C\rho-\mu)^2}{K}m'}$ .

Proof: As demonstrated in [?], the solution obtained from RKA lies in the space formed by the hyperplanes of the linear inequality problem with the following convergence rate:

$$\mathbb{E} \{ \hat{h}(\mathbf{x}_i, \hat{\mathbf{x}}) \} \leq \left( 1 - \frac{1}{\kappa^2(\mathbf{A})} \right)^i \hat{h}(\mathbf{x}_0, \hat{\mathbf{x}}), \quad (36)$$

where  $\hat{\mathbf{x}}$  is a point inside the space created by a polyhedron  $\mathcal{P}_{\mathbf{x}}$ . The convergence rate (??) only ensures that the solution will lie within the space created by the hyperplanes, not necessarily within the ball around the desired solution. However, in order to guarantee a perfect reconstruction and ensure that the solution of the linear feasibility problem lies within the ball around the desired solution with radius  $\rho$ , it is essential to have a sufficient number of samples. Until we reach the required number of samples, the upper bound of convergence to a solution lying within the space  $\mathcal{P}_{\mathbf{x}}$ , differs (and is smaller than) from that of the convergence to a solution within the ball around the desired point, since  $\text{vol}(\mathcal{P}_{\mathbf{x}}) \not\subseteq \text{vol}(\mathcal{B}_{\rho}(\mathbf{x}_{\star}))$ . However, once we obtain a sufficient number of samples to have the volume created by the intersections of hyperplanes inside the ball, i.e.,  $\text{vol}(\mathcal{P}_{\mathbf{x}}) \subseteq \text{vol}(\mathcal{B}_{\rho}(\mathbf{x}_{\star}))$ , we then have  $\hat{\mathbf{x}}$  lying within the ball around the desired point  $\mathbf{x}_{\star}$ , i.e.,  $\hat{\mathbf{x}} \in \mathcal{B}_{\rho}(\mathbf{x}_{\star})$ . To address this discrepancy between the two scenarios, we introduce a second term that is dependent on the difference between  $\hat{\mathbf{x}}$  and  $\mathbf{x}_{\star}$ , as follows:

$$\begin{aligned} \mathbb{E} \{ \|\mathbf{x}_i - \mathbf{x}_{\star}\|_2^2 \} &= \mathbb{E} \{ \|\mathbf{x}_i - \hat{\mathbf{x}} + \hat{\mathbf{x}} - \mathbf{x}_{\star}\|_2^2 \} \\ &\leq \mathbb{E} \{ \|\mathbf{x}_i - \hat{\mathbf{x}}\|_2^2 \} + \mathbb{E} \{ \|\mathbf{x}_{\star} - \hat{\mathbf{x}}\|_2^2 \}, \end{aligned} \quad (37)$$

where from (??) and the fact that the error between  $\hat{\mathbf{x}}$  and the original signal remains deterministic with respect to each iteration, we can write

$$\mathbb{E} \{ \|\mathbf{x}_i - \mathbf{x}_{\star}\|_2^2 \} \leq \left( 1 - \frac{1}{\kappa^2(\mathbf{A})} \right)^i \hat{h}(\mathbf{x}_0, \hat{\mathbf{x}}) + \|\mathbf{x}_{\star} - \hat{\mathbf{x}}\|_2^2. \quad (38)$$

The convergence to  $\mathbf{x}_{\star}$  is ensured only when the second term,  $\|\mathbf{x}_{\star} - \hat{\mathbf{x}}\|_2^2$ , is bounded. As demonstrated in Theorem ??, with a minimum probability of  $1 - e^{-\frac{c_1(C\rho-\mu)^2}{K}m'}$ , we establish that  $\text{vol}(\mathcal{P}_{\mathbf{x}}) \subseteq \text{vol}(\mathcal{B}_{\rho}(\mathbf{x}_{\star}))$  and consequently have  $\|\mathbf{x}_{\star} - \hat{\mathbf{x}}\|_2^2 \leq \rho^2$ , which proves the proposition. ■

In the following corollary, we present the required number of iterations  $i$  such that ORKA obtains an upper recovery bound  $\mathbb{E} \{ \|\mathbf{x}_i - \mathbf{x}_{\star}\|_2^2 \} \leq \epsilon_0$  at the  $i$ -th iteration:

Corollary 6. Based on the assumptions in Proposition ??, ORKA meets an upper recovery bound  $\mathbb{E} \left\{ \|\mathbf{x}_i - \mathbf{x}_\star\|_2^2 \right\} \leq \epsilon_0$  in the sample abundance scenario with the required number of iterations  $i$  satisfies  $i = \mathcal{O} \left( d \log \left( \frac{1}{\epsilon_0 - \rho^2} \right) \right)$ .

The proof of Corollary ?? is provided in Appendix ??.

Based on (??) and its proof, it is evident that the constant  $\delta_s$  contains  $\rho^{-2}$ , which leads to the conclusion that  $\rho = \mathcal{O} \left( m'^{-\frac{1}{3}} \right)$  and  $\rho^2 = \mathcal{O} \left( m'^{-\frac{2}{3}} \right)$ . As a result, the upper bound of ORKA error decays with a rate of  $\mathcal{O} \left( m'^{-\frac{2}{3}} \right)$  with respect to the number of samples. Similarly, for low-rank matrix sensing according to (??), the second term  $\rho^2 = \mathcal{O} \left( m'^{-\frac{2}{3}} \right)$ . Since the constant  $\delta_r$  also contains  $\rho^{-2}$ , the relation includes  $\rho$  as well. Hence, the upper bound of ORKA for low-rank matrix sensing decays with a rate of  $\mathcal{O} \left( m'^{-\frac{2}{3}} \right)$ .

Note that if any other randomized algorithm is utilized for one-bit sensing instead of RKA to achieve uniform reconstruction, the convergence rate will maintain the same structure as Proposition ??. However, there will be a difference in the first term, which will be substituted by the algorithm's convergence rate to a point inside the feasible space of hyperplanes.

## V. ORKA With Noisy Measurements

In addition to the theoretical assurances offered by our proposed algorithms, it is crucial to assess their effectiveness in the presence of noise. Previous studies, such as [?, ?], have examined one-bit noisy models with a linear measurement framework incorporating additive Gaussian noise. In these models, the input signal was recovered using a MLE approach, employing the Gaussian likelihood function. However, when dealing with non-Gaussian contamination, the MLE objective becomes nonconcave, leading to non-unique solutions for signal recovery. Moreover, MLE-based recovery is computationally more complex for high-dimensional signals. Herein, we formulate the noisy version of one-bit sampling with time-varying thresholds. Denote  $\mathbf{z} = [z_j] \in \mathbb{R}^n$  as a noise vector which has been added to the linear system of equations  $\mathbf{y} = \mathbf{A}\mathbf{x}$ . Then, the corresponding noisy one-bit samples are generated as

$$r_j^{(\ell)} = \begin{cases} +1 & \mathbf{a}_j \mathbf{x} + z_j > \tau_j^{(\ell)}, \\ -1 & \mathbf{a}_j \mathbf{x} + z_j < \tau_j^{(\ell)}, \end{cases} \quad j \in [n], \ell \in [m], \quad (39)$$

where  $\mathbf{a}_j$  denotes the  $j$ -th row of a sampling matrix  $\mathbf{A}$ . Consequently, the one-bit polyhedron associated with (??) is rewritten as



$$\mathbf{P}\mathbf{x} + \mathbf{v} \succeq \text{vec}(\mathbf{R}) \odot \text{vec}(\mathbf{\Gamma}), \quad (40)$$

where  $\mathbf{P}$  is defined in (??) and  $\mathbf{v} = \tilde{\mathbf{\Omega}}\mathbf{z}$  is the noise of our system with  $\tilde{\mathbf{\Omega}}$  defined in (??). For instance, assuming a zero-mean Gaussian noise vector  $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_z)$  with the covariance matrix  $\mathbf{\Sigma}_z$ , the distribution of  $\mathbf{v}$  will be  $\mathcal{N}(\mathbf{0}, \tilde{\mathbf{\Omega}}\mathbf{\Sigma}_z\tilde{\mathbf{\Omega}}^H)$ .

The robustness of the RKA against noise has been demonstrated in [?] and [?]. Furthermore, the authors of [?] specifically explored the performance of the RKA in the presence of Gaussian and Poisson noise, highlighting its robustness even when dealing with Poisson noisy measurements. In our discussion, in Section ?? we will explore how the inconsistency of a linear system in a noisy scenario manifests itself in the recovery error of the RKA. Next, in Section ?? we will propose a novel algorithm to have a robust recovery performance in the presence of impulsive noise.

#### A. Robustness of ORKA Against Noise

Given a linear system of equations  $\mathbf{U}\mathbf{x} = \mathbf{b}$  that is highly over-determined and subject to a noise vector  $\mathbf{n} = [n_j]$  resulting in a corrupted system of equations  $\mathbf{U}\mathbf{x} \approx \mathbf{b} + \mathbf{n}$ . The convergence rate of the noisy RKA was comprehensively discussed in [?, Theorem 2.1] for the case of  $\mathbf{U}\mathbf{x} \approx \mathbf{b} + \mathbf{n}$ . The primary contrast between the convergence rates of RKA and noisy RKA, as demonstrated in [?, Theorem 2.1], lies in the second term of convergence rate  $\kappa^2 \max_j \frac{n_j^2}{\|\mathbf{u}_j\|_2^2}$ . This term indicates the degree to which the error in the corrupted system  $\mathbf{U}\mathbf{x} \approx \mathbf{b} + \mathbf{n}$  deviates from the main solution.

Drawing inspiration from the convergence rate of the noisy RKA, we can similarly derive the convergence rate of noisy RKA in the case of noisy linear system of inequalities  $\mathbf{C}\mathbf{x} + \mathbf{n} \succeq \mathbf{b}$  using the following proposition:

**Proposition 2.** Let  $\mathbf{C} \in \mathbb{R}^{m \times n}$  have full column rank and assume  $\hat{\mathbf{x}}$  is the solution of the noisy linear feasibility problem  $\mathbf{C}\mathbf{x} + \mathbf{n} \succeq \mathbf{b}$ . Let  $\bar{\mathbf{x}}_i$  be the  $i$ -th iterate of the noisy RKA run with  $\mathbf{C}\mathbf{x} \succeq \mathbf{b}$ , and let  $n_j$  denote the  $j$ -th element of  $\mathbf{n}$ , respectively. Then we have

$$\mathbb{E} \{ \bar{h}(\bar{\mathbf{x}}_i, \hat{\mathbf{x}}) \} \leq \left( 1 - \frac{1}{\kappa^2(\mathbf{C})} \right)^i \bar{h}(\mathbf{x}_0, \hat{\mathbf{x}}) + \kappa^2 \max_j \gamma_j, \quad (41)$$

where  $\gamma_j = \frac{(n_j^+)^2}{\|\mathbf{c}_j\|_2^2}$ .

The proof of Proposition ?? is presented in Appendix ?. Based on the noisy one-bit polyhedron (??), the parameters  $\mathbf{C} = \mathbf{P}$ ,  $\mathbf{n} = \mathbf{v}$ , and  $\mathbf{b} = \text{vec}(\mathbf{R}) \odot \text{vec}(\mathbf{\Gamma})$  can be replaced in the convergence rate presented in (??) to get the similar result in the one-bit noisy scenario. Note that as can be observed in (??), a small perturbation in the linear feasibility problem  $\mathbf{C}\mathbf{x} \succeq \mathbf{b}$  may slightly deviate the solution of the noisy RKA from the main solution.

## B. Upper Quantile-Based ORKA

In Section ??, the robustness of the noisy RKA in the presence of a small perturbation has been discussed. However, as can be observed in the convergence rate of the noisy RKA (??), a large perturbation in the linear feasibility  $\mathbf{C}\mathbf{x} \succeq \mathbf{b}$  can lead to a significant error in the input signal recovery. In such a scenario, an accurate input signal recovery is quite a challenging task. A well-known example of large perturbation is impulsive noise. Impulsive noise poses substantial challenges in the realm of signal processing and imaging applications. Within the domain of audio signal processing, it manifests as disruptive events such as clicks, pops, or random bursts, leading to a degradation in sound quality [?, ?]. In the field of magnetic resonance imaging (MRI), the presence of impulsive noise gives rise to unwanted anomalies and distortions in the acquired images [?]. In this section, our goal is to propose a novel RKA-based algorithm which is robust to impulsive noise. The noisy linear inequality feasibility problem is defined as  $\mathbf{C}\mathbf{x} + \mathbf{n} \succeq \mathbf{b}$ , where  $\mathbf{n}$  is the noise of our system. If the noise does not have a significant impact on the inequality, we can rephrase the system with noise as  $\mathbf{C}\mathbf{x} \succeq \mathbf{b}$  and solve it using the RKA problem. However, if the noise has the potential to alter the direction of the inequalities, we need to handle it differently. In this section, specifically, we introduce an algorithm that identifies the “orthants” that are immune to corruption (where noise cannot change their direction), and only incorporate them in the updates of RKA.

The probability of orthants that are corrupted with noise is formulated as the following upper quantile

$$\alpha_j = \Pr(\mathbf{c}_j\mathbf{x} - b_j \geq -n_j), \quad (42)$$

where  $b_j$  and  $n_j$  are  $j$ -th elements of  $\mathbf{b}$  and  $\mathbf{n}$ , respectively. By using this formulation, we may be able to determine a threshold for our residuals  $\{\mathbf{c}_j\mathbf{x} - \mathbf{b}_j\}$  that can be used to

identify the corrupted ones. The threshold is calculated based on the empirical  $q$ -quantile of the noise:

$$\mathcal{Q}(\mathbf{x}) \triangleq q\text{-quantile} \{|\mathbf{c}_j \mathbf{x} - b_j|, j \in [m]\}. \quad (43)$$

If a residual exceeds this threshold, it suggests that the noise may not have been strong enough to change its direction. By applying the thresholding process, we ensure that a sufficient distance is maintained between  $\mathbf{c}_j \mathbf{x}$  and  $b_j$  to prevent any noise from impacting the inequality system. This helps to preserve the integrity of the solution throughout the algorithm. Assume  $\mathbf{p}_j$  is the  $j$ -th row of  $\mathbf{P}$  randomly chosen at each iteration  $i$ , the proposed algorithm for the noisy one-bit sampled systems (??), upper quantile-based ORKA is written as follows: (i) update the RKA projection if  $|\mathbf{p}_j \mathbf{x}_i - r_j^{(\ell)} \tau_j^{(\ell)}| \geq \mathcal{Q}(\mathbf{x}_i)$ , and (ii) otherwise, set  $\mathbf{x}_i = \mathbf{x}_{i+1}$ .

## VI. Judicious Sampling With Adaptive Thresholding for ORKA

By the spirit of using the iterative RKA, suitable time-varying sampling thresholds can be selected in order to enhance the recovery performance. In the sample abundance ORKA, we face a highly overdetermined linear feasibility problem creating a finite-volume space. To capture the desired signal  $\mathbf{x}_\star$  more efficiently, the right-hand side of the inequalities in (??), i.e.  $\text{vec}(\mathbf{R}) \odot \text{vec}(\mathbf{\Gamma})$ , must be determined in a way that each associated hyperplane passes through the desired feasible region within  $\mathcal{F}_\mathbf{x}$ . Therefore, an algorithm is proposed to ensure that this occurs in practice. Accordingly, we propose an iterative algorithm generating adaptive sampling thresholds to accurately obtain the desired solution. To have a smaller area of the finite-volume space around the desired signal  $\mathbf{x}_\star$ , one can somehow choose thresholds to reduce the distances between the desired point and the associated hyperplanes in (??). To do so, we update the time-varying thresholds as

$$\begin{cases} \boldsymbol{\tau}_{k+1}^{(\ell)} = \mathbf{A} \mathbf{x}_k - \frac{1}{2} \left( \mathbf{r}^{(\ell)} \odot \boldsymbol{\epsilon}_k^{(\ell)} \right), \\ \boldsymbol{\epsilon}_k^{(\ell)} = \mathbf{r}^{(\ell)} \odot \left( \mathbf{A} \mathbf{x}_k - \boldsymbol{\tau}_k^{(\ell)} \right), \end{cases} \quad (44)$$

where  $\mathbf{x}_k$  and  $\boldsymbol{\tau}_k^{(\ell)}$  denote the  $k$ -th updates of  $\mathbf{x}$  and  $\boldsymbol{\tau}^{(\ell)}$  in our proposed adaptive threshold design algorithm. The proposed sampling algorithm is summarized in Algorithm ??. By employing this adaptive thresholding algorithm, it is notable that each distance defined in (??) tends to decrease, resulting in a reduction of  $T_{\text{ave}}$  with a high probability. Therefore, a smaller number of time-varying sampling threshold sequences can be utilized in ORKA with similar

---

**Algorithm 3** Adaptive Thresholding for ORKA
 

---

Input: One-bit data, time-varying sampling thresholds, and the sampling matrix.

Output: A solution  $\bar{\mathbf{x}}$  in the one-bit polyhedron (??).

- 1:  $\mathbf{b}^{(\ell)} \leftarrow \boldsymbol{\tau}_k^{(\ell)}$ .
  - 2:  $\mathcal{P}_k \leftarrow \left\{ \mathbf{x}_k \mid r_j^{(\ell)} \mathbf{a}_j \mathbf{x}_k \geq r_j^{(\ell)} b_j^{(\ell)}, j \in [n], \ell \in [m] \right\}$ .
  - 3: Obtain  $\mathbf{x}_k$  in  $\mathcal{P}_k$  by RKA, PrSKM or Block SKM.
  - 4:  $\boldsymbol{\epsilon}_k^{(\ell)} \leftarrow \mathbf{r}^{(\ell)} \odot (\mathbf{A}\mathbf{x}_k - \mathbf{b}^{(\ell)})$ .
  - 5:  $\boldsymbol{\tau}_{k+1}^{(\ell)} \leftarrow \mathbf{A}\mathbf{x}_k - \frac{1}{2} \left( \mathbf{r}^{(\ell)} \odot \boldsymbol{\epsilon}_k^{(\ell)} \right)$ .
  - 6: Increase  $k$  by one.
  - 7: Repeat Steps (1)-(6) until  $\sum_{\ell=1}^m \left\| \boldsymbol{\tau}_{k+1}^{(\ell)} - \boldsymbol{\tau}_k^{(\ell)} \right\|_2 \leq \delta$ .
  - 8: return  $\bar{\mathbf{x}}$
- 

recovery performance. Additionally, non-informative sampling thresholds, which appear as extra inequality constraints in the random time-varying sampling thresholds scenario, may be efficiently removed by choosing the adaptive thresholds with closer hyperplanes to the desired point.

## VII. One-Bit Low-Rank Matrix Sensing

Low-rank matrix sensing is an excellent example for problems that assume the form in (??), and that can be tackled using our methodology. In Section ??, we first briefly introduce the nuclear norm minimization form of the low-rank matrix sensing problem. Subsequently, we apply ORKA to this problem without considering the associated costly constraints. As mentioned previously, there exists a trade-off between the number of samples and the computational complexity of the reconstruction algorithm in the problem of signal parameters recovery. In Section ??, we specifically address the scenario where, due to practical limitations or other factors, an adequate number of measurements or one-bit samples may not be available to satisfy the FVP described in Section ?. To address this challenge, we will introduce a novel algorithm called SVP-ORKA.

### A. Problem Formulation

The problem of the low-rank matrix sensing is formulated as:

$$\text{find } \mathbf{X} \in \Omega_c \quad \text{subject to} \quad \mathcal{A}(\mathbf{X}) = \mathbf{y}, \quad \text{rank}(\mathbf{X}) \leq r, \quad (45)$$

where  $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2}$  is the matrix of unknowns,  $\mathbf{y} \in \mathbb{R}^n$  is the measurement vector, and  $\mathcal{A}$  is a linear transformation such that  $\mathcal{A} : \mathbb{R}^{n_1 \times n_2} \mapsto \mathbb{R}^n$ . In general,  $\Omega_c$  can be chosen such as the set of semi-definite matrices, symmetric matrices, upper or lower triangle matrices, Hessenberg matrices and a specific constraint on the matrix elements  $\|\mathbf{X}\|_\infty \leq \alpha$  or on its eigenvalues, i.e.,  $\lambda_i \leq \epsilon$  where  $\{\lambda_i\}$  are eigenvalues of  $\mathbf{X}$  [?, ?, ?]. The problem (??) can be rewritten as an optimization problem:

$$\underset{\mathbf{X} \in \Omega_c}{\text{minimize}} \quad \text{rank}(\mathbf{X}) \quad \text{subject to} \quad \mathcal{A}(\mathbf{X}) = \mathbf{y}. \quad (46)$$

This problem is known to be NP-hard, whose solution is difficult to approximate [?, ?]. Recall that the rank of  $\mathbf{X}$  is equal to the number of nonzero singular values. In the case when the singular values are all equal to one, the sum of the singular values is equal to the rank. When the singular values are less than or equal to one, the sum of the singular values is a convex function that is strictly less than the rank. Therefore, it has been popular for this problem to replace the rank function with the sum of the singular values of  $\mathbf{X}$ ; i.e., its nuclear norm. The nuclear norm minimization alternative of the problem is given by [?, ?, ?]:

$$\underset{\mathbf{X} \in \Omega_c}{\text{minimize}} \quad \|\mathbf{X}\|_* \quad \text{subject to} \quad \mathcal{A}(\mathbf{X}) = \mathbf{y}. \quad (47)$$

In this problem, the feasible set  $\mathcal{F}_{\mathbf{X}}$  is obtained as  $\mathcal{F}_{\mathbf{X}} = \{\mathcal{P}_{\mathbf{X}}^* \cap \Omega_c\}$ , where  $\mathcal{P}_{\mathbf{X}}^*$  is defined as follows

$$\mathcal{P}_{\mathbf{X}}^* = \{\mathbf{X} \mid \|\mathbf{X}\|_* \leq \tau\}, \quad \tau \in \mathbb{R}^+. \quad (48)$$

Next, we will apply ORKA to (??) to make its costly constraints redundant by using abundant number of one-bit samples  $m'$ .

In low-rank matrix sensing, the linear operator  $\mathcal{A}(\mathbf{X})$  is obtained as [?],

$$\mathcal{A}(\mathbf{X}) = \frac{1}{\sqrt{n}} [\text{Tr}(\mathbf{A}_1^\top \mathbf{X}) \cdots \text{Tr}(\mathbf{A}_n^\top \mathbf{X})]^\top, \quad (49)$$

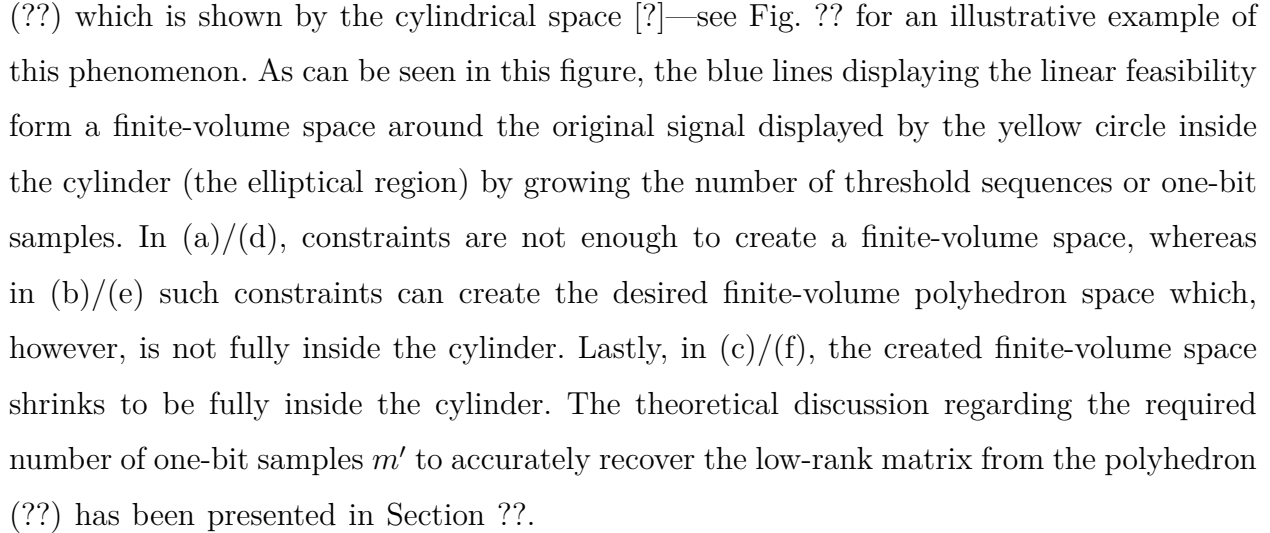
where  $\mathbf{A}_j \in \mathbb{R}^{n_1 \times n_2}$  is the  $j$ -th sensing matrix. The one-bit polyhedron for the low-rank matrix sensing is given by

$$\mathcal{P}^{(M)} = \left\{ \mathbf{X} \mid r_j^{(\ell)} \text{Tr}(\mathbf{A}_j^\top \mathbf{X}) \geq r_j^{(\ell)} \tau_j^{(\ell)}, \quad j \in [n], \quad \ell \in [m] \right\}. \quad (50)$$

Writing the update process of ORKA in matrix form yields the following representation:

$$\mathbf{X}_{i+1} = \mathbf{X}_i + \frac{\left( r_j^{(\ell)} \tau_j^{(\ell)} - r_j^{(\ell)} \text{Tr}(\mathbf{A}_j^\top \mathbf{X}_i) \right)^+}{\|\mathbf{A}_j\|_F^2} \mathbf{A}_j. \quad (51)$$

By applying the vectorization operator to both sides of (??) and utilizing the fact that  $\|\mathbf{A}_j\|_F = \|\text{vec}(\mathbf{A}_j)\|_2$ , we obtain an update process identical to that of RKA. Therefore, the convergence rate of this process is equivalent to that of RKA. Denote the update process in (??) by  $\mathbf{X}_{i+1} = \text{KA}_r(\mathbf{X}_i)$ . One can similarly utilize other variants of RKA, PrSKM and Block SKM, and generalize the update process (??) as  $\mathbf{X}_{i+1} = \text{KA}_p(\mathbf{X}_i)$  and  $\mathbf{X}_{i+1} = \text{KA}_b(\mathbf{X}_i)$ .

A numerical investigation of (??) reveals that by increasing the number of time-varying sampling threshold sequences  $m$ , the space formed by the intersection of half-spaces (inequality constraints) can fully shrink to the desired signal  $\mathbf{X}_\star$  inside the feasible region of (??) which is shown by the cylindrical space [?]—see Fig. ?? for an illustrative example of this phenomenon. As can be seen in this figure, the blue lines displaying the linear feasibility form a finite-volume space around the original signal displayed by the yellow circle inside the cylinder (the elliptical region) by growing the number of threshold sequences or one-bit samples. In (a)/(d), constraints are not enough to create a finite-volume space, whereas in (b)/(e) such constraints can create the desired finite-volume polyhedron space which, however, is not fully inside the cylinder. Lastly, in (c)/(f), the created finite-volume space shrinks to be fully inside the cylinder. The theoretical discussion regarding the required number of one-bit samples  $m'$  to accurately recover the low-rank matrix from the polyhedron (??) has been presented in Section ??.

## B. ORKA With Low-Rank Matrix Factorization

Although Kaczmarz algorithms use a simple and low-complexity update process, the computational cost can still be prohibitively large for a matrix  $\mathbf{X} \in \mathbb{R}^{n_1 \times n_2}$  when  $n_1$  and  $n_2$  are large values, typically requiring  $\mathcal{O}(n_1 n_2)$  operations. To address this issue, we employ the well-known low-rank matrix factorization technique. Instead of using the full matrix  $\mathbf{X}$ , we use its low-rank factorization  $\mathbf{X} = \mathbf{L}\mathbf{W}^\top$ , where  $\mathbf{L} \in \mathbb{R}^{n_1 \times r}$  and  $\mathbf{W} \in \mathbb{R}^{n_2 \times r}$  are low-rank factors. The advantage of this approach is that the economical representation of the low-rank matrix results in lower storage requirements, lower per-iteration computational cost, and better amenability to parallelization. It also makes the method more scalable to larger problem sizes, especially when using iterative optimization methods like ORKA. The combination of low-rank matrix factorization with gradient descent methods has been extensively discussed in the literature [?]. In this section, we explore the integration of low-

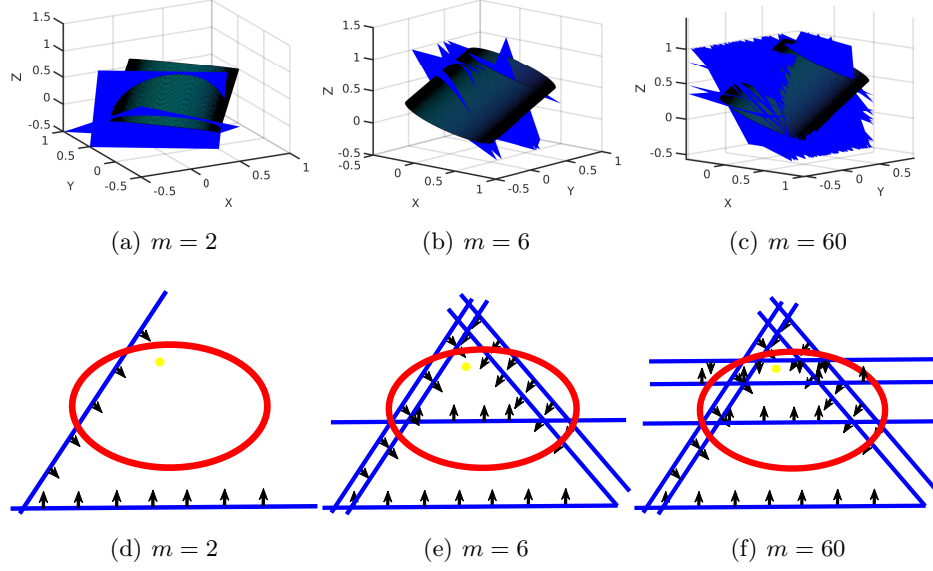


Figure 2. Shrinkage of the one-bit polyhedron (??) in blue, ultimately placed within the unit ball of the nuclear norm  $\|\mathbf{X}\|_* \leq 1$  shown with black cylindrical region and its red contours, when the number constraints (samples) grows large. The arrows point to the half-space associated with each inequality constraint. The evolution of the feasible regime is depicted with increasing samples in three cases: (a) and (d) small sample-size regime, constraints not forming a finite-value polyhedron; (b) and (e) medium sample-size regime, constraints forming a finite-volume polyhedron, parts of which are outside the cylinder; (c) and (f) large sample-size regime, constraints forming a finite-volume polyhedron inside the nuclear norm cylinder, making its constraint redundant. The original signal representing the signal to be recovered is shown by yellow.

rank matrix factorization with ORKA and the alternating minimization (AltMin) algorithm discussed in [?].

The one-bit polyhedron  $\mathcal{P}^{(M)}$  associated with the low-rank matrix factorization approach is written as

$$\mathcal{P}_0^{(M)} = \left\{ \mathbf{L}, \mathbf{W} \mid r_j^{(\ell)} \text{Tr}(\mathbf{A}_j^\top \mathbf{L} \mathbf{W}^\top) \geq r_j^{(\ell)} \tau_j^{(\ell)}, j \in [n], \ell \in [m] \right\}. \quad (52)$$

To find the solution from  $\mathcal{P}_0^{(M)}$ , we use the idea of AltMin algorithm. We split  $\mathcal{P}_0^{(M)}$  into two linear feasibility sub-problems with respect to  $\mathbf{L}$  and  $\mathbf{W}$ , respectively. Specifically, with respect to  $\mathbf{L}$  when  $\mathbf{W}$  is fixed we have:

$$\mathcal{P}_{\mathbf{L}}^{(M)} = \left\{ \mathbf{L} \mid r_j^{(\ell)} \text{Tr}(\mathbf{W}^\top \mathbf{A}_j^\top \mathbf{L}) \geq r_j^{(\ell)} \tau_j^{(\ell)}, j \in [n], \ell \in [m] \right\}, \quad (53)$$

and with respect to  $\mathbf{W}$  when  $\mathbf{L}$  is fixed we have:

$$\mathcal{P}_{\mathbf{W}}^{(M)} = \left\{ \mathbf{W} \mid r_j^{(\ell)} \text{Tr}(\mathbf{A}_j^\top \mathbf{L} \mathbf{W}^\top) \geq r_j^{(\ell)} \tau_j^{(\ell)}, j \in [n], \ell \in [m] \right\}. \quad (54)$$

---

**Algorithm 4** ORKA (with Low-Rank Matrix Factorization)

---

Input: One-bit samples, time-varying thresholds and the sensing matrices presented in (??), total number of required iterations  $T$ , and  $\text{rank}(\mathbf{X}_\star) = r$ .

Output: A solution  $\bar{\mathbf{X}} \in \mathbb{R}^{n_1 \times n_2}$  in the polyhedron (??).

Note:  $\mathbf{L}_t \in \mathbb{R}^{n_1 \times r}$  and  $\mathbf{W}_t \in \mathbb{R}^{n_2 \times r}$  denote the obtained matrices at  $t$ -th iteration,  $\mathbf{H}_{i+1} = \text{KA}_r(\mathbf{H}_i; \tilde{\mathbf{H}})$  denotes the update process of ORKA using the RKA when  $\tilde{\mathbf{H}}$  is fixed.

- 1:  $\mathbf{W}_{t+1} \leftarrow \text{KA}_r(\mathbf{W}_t; \mathbf{L}_t) \triangleright$  Update process for polyhedron (??).
  - 2:  $\mathbf{L}_{t+1} \leftarrow \text{KA}_r(\mathbf{L}_t; \mathbf{W}_{t+1}) \triangleright$  Update process for polyhedron (??).
  - 3: Repeat steps (1) and (2) until convergence.
  - 4:  $\mathbf{L}_\star \leftarrow \mathbf{L}_T$  and  $\mathbf{W}_\star \leftarrow \mathbf{W}_T$ .
  - 5: return  $\bar{\mathbf{X}} \leftarrow \mathbf{L}_\star \mathbf{W}_\star^\top$ .
- 

If either  $\mathbf{L}$  or  $\mathbf{W}$  are fixed, finding the solution with respect to the other variable is achieved via ORKA as indicated in Algorithm ??.

### C. Curse of Dimensionality: SVP-ORKA

While acquiring a large number of samples is not typically an issue for one-bit sampling, it is still practical to avoid using excess samples, particularly in signal processing applications where access to sufficient measurements may be limited. Adhering to the law of parsimony, “Entia non svnt mvltiplicanda præter necessitatem,” i.e., entities should not be multiplied beyond necessity, this section focuses on developing ORKA to facilitate low-rank matrix sensing with a reduced number of one-bit samples. As with any technique, there is a trade-off between the number of samples and the complexity of that technique, and this section aims to address the challenges associated with a limited number of samples.

Define  $r$  as the predefined rank of the unknown matrix  $\mathbf{X}$ . In order to obtain the solution within a reduced number of samples in the polyhedron  $\mathcal{P}^{(M)}$  defined in (??), we impose a rank constraint,  $\text{rank}(\mathbf{X}) \leq r$ , to shrink the entire space, as shown by the following polyhedron:

$$\mathcal{P}_1^{(M)} = \left\{ \mathbf{X} \mid r_j^{(\ell)} \text{Tr}(\mathbf{A}_j^\top \mathbf{X}) \geq r_j^{(\ell)} \tau_j^{(\ell)}, \text{rank}(\mathbf{X}) \leq r \right\}, \quad (55)$$



where  $j \in [n]$ ,  $\ell \in [m]$ . To tackle this problem, we apply the SVP method to ORKA. The SVP was introduced as a solution to the general affine rank minimization problem (ARMP). In [?], it was demonstrated that the SVP can effectively recover the minimum rank solution even in the presence of noise and when the affine constraints satisfy RIP. Moreover, some theoretical guarantees for this approach were also established. This method utilizes the operator  $P_r$  to modify the gradient descent process at each iteration. The operator  $P_r$  calculates the  $r$  largest singular values of a matrix and subsequently rewrites its SVD based on these  $r$  singular values and their corresponding singular vectors. Similar to Section ??, denote  $\text{KA}_r(\cdot)$  as the update process of ORKA using the RKA defined in (??). Through the integration of SVP into each iteration of ORKA, we can achieve the following update process:

$$\begin{cases} \mathbf{Z}_{i+1} = \text{KA}_r(\mathbf{X}_i), \\ \mathbf{X}_{i+1} = P_r(\mathbf{Z}_{i+1}). \end{cases} \quad (56)$$

The first step of the update process (??),  $\mathbf{Z}_{i+1} = \text{KA}_r(\mathbf{X}_i)$ , can be also replaced with  $\mathbf{Z}_{i+1} = \text{KA}_p(\mathbf{X}_i)$  or  $\mathbf{Z}_{i+1} = \text{KA}_b(\mathbf{X}_i)$  to be consistent with the updates of PrSKM and Block SKM, respectively. The convergence guarantee of SVP-ORKA is concluded according to the following lemma:

Lemma 3. The update process of SVP-ORKA presented in (??) converges linearly in expectation to a ball centered at the original signal  $\mathcal{B}_\rho(\text{vec}(\mathbf{X}_*))$  with the number of samples satisfying Theorem ?? and probability exceeding  $1 - \eta$ , as follows:

$$\mathbb{E} \{ \|\mathbf{X}_{i+1} - \mathbf{X}_*\|_{\text{F}}^2 \} \leq \left( 1 - \frac{1}{\kappa^2(\mathbf{V})} \right)^i \left\| \mathbf{X}_0 - \widehat{\mathbf{X}} \right\|_{\text{F}}^2 + \rho^2, \quad (57)$$

where  $\widehat{\mathbf{X}} \in \mathcal{P}_1^{(M)}$ , and  $\mathbf{V}$  is the matrix with vectorized sensing matrices  $\{\text{vec}(\mathbf{A}_j)\}_{j=1}^n$  as its rows.

The proof of Lemma (??) is studied in Appendix ??.

### VIII. One-Bit Compressed Sensing: From Optimization to Linear Feasibility

CS is an interesting and rapidly growing area of research that has attracted considerable attention in electrical engineering, applied mathematics, statistics, and computer science

[?, ?]. In CS, the objective is to recover a sparse high-dimensional signal from incomplete measurements, which may be formulated as follows [?]:

$$\underset{\mathbf{x}}{\text{minimize}} \quad \|\mathbf{x}\|_1 \quad \text{subject to} \quad \mathbf{A}\mathbf{x} = \mathbf{y}, \quad (58)$$

where  $\mathbf{A} \in \mathbb{R}^{n \times d}$ , and  $n \ll d$ .

The concise summary of CS theory mentioned earlier implies that, under ideal circumstances, the measurement vector  $\mathbf{y}$  is assumed to be represented with infinite precision. However, real-world sensing models typically involve digitalization and finite precision data representations for purposes such as storing, transmitting, or processing acquired observations.

Following our motivation in one-bit low-rank matrix recovery, herein we are inspired to investigate the one-bit CS problem in (i) sample abundance, and (ii) sample restricted scenarios. In Section ??, our goal is to examine the reconstruction performance of one-bit CS with multiple time-varying threshold sequences under a sample-abundance regime without any form of regularization. Beyond the sample abundance scenario, in Section ?? we also develop the ORKA for conditions where enough measurements to create a highly-overdetermined feasibility problem are not available. It is crucial to highlight that the choice of a particular one-bit reconstruction scheme for compressed sensing measurements is heavily influenced by factors such as available computational resources, implementation complexity, and cost considerations associated with the sensor.

#### A. One-Bit CS via Sample Abundance

As discussed earlier, by deploying one-bit quantization, the opportunity exists to increase the number of samples in (??). The one-bit CS is thus solely accomplished by creating a highly-constrained one-bit polyhedron:

$$\mathcal{P}^{(C)} = \left\{ \mathbf{x} \mid r_j^{(\ell)} \mathbf{a}_j \mathbf{x} \geq r_j^{(\ell)} \tau_j^{(\ell)}, \quad j \in [n], \quad \ell \in [d] \right\}. \quad (59)$$

In other words, instead of solving an optimization problem with costly constraints, the problem may be tackled by the following update process:

$$\mathbf{x}_{i+1} = \text{KA}_r(\mathbf{x}_i), \quad (60)$$

where  $\text{KA}_r(\cdot)$  denotes the RKA update process presented as  $\mathbf{x}_{i+1} = \mathbf{x}_i + \frac{(r_j^{(\ell)} \tau_j^{(\ell)} - r_j^{(\ell)} \mathbf{a}_j \mathbf{x}_i)^+}{\|\mathbf{a}_j\|_2^2} \mathbf{a}_j^H$ . The update process in (??) can be also replaced with  $\mathbf{x}_{i+1} = \text{KA}_p(\mathbf{x}_i)$  or  $\mathbf{x}_{i+1} = \text{KA}_b(\mathbf{x}_i)$

to be consistent with the updates of PrSKM and Block SKM, respectively. The theoretical discussion about the minimum number of one-bit samples  $m'$  to precisely recover the sparse vector from the polyhedron (??) has been presented in Section ??.

## B. Regularized ORKA for One-Bit CS: ST-ORKA

In this section, we continue our exploration of algorithms for one-bit CS by introducing the  $\ell_1$  regularized ORKA. Building on the motivation behind our earlier proposal of the SVP-ORKA, ORKA is developed to efficiently recover a sparse solution from the one-bit CS polyhedron using a reduced number of one-bit samples. This feature proves particularly useful in scenarios where the number of measurements is severely limited. The regularized one-bit CS polyhedron is obtained as

$$\mathcal{P}_1^{(C)} = \left\{ \mathbf{x} \mid r_j^{(\ell)} \mathbf{a}_j \mathbf{x} \geq r_j^{(\ell)} \tau_j^{(\ell)}, \|\mathbf{x}\|_1 \leq \kappa \right\}, \quad (61)$$

where  $j \in [n]$  and  $\ell \in [m]$ . Define the soft thresholding (ST) operator as  $S_\kappa(\mathbf{x}) = \text{sgn}(\mathbf{x})(\|\mathbf{x}\| - \kappa)^+$ , where  $\kappa$  is the predefined threshold. The ST-ORKA utilizes the ST operator  $S_\kappa$  to project each iteration of ORKA to the set  $\{\|\mathbf{x}\|_1 \leq \kappa\}$  with the following update process:

$$\begin{cases} \mathbf{z}_{i+1} = \text{KA}_r(\mathbf{x}_i), \\ \mathbf{x}_{i+1} = S_\kappa(\mathbf{z}_{i+1}), \end{cases} \quad (62)$$

where  $\text{KA}_r(\cdot)$  denotes the RKA updates. The first step of the update process (??) can be also replaced with  $\mathbf{z}_{i+1} = \text{KA}_p(\mathbf{x}_i)$  or  $\mathbf{z}_{i+1} = \text{KA}_b(\mathbf{x}_i)$  to be consistent with the updates of PrSKM and Block SKM, respectively. The convergence rate of ST-ORKA is studied in the following lemma:

Lemma 4. Consider a ball centered at the original signal  $\mathcal{B}_\rho(\mathbf{x}_\star)$  and  $\hat{\mathbf{x}} \in \mathcal{P}_1^{(C)}$ , the convergence of ST-ORKA presented in (??) is given by

$$\mathbb{E} \left\{ \|\mathbf{x}_{i+1} - \mathbf{x}_\star\|^2 \right\} \leq \left( 1 - \frac{1}{\kappa^2(\mathbf{A})} \right)^i \|\mathbf{x}_0 - \hat{\mathbf{x}}\|_2^2 + \rho^2, \quad (63)$$

with a probability exceeding  $1 - \eta$ , and the number of samples satisfying Theorem ??.

The proof of Lemma ?? is studied in Appendix ??. Note that to have a guaranteed convergence, one can integrate ORKA with different operators that satisfy Lipschitz continuity. This is important in various applications and provides a chance to go beyond just ST-ORKA

and SVP-ORKA. By doing so, we can achieve a convergence rate that is described in the following lemma:

Lemma 5. Assume  $f(\cdot)$  is an operator such that for any  $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^d$  we have  $\|f(\mathbf{x}_1) - f(\mathbf{x}_2)\|_2^2 \leq L \|\mathbf{x}_1 - \mathbf{x}_2\|_2^2$ . Then the integration of RKA with the operator  $f$  in each iteration of solving the feasibility problem  $\mathbf{C}\mathbf{x} \succeq \mathbf{b}$  with  $\mathbf{C} \in \mathbb{R}^{n \times d}$  has the following convergence rate:

$$\mathbb{E} \{ \|\mathbf{x}_{i+1} - \hat{\mathbf{x}}\|_2^2 \} \leq L \left( 1 - \frac{1}{\kappa^2(\mathbf{C})} \right)^i \|\mathbf{x}_0 - \hat{\mathbf{x}}\|_2^2. \quad (64)$$

A notable example for Lemma ?? is the integration of ORKA with the hard thresholding (HT) operator  $\mathcal{T}_s(\cdot)$  to reconstruct the  $s$ -sparse signal in the one-bit CS problem. This operator selects the best  $s$ -sparse approximation of the solution at each iteration. The algorithm is named HT-ORKA with the update process presented as

$$\begin{cases} \mathbf{z}_{i+1} = \mathbf{K}\mathbf{A}_r(\mathbf{x}_i), \\ \mathbf{x}_{i+1} = \mathcal{T}_s(\mathbf{z}_{i+1}). \end{cases} \quad (65)$$

The convergence rate of HT-ORKA can be determined using the following lemma:

Lemma 6. Assume  $\mathcal{B}_\rho(\mathbf{x}_\star)$  be a ball centered at the original signal and  $\hat{\mathbf{x}} \in \mathcal{P}_1^{(C)}$ , the convergence of HT-ORKA presented in (??) is given by

$$\mathbb{E} \{ \|\mathbf{x}_{i+1} - \mathbf{x}_\star\|^2 \} \leq 2 \left( 1 - \frac{1}{\kappa^2(\mathbf{A})} \right)^i \|\mathbf{x}_0 - \hat{\mathbf{x}}\|_2^2 + \rho^2. \quad (66)$$

with probability higher than  $1 - \eta$ , and the number of samples satisfying Theorem ??.

The proof of Lemma ?? is provided in Appendix ??. In their work, the authors of [?] introduced the random hyperplane tessellations theorem for the ditherless scenario in one-bit CS, specifically targeting the reconstruction of signal direction. This theorem was developed in [?] for one-bit CS with Gaussian dithering, where based on the augmentation trick [?, ?], random hyperplane tessellations are limited to Gaussian random variables for the sampling matrix and normal distribution for the dithering. The random hyperplane tessellations theorem has previously been extended for certain non-Gaussian sampling matrices in [?, ?]. In their work, the authors derived theoretical guarantees for Uniform dithering and focused on the Hamming distance. The main distinction between the FVP and random hyperplane tessellations lies in the modeling of the created finite volume around the original signal. While random hyperplane tessellations utilize Hamming distance to capture the direction

of measurements, we incorporate Euclidean distance between the one-bit hyperplanes and the original signal, considering the importance of amplitude in dithered one-bit sensing. By using the Euclidean distances between the original signal and the surrounding hyperplanes as the basis of our work, we were able to obtain theoretical guarantees beyond Uniform dithering. This is a crucial distinction, as it highlights the significance of the difference between measurements and thresholds in one-bit sensing. Intuitively, if measurements and thresholds become closer and closer, the recovery performance improves.

## IX. Numerical Investigations

In this section, we conduct numerical evaluations to assess the performance of our proposed algorithms in two distinct scenarios: (i) sample abundance and (ii) sample restriction. All presented results are averaged over 1000 experiments.

### A. Sample abundance

In this particular scenario, we examine two examples: one-bit low-rank matrix sensing and one-bit CS. For both cases, we employ the Block SKM algorithm to recover the desired signal and then evaluate its performance against the adaptive threshold strategy introduced in [?, Section VI]. The only distinction is that, in this study, we did not update the one-bit data to account for an exceedingly low-complexity hardware implementation. In all experiments, we have taken into account the presence of Gaussian additive noise, also known as prequantization error, with a standard deviation of 0.1.

One-bit low-rank matrix sensing. We considered a set of sampling matrices  $\{\mathbf{A}_j\}_{j=1}^{1800}$ , where each entry is independently drawn from a standard normal distribution. We have generated the desired matrix  $\mathbf{X}_\star \in \mathbb{R}^{30 \times 30}$  such that  $\text{rank}(\mathbf{X}_\star) = 2$ . The number of time-varying sampling threshold sequences were set to  $m \in \{1, 10, 20, 30\}$ . Accordingly, we have generated sequences of time-varying sampling thresholds as  $\left\{ \boldsymbol{\tau}^{(\ell)} \sim \mathcal{N}\left(\mathbf{0}, \frac{\beta_{\mathbf{y}}^2}{9} \mathbf{I}\right) \right\}_{\ell=1}^m$ , where  $\beta_{\mathbf{y}}$  denotes the dynamic range of the high-resolution measurements  $\mathbf{y}$ . Figure ??(a) illustrates a comparison between the recovery performance of Block SKM using random thresholds and adaptive thresholds. It is evident that the utilization of adaptive thresholds enhances the recovery performance compared to random thresholds.

One-bit CS. We have generated a sensing matrix  $\mathbf{A} \in \mathbb{R}^{500 \times 100}$  in which each element follows a standard normal distribution. The desired signal  $\mathbf{x}_\star \in \mathbb{R}^{100}$  was assumed to have the level

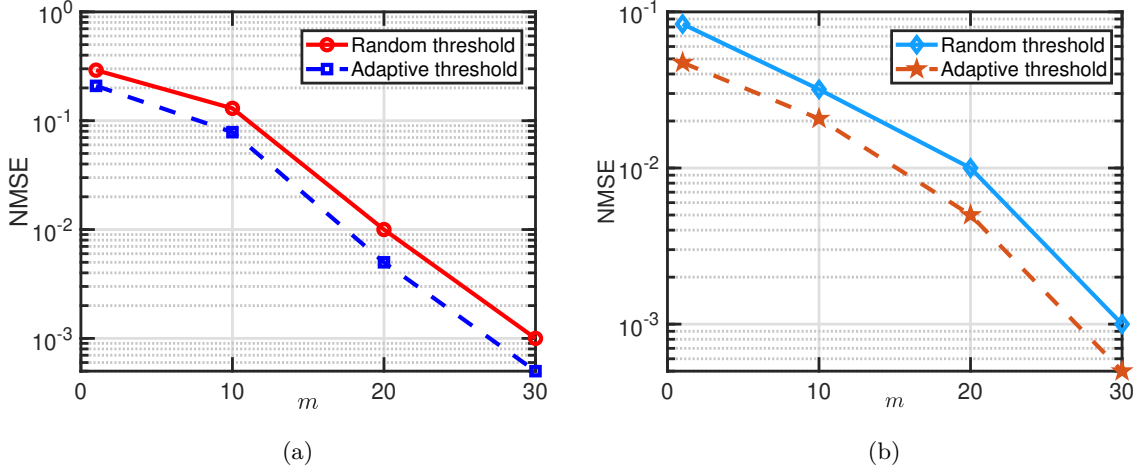


Figure 3. Comparison between the recovery performance of Block SKM using random thresholds and adaptive thresholds in the sample abundance scenario for (a) one-bit low-rank matrix sensing, and (ii) one-bit CS.

of sparsity  $s = 10$ . The settings for time-varying sampling thresholds were considered to be the same as the one-bit low-rank matrix sensing case. Fig. ??(b) displays the recovery performance of Block SKM using random thresholds in comparison with adaptive thresholds. Consistent with previous observations, the utilization of adaptive thresholds improves the recovery performance.

## B. Sample Scarcity

Similar to sample abundance, herein we investigate the performance of our proposed methods for one-bit low-rank matrix sensing and one-bit CS when we have a limited number of samples. Note that in all experiments, the high-resolution measurements were contaminated by the additive Gaussian noise with the standard deviation 0.1 except the case related to low-rank matrix sensing by SVP-ORKA which was considered to be noiseless.

**One-bit low-rank matrix sensing.** We generated a collection of sampling matrices  $\{\mathbf{A}_j\}_{j=1}^n$ , where each entry is independently sampled from a standard normal distribution. The desired matrix  $\mathbf{X}_\star \in \mathbb{R}^{30 \times 30}$  was generated with  $\text{rank}(\mathbf{X}_\star) = 2$ . Define the oversampling factor as  $\lambda = \frac{n}{n_1 r} = \frac{n}{60}$ . In our experiments, we have set  $\log(\lambda) \in \{3, 4, 5, 6\}$ . The number of time-varying sampling threshold sequences was fixed at  $m = 1$ . The generation of time-varying sampling thresholds followed the same procedure as in the previous cases. Fig. ??(a) compares the recovery performance of SVP-ORKA with hard singular value thresholding (HSVT) algorithm [?] in the noiseless scenario. As can be observed, SVP-

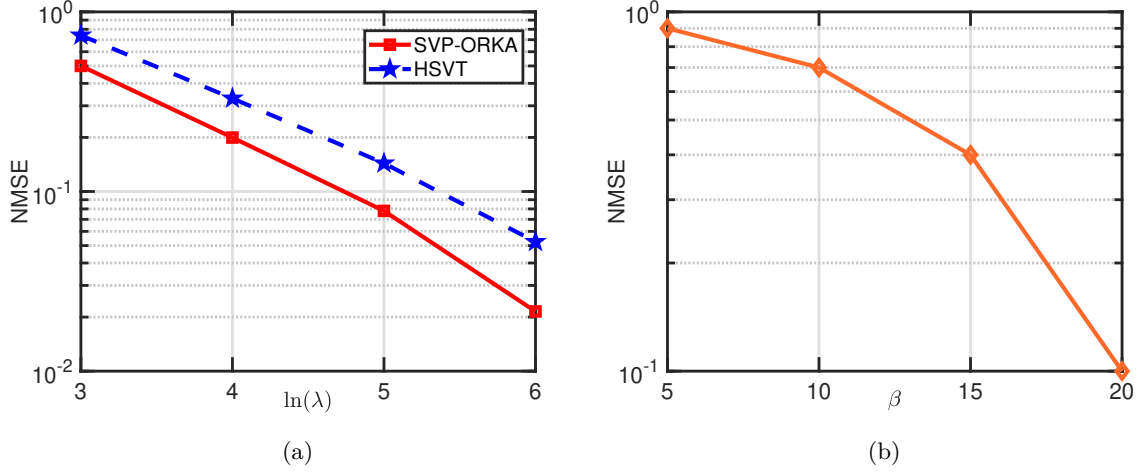


Figure 4. (a) Comparison between the recovery performance of SVP-ORKA and HSVT algorithm over different values of oversampling factor  $\lambda$ . (b) Recovery performance of Algorithm ?? (ORKA with low-rank matrix factorization) over different values of sampling factor  $\beta$ .

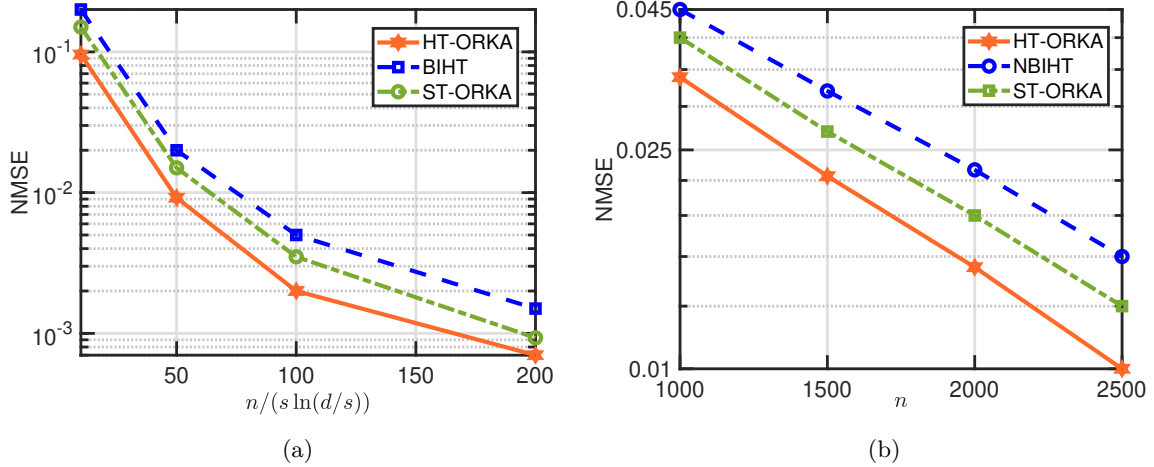


Figure 5. Comparison between the recovery performance of HT-ORKA, ST-ORKA and (a) BIHT with random thresholds, and (b) NBIHT in ditherless scenario.

ORKA outperforms HSVT over different values of the oversampling factor. In another experimental setting aimed at assessing the performance of Algorithm ?? (ORKA with low-rank matrix factorization), we generated the desired matrix  $\mathbf{X}_\star \in \mathbb{R}^{30 \times 30}$  with  $\text{rank}(\mathbf{X}_\star) = 1$ . The remaining parameter settings were identical to the previous example. Note that in this case we define the oversampling factor as  $\beta = \frac{n}{n_1^2 r} = \frac{n}{900}$ . In our simulations, we have set  $\beta = \{5, 10, 15, 20\}$ . As can be seen in Fig. ??(b), the recovery performance of Algorithm ?? enhances as the value of the oversampling factor  $\beta$  grows large.

One-bit CS. Each element of the sensing matrix  $\mathbf{A} \in \mathbb{R}^{n \times 100}$  was independently drawn from a standard normal distribution. The desired signal  $\mathbf{x}_* \in \mathbb{R}^{100}$  was assumed to have a sparsity level of  $s = 15$ . Define the oversampling factor as  $n/(s \log(d/s))$ . In our experiments, we have set the oversampling factor to the values  $\{10, 50, 100, 200\}$ . The number of time-varying sampling threshold sequences was fixed at  $m = 1$ , and the generation of time-varying sampling thresholds followed the same procedure as in the previous cases. In Fig. ??(a), we compare the recovery performance of HT-ORKA, ST-ORKA, and BIHT algorithm with random time-varying sampling thresholds [?]. It is evident that HT-ORKA outperforms both ST-ORKA and BIHT in recovering the  $s$ -sparse signal in the one-bit CS problem. In the next example, we examine the effectiveness of our proposed algorithms, HT-ORKA and ST-ORKA, in recovering a sparse signal in a ditherless scenario. The sensing matrix  $\mathbf{A} \in \mathbb{R}^{n \times 256}$  was generated in the same manner as in the previous example. The number of high-resolution samples was considered to be  $n \in \{1000, 1500, 2000, 2500\}$ . The desired signal  $\mathbf{x}_* \in \mathbb{R}^{256}$  was assumed to have a sparsity level of  $s = 25$ . In Fig. ??(b), we compare the recovery performance of HT-ORKA, ST-ORKA and the NBIHT algorithm [?]. Once again, similar to the previous example, HT-ORKA exhibits superior recovery performance compared to ST-ORKA and the NBIHT method.

## X. Discussion

In this paper, we have established the theoretical guarantees for uniform perfect reconstruction in dithered one-bit sensing. Our approach involves transforming the one-bit signal reconstruction problem into a linear feasibility problem. We then introduced the FVP theorem to analyze the possibility of creating a finite volume formed by the hyperplanes around the original signal. The FVP theorem allows us to determine the minimum probability of achieving perfect reconstruction and the required number of samples to attain uniform perfect reconstruction. What sets this theorem apart from others, such as the random hyperplane tessellations theorem, is that it approaches one-bit sensing from the perspective of a linear feasibility problem. It investigates the number of samples needed to capture the original signal within the space of hyperplanes. An intriguing aspect of the FVP theorem is its capability to provide guarantees not only for Uniform dithering and restricted sampling matrices considered in previous efforts. Leveraging this theorem, we were able to derive theoretical guarantees for deterministic matrices like DCT as well. The only



restriction in the FVP theorem is that the distances between the original signals and the surrounding hyperplanes defined in (??), should be considered as sub-Gaussian. It remains an open problem for other distributions, which may be explored in future research.

This work represents a pioneering effort in the literature, as it explores the performance of randomized algorithms in one-bit sensing for the first time. We introduced two novel variations of the Kaczmarz algorithm, PrSKM and Block SKM, which served as the foundation for our proposed algorithm, ORKA. In our investigation of ORKA, we analyzed its upper recovery bound and demonstrated that it decays concerning the number of measurements. Specifically, for both compressed sensing and low-rank matrix recovery, the decay rate is  $\mathcal{O}\left(m'^{-\frac{2}{3}}\right)$ . These findings contribute valuable insights into the potential of randomized algorithms in one-bit sensing applications. To the best of our knowledge, we are the first to derive the convergence rate of RKA for a noisy linear inequality system. This novel finding highlights the robustness of the algorithm, even in the presence of noise.

We have introduced an improved update process for designing the thresholding process. Unlike the sigma-delta thresholding design discussed in references [?, ?], our approach does not require updating the one-bit data, eliminating the need for feedback in the sampling scheme. Through numerical demonstrations, we showed that this adaptive thresholding process enhances signal reconstruction performance from one-bit data. Through numerical experiments, we experimentally demonstrated that the proposed algorithm exhibits superior reconstruction performance in one-bit CS compared to the NBIHT for the ditherless scenario and BIHT adapted with dithering for the dithered scenario. Furthermore, our results show that the proposed SVP-ORKA outperforms the HSVT algorithm in terms of recovery accuracy.

In addition to the common situation of sample abundance in dithered one-bit quantization, we also address scenarios with sample restrictions. For these scenarios, we further develop the proposed randomized algorithms into storage-friendly approaches, such as random sketching and low-rank matrix factorization. This extension allows for efficient handling of limited samples, broadening the applicability of the algorithms to a wider range of practical settings. The convergence rate of ORKA with low-rank matrix factorization remains an open problem. This is because it employs the idea of cyclic algorithms or AltMin, whose convergence is still an ongoing topic of research in the literature [?].

## Appendix A

### Proof of Theorem ??

Since  $\mathbf{N} \in \mathbb{R}^{m \times s}$  is a Gaussian matrix with sketch size  $s = \mathcal{O}(n/\epsilon^2)$  and  $\epsilon = \frac{1}{2}$ , the  $\epsilon$ -subspace embedding property holds with high probability for all  $\mathbf{x} \in \mathbb{R}^n$  [?],

$$(1 - \epsilon) \|\mathbf{C}\mathbf{x}\|_2^2 \leq \|\mathbf{N}^\top \mathbf{C}\mathbf{x}\|_2^2 \leq (1 + \epsilon) \|\mathbf{C}\mathbf{x}\|_2^2. \quad (67)$$

As studied in [?], we can use this result to show that the preconditioned system  $\mathbf{C}\mathbf{R}_s^{-1}$  preserves the length of vectors in the range of  $\mathbf{C}$ . To see how, substitute  $\mathbf{x}$  in (??) with  $\mathbf{R}_s^{-1}\mathbf{z}$ . By the subspace embedding property,

$$\frac{1}{1 + \epsilon} \|\mathbf{N}^\top \mathbf{C}\mathbf{R}_s^{-1}\mathbf{z}\|_2^2 \leq \|\mathbf{C}\mathbf{R}_s^{-1}\mathbf{z}\|_2^2 \leq \frac{1}{1 - \epsilon} \|\mathbf{N}^\top \mathbf{C}\mathbf{R}_s^{-1}\mathbf{z}\|_2^2. \quad (68)$$

Now using  $\|\mathbf{N}^\top \mathbf{C}\mathbf{R}_s^{-1}\mathbf{z}\|_2^2 = \|\mathbf{Q}_s \mathbf{z}\|_2^2 = \|\mathbf{z}\|_2^2$ , the relation (??) becomes

$$\frac{1}{1 + \epsilon} \|\mathbf{z}\|_2^2 \leq \|\mathbf{C}\mathbf{R}_s^{-1}\mathbf{z}\|_2^2 \leq \frac{1}{1 - \epsilon} \|\mathbf{z}\|_2^2, \quad (69)$$

which bounds the condition number  $\varrho(\mathbf{C}\mathbf{R}_s^{-1}) \leq \sqrt{\frac{1+\epsilon}{1-\epsilon}}$ . Setting  $\epsilon = \frac{1}{2}$ , we have  $\varrho(\mathbf{C}\mathbf{R}_s^{-1}) \leq \sqrt{3}$  and based on (??),  $\kappa(\mathbf{C}\mathbf{R}_s^{-1}) \leq \sqrt{3n}$ . Therefore, according to (??) and the result that we have obtained here, one can conclude the convergence bound (??).

## Appendix B

### Convergence analysis of Block SKM with the sparse Gaussian sketch (??)

The Block SKM algorithm can be considered to be a special case of the more general sketch-and-project method with a sparse block sketch matrix as defined in [?].

$$\mathbf{x}_{i+1} = \underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{x} - \mathbf{x}_i\|_2^2 \quad \text{subject to} \quad \mathbf{S}^\top \mathbf{B}\mathbf{x} \succeq \mathbf{S}^\top \mathbf{b}, \quad (70)$$

where  $\mathbf{S} \in \mathbb{R}^{mn \times n}$  is the sketch matrix choosing a block uniformly at random from the main matrix  $\mathbf{B}$  similar to step 1 of Algorithm ??. The second step of sketch-and-project method follows the Motzkin sampling where the index  $j_i^*$  is chosen in  $i$ -th iteration as

$$j_i^* = \underset{j}{\operatorname{argmax}} \left\{ \left( (\mathbf{S}^\top \mathbf{b})_j - (\mathbf{S}^\top \mathbf{B})_j \mathbf{x}_i \right)^+ \right\}, \quad (71)$$

with  $(\cdot)_j$  denoting the  $j$ -th row of the matrix/vector argument (this step is similar to steps 2–4 of Algorithm ?? with  $k' = 1$ ). In the Block SKM algorithm, the sketch matrix is given by

$$\mathbf{S} = \left[ \mathbf{0}_{n \times p} \mid \mathbf{I}_n \mid \mathbf{0}_{n \times (mn - n - p)} \right]^\top, \quad \mathbf{S} \in \mathbb{R}^{mn \times n}, \quad (72)$$

where  $p = n\alpha$ ,  $\alpha \in \{1, \dots, m-1\}$ . Note that the literature does not offer any theoretical guarantees for the convergence of the Block SKM with the identity matrix [?]. To derive our theoretical guarantees for the Block SKM algorithm, we change the sketch matrix to the sparse Gaussian sketch matrix as follows:

$$\mathbf{S} = \left[ \mathbf{0}_{n \times p} \mid \mathbf{G} \mid \mathbf{0}_{n \times (mn-n-p)} \right]^\top, \quad \mathbf{S} \in \mathbb{R}^{mn \times n}, \quad (73)$$

where  $\mathbf{G}$  is a  $n \times n$  Gaussian matrix, whose entries are i.i.d. following the distribution  $\mathcal{N}(0, 1)$ . In this framework, we are able to provide some theoretical guarantees by taking advantage of the favorable properties of Gaussian random variables. Assume that  $\mathcal{P}_{\mathbf{x}}$  denotes a nonempty solution set of  $\mathbf{B}\mathbf{x} \succeq \mathbf{b}$  and  $\hat{\mathbf{x}} \in \mathcal{P}_{\mathbf{x}}$ . Considering the sparse Gaussian sketch (??) which satisfies the set of inequalities  $\mathbf{S}^\top \mathbf{B}\mathbf{x} \succeq \mathbf{S}^\top \mathbf{b}$ , we have,

$$\begin{aligned} \|\mathbf{x}_{i+1} - \hat{\mathbf{x}}\|_2^2 &= \left\| (\mathbf{x}_i - \hat{\mathbf{x}}) + \frac{\left( (\mathbf{S}^\top \mathbf{b})_{j_i^*} - (\mathbf{S}^\top \mathbf{B})_{j_i^*} \mathbf{x}_i \right)^+ (\mathbf{S}^\top \mathbf{B})_{j_i^*}^\mathbf{H}}{\left\| (\mathbf{S}^\top \mathbf{B})_{j_i^*} \right\|_2^2} \right\|_2^2 \\ &= \|\mathbf{x}_i - \hat{\mathbf{x}}\|_2^2 + \frac{\left( \left( (\mathbf{S}^\top \mathbf{b})_{j_i^*} - (\mathbf{S}^\top \mathbf{B})_{j_i^*} \mathbf{x}_i \right)^+ \right)^2}{\left\| (\mathbf{S}^\top \mathbf{B})_{j_i^*} \right\|_2^2} + \\ &\quad \frac{2 \left( (\mathbf{S}^\top \mathbf{b})_{j_i^*} - (\mathbf{S}^\top \mathbf{B})_{j_i^*} \mathbf{x}_i \right)^+ (\mathbf{S}^\top \mathbf{B})_{j_i^*}^\mathbf{H} (\mathbf{x}_i - \hat{\mathbf{x}})}{\left\| (\mathbf{S}^\top \mathbf{B})_{j_i^*} \right\|_2^2}. \end{aligned} \quad (74)$$

We can observe that  $(\mathbf{S}^\top \mathbf{B})_{j_i^*}^\mathbf{H} (\mathbf{x}_i - \hat{\mathbf{x}}) \leq (\mathbf{S}^\top \mathbf{B})_{j_i^*}^\mathbf{H} \mathbf{x}_i - (\mathbf{S}^\top \mathbf{b})_{j_i^*}$ , therefore, the left-hand side of (??) is less than or equal to,

$$\begin{aligned} \|\mathbf{x}_i - \hat{\mathbf{x}}\|_2^2 &+ \frac{\left( \left( (\mathbf{S}^\top \mathbf{b})_{j_i^*} - (\mathbf{S}^\top \mathbf{B})_{j_i^*} \mathbf{x}_i \right)^+ \right)^2}{\left\| (\mathbf{S}^\top \mathbf{B})_{j_i^*} \right\|_2^2} + \\ &\frac{2 \left( (\mathbf{S}^\top \mathbf{b})_{j_i^*} - (\mathbf{S}^\top \mathbf{B})_{j_i^*} \mathbf{x}_i \right)^+ \left( (\mathbf{S}^\top \mathbf{B})_{j_i^*}^\mathbf{H} \mathbf{x}_i - (\mathbf{S}^\top \mathbf{b})_{j_i^*} \right)}{\left\| (\mathbf{S}^\top \mathbf{B})_{j_i^*} \right\|_2^2} \\ &= \|\mathbf{x}_i - \hat{\mathbf{x}}\|_2^2 - \frac{\left( \left( (\mathbf{S}^\top \mathbf{b})_{j_i^*} - (\mathbf{S}^\top \mathbf{B})_{j_i^*} \mathbf{x}_i \right)^+ \right)^2}{\left\| (\mathbf{S}^\top \mathbf{B})_{j_i^*} \right\|_2^2}. \end{aligned} \quad (75)$$

Based on the definition of  $j_i^*$ , one can rewrite (??) as,

$$\|\mathbf{x}_{i+1} - \hat{\mathbf{x}}\|_2^2 \leq \|\mathbf{x}_i - \hat{\mathbf{x}}\|_2^2 - \frac{\left\|(\mathbf{S}^\top \mathbf{b} - \mathbf{S}^\top \mathbf{B} \mathbf{x}_i)^+\right\|_\infty^2}{\left\|(\mathbf{S}^\top \mathbf{B})_{j_i^*}\right\|_2^2}. \quad (76)$$

By taking the expectation over the error, we have

$$\mathbb{E}_{\mathbf{S}} \left\{ \|\mathbf{x}_{i+1} - \hat{\mathbf{x}}\|_2^2 \right\} \leq \|\mathbf{x}_i - \hat{\mathbf{x}}\|_2^2 - \mathbb{E}_{\mathbf{S}} \left\{ \frac{\left\|(\mathbf{S}^\top \mathbf{b} - \mathbf{S}^\top \mathbf{B} \mathbf{x}_i)^+\right\|_\infty^2}{\left\|(\mathbf{S}^\top \mathbf{B})_{j_i^*}\right\|_2^2} \right\}. \quad (77)$$

In addition, we have that

$$\mathbb{E}_{\mathbf{S}} \left\{ \left\|(\mathbf{S}^\top \mathbf{B})_{j_i^*}\right\|_2^2 \right\} = \sum_{k=1}^d \mathbb{E}_{\mathbf{S}} \left\{ \left( \sum_{l=1}^{mn} \mathbf{S}_{j_i^* l} \mathbf{B}_{lk} \right)^2 \right\}, \quad (78)$$

or equivalently, in terms of  $\mathbf{G}$  in (??),

$$\begin{aligned} \sum_{k=1}^d \mathbb{E}_{\mathbf{G}} \left\{ \left( \sum_{l=1}^n \mathbf{G}_{j_i^* l}^\top \mathbf{B}_{lk} \right)^2 \right\} = \\ \sum_{k=1}^d \sum_{l=1}^n \mathbb{E}_{\mathbf{G}} \left\{ \left( \mathbf{G}_{j_i^* l}^\top \right)^2 \right\} \mathbf{B}_{lk}^2, \end{aligned} \quad (79)$$

with  $\mathbb{E}_{\mathbf{G}} \left\{ \left( \mathbf{G}_{j_i^* l}^\top \right)^2 \right\} = 1$ , which helps to simplify (??) as

$$\sum_{k=1}^d \sum_{l=1}^n \mathbf{B}_{lk}^2 = \|\hat{\mathbf{B}}\|_{\text{F}}^2, \quad (80)$$

where  $\hat{\mathbf{B}}$  is the  $n \times d$  submatrix of  $\mathbf{B}$  (one of the candidates of  $\mathbf{B}_j$  in Algorithm ??). Due to the fact that the second term in the right-hand side of (??) is an expectation over the convex function  $f(x, y) = x^2/y$ , we can apply Jensen's inequality as follows:

$$\mathbb{E}_{\mathbf{S}} \left\{ \frac{\left\|(\mathbf{S}^\top \mathbf{b} - \mathbf{S}^\top \mathbf{B} \mathbf{x}_i)^+\right\|_\infty^2}{\left\|(\mathbf{S}^\top \mathbf{B})_{j_i^*}\right\|_2^2} \right\} \geq \frac{\left( \mathbb{E}_{\mathbf{S}} \left\{ \left\|(\mathbf{S}^\top \mathbf{b} - \mathbf{S}^\top \mathbf{B} \mathbf{x}_i)^+\right\|_\infty \right\} \right)^2}{\mathbb{E}_{\mathbf{S}} \left\{ \left\|(\mathbf{S}^\top \mathbf{B})_{j_i^*}\right\|_2^2 \right\}}. \quad (81)$$

Consider the following lemma regarding the estimate of the maximum of independent normal random variables:

Lemma 7. [?, Section 2.5.2] Let  $X_1, \dots, X_n$  be independent  $\mathcal{N}(0, 1)$  random variables. Then we have

$$\mathbb{E} \left\{ \max_{i \leq n} X_i \right\} \geq c \sqrt{\log n}, \quad (82)$$

where  $c$  is an absolute constant.

By taking advantage of Lemma ??, we have

$$\begin{aligned}
\mathbb{E}_{\mathbf{S}} \left\{ \left\| (\mathbf{S}^\top \mathbf{b} - \mathbf{S}^\top \mathbf{B} \mathbf{x}_i)^+ \right\|_\infty \right\} &= \mathbb{E}_{\mathbf{S}} \left\{ \max_{t \in [n]} (\langle \mathbf{s}_t, \mathbf{b} - \mathbf{B} \mathbf{x}_i \rangle)^+ \right\} \\
&\geq \mathbb{E}_{\mathbf{S}} \left\{ \max_{t \in [n]} \langle \mathbf{s}_t, (\mathbf{b} - \mathbf{B} \mathbf{x}_i)^+ \rangle \right\} \\
&= \mathbb{E}_{\mathbf{G}} \left\{ \max_{t \in [n]} \left\langle \mathbf{g}_t, (\hat{\mathbf{b}} - \hat{\mathbf{B}} \mathbf{x}_i)^+ \right\rangle \right\} \\
&\geq c \left\| (\hat{\mathbf{b}} - \hat{\mathbf{B}} \mathbf{x}_i)^+ \right\|_2 \sqrt{\log n},
\end{aligned} \tag{83}$$

where  $\hat{\mathbf{b}} \in \mathbb{R}^n$  is a block of  $\mathbf{b}$ ,  $\mathbf{s}_t$  and  $\mathbf{g}_t$  are the  $t$ -th columns of  $\mathbf{S}$  and  $\mathbf{G}$ , respectively,  $[n] = \{1, 2, \dots, n\}$ , and  $c$  is a positive value. By plugging the inequality (??) into (??) and using the Hoffman bound [?, Theorem 4.2], we have

$$\begin{aligned}
\mathbb{E} \left\{ \|\mathbf{x}_{i+1} - \hat{\mathbf{x}}\|_2^2 \right\} &\leq \|\mathbf{x}_i - \hat{\mathbf{x}}\|_2^2 - \frac{c \|(\hat{\mathbf{b}} - \hat{\mathbf{B}} \mathbf{x}_i)^+\|_2^2 \log n}{\|\hat{\mathbf{B}}\|_{\text{F}}^2} \\
&\leq \|\mathbf{x}_i - \hat{\mathbf{x}}\|_2^2 - \frac{c \sigma_{\min}^2(\hat{\mathbf{B}}) \log n}{\|\hat{\mathbf{B}}\|_{\text{F}}^2} \|\mathbf{x}_i - \hat{\mathbf{x}}\|_2^2 \\
&\leq \left( 1 - \frac{c \sigma_{\min}^2(\hat{\mathbf{B}}) \log n}{\|\hat{\mathbf{B}}\|_{\text{F}}^2} \right) \|\mathbf{x}_i - \hat{\mathbf{x}}\|_2^2,
\end{aligned} \tag{84}$$

which can be recast as the following convergence rate, after  $K$  updates:

$$\mathbb{E} \left\{ \|\mathbf{x}_{i+1} - \hat{\mathbf{x}}\|_2^2 \right\} \leq \left( 1 - \frac{c \sigma_{\min}^2(\hat{\mathbf{B}}) \log n}{\|\hat{\mathbf{B}}\|_{\text{F}}^2} \right)^K \|\mathbf{x}_0 - \hat{\mathbf{x}}\|_2^2. \tag{85}$$

## Appendix C

### Proof of Theorem ??

To prove Theorem ??, define the set  $\mathcal{T} = \{\mathbf{x} \in \mathbb{R}^d \mid \|\mathbf{x}\|_2 \leq 1\}$ . Assume that  $\mathbb{E}\{T_{\text{ave}}\} = \mu$  and  $\left\| d_j^{(\ell)} \right\|_{\psi_2}^2 \leq K$ . Then our goal is to obtain the required number of one-bit samples  $m'$  to achieve the uniform perfect reconstruction criterion with  $\bar{\mathbf{x}} \in \mathcal{B}_\rho(\mathbf{x}_*)$  considering all  $\mathbf{x}_*, \bar{\mathbf{x}} \in \mathcal{T}$ . The condition  $\bar{\mathbf{x}} \in \mathcal{B}_\rho(\mathbf{x}_*)$  or equivalently  $\|\bar{\mathbf{x}} - \mathbf{x}_*\|_2 \leq \rho$  implies a creation of finite-volume in  $\mathbb{R}^d$  with a maximum radius of  $\rho$ . To see the connection between this condition,  $\|\bar{\mathbf{x}} - \mathbf{x}_*\|_2 \leq \rho$ , and  $T_{\text{ave}}$ , there exists a set  $\mathcal{G}$  with the minimum cardinality  $\text{card}(\mathcal{G}) \sim d^2$

<sup>2</sup>In  $\mathbb{R}^d$ , we need at least  $d+1$  number of hyperplanes to create a finite-volume space.

which contains  $\text{card}(\mathcal{G})$  minimum distances  $d_j^{(\ell)}$  such that  $\frac{1}{\text{card}(\mathcal{G})} \sum_{d_j^{(\ell)} \in \mathcal{G}} d_j^{(\ell)} \sim \rho$ . For the complement of the set  $\mathcal{G}$  denoted by  $\bar{\mathcal{G}}$ , we have  $d_j^{(\ell)} \geq \rho$  for all  $d_j^{(\ell)} \in \bar{\mathcal{G}}$  which leads to  $\frac{1}{\text{card}(\bar{\mathcal{G}})} \sum_{d_j^{(\ell)} \in \bar{\mathcal{G}}} d_j^{(\ell)} \geq \rho$ . Therefore, we can conclude that  $T_{\text{ave}} \geq \rho$ . To see why the set  $\mathcal{G}$  with the minimum cardinality  $\text{card}(\mathcal{G}) \sim d$  exists such that  $\frac{1}{\text{card}(\mathcal{G})} \sum_{d_j^{(\ell)} \in \mathcal{G}} d_j^{(\ell)} \sim \rho$ , we present the following lemma:

Lemma 8. In  $\mathbb{R}^d$ , with high probability there exists a set  $\mathcal{G}$  with the minimum cardinality  $\text{card}(\mathcal{G}) \sim d$  which contains  $\text{card}(\mathcal{G})$  minimum distances  $d_j^{(\ell)}$  such that  $\frac{1}{\text{card}(\mathcal{G})} \sum_{d_j^{(\ell)} \in \mathcal{G}} d_j^{(\ell)} \sim \rho$  implying  $\|\bar{\mathbf{x}} - \mathbf{x}_\star\|_2 \leq \rho$ .

Proof: For simplicity denote  $T_{\text{ave}}^{\mathcal{G}} = \frac{1}{\text{card}(\mathcal{G})} \sum_{d_j^{(\ell)} \in \mathcal{G}} d_j^{(\ell)}$ . If  $\exists \rho' > \rho$  such that  $T_{\text{ave}}^{\mathcal{G}} \sim \rho'$ , then  $\exists \mathcal{G}' \subseteq \mathcal{G}$  such that  $T_{\text{ave}}^{\mathcal{G}'} \sim \rho$  with  $\text{card}(\mathcal{G}') < \text{card}(\mathcal{G}) \sim d$ . This property,  $\text{card}(\mathcal{G}') < \text{card}(\mathcal{G}) \sim d$ , contradicts the creation of a finite-volume with the maximum radius of  $\rho$  in  $\mathbb{R}^d$ . Therefore, we can only assume that  $T_{\text{ave}}^{\mathcal{G}} < \rho$ . Define a set  $\hat{\mathcal{G}} = \mathcal{G} \cup \tilde{\mathcal{G}}$  such that  $T_{\text{ave}}^{\hat{\mathcal{G}}} \sim \rho$ , where  $\tilde{\mathcal{G}}$  is a set with  $J$  minimum distances of the set  $\bar{\mathcal{G}}$ . If  $J < \text{card}(\bar{\mathcal{G}})$ , then such set  $\hat{\mathcal{G}}$  exists which partitions all distances  $d_j^{(\ell)}$  into  $\hat{\mathcal{G}}$  with  $T_{\text{ave}}^{\hat{\mathcal{G}}} \sim \rho$  and  $\bar{\mathcal{G}} \setminus \tilde{\mathcal{G}}$  with  $T_{\text{ave}}^{\bar{\mathcal{G}} \setminus \tilde{\mathcal{G}}} \geq \rho$ . The existence of the set  $\hat{\mathcal{G}}$  with  $\text{card}(\hat{\mathcal{G}}) \sim d + J$  such that  $T_{\text{ave}}^{\hat{\mathcal{G}}} \sim \rho$  implies the creation of a finite-volume space with the maximum radius of  $\rho$  in  $\mathbb{R}^d$ . If  $J > \text{card}(\bar{\mathcal{G}})$ , the set  $\hat{\mathcal{G}}$  does not exist which informs  $T_{\text{ave}} < \rho$ . Since  $m' \geq \text{card}(\mathcal{G}) \sim d$ , the result  $T_{\text{ave}} < \rho$  implies the creation of a finite-volume space in  $\mathbb{R}^d$  with the maximum radius less than  $\rho$  with high probability. In fact, one can apply the general Hoeffding's inequality [?, Theorem 2.6.2] to the random variable  $T_{\text{ave}}$  as follows:

$$\Pr(T_{\text{ave}} \leq \mu - \delta) \leq e^{-\frac{c_1 \delta^2 m'}{K}}, \quad (86)$$

where  $\delta$  and  $c_1$  are positive constants. If  $\delta \geq \mu - \rho$ , then with probability at most  $e^{-\frac{c_1 \delta^2 m'}{K}}$  we have  $T_{\text{ave}} \leq \rho$ . Based on this result and the fact  $m' \geq \text{card}(\mathcal{G}) \sim d$ , we can conclude that  $T_{\text{ave}} \geq \rho$  with high probability. This, in turn, indicates the existence of a set  $\mathcal{G}$  with the minimum cardinality  $\text{card}(\mathcal{G}) \sim d$  such that  $T_{\text{ave}}^{\mathcal{G}} \sim \rho$ . ■

Based on Lemma ??, existing a set  $\mathcal{G}$  with the minimum cardinality  $\text{card}(\mathcal{G}) \sim d$  such that  $\frac{1}{\text{card}(\mathcal{G})} \sum_{d_j^{(\ell)} \in \mathcal{G}} d_j^{(\ell)} \sim \rho$  with high probability informs that  $\|\bar{\mathbf{x}} - \mathbf{x}_\star\|_2 \leq \rho$ . To meet  $\|\bar{\mathbf{x}} - \mathbf{x}_\star\|_2 \leq \rho$ ,  $\exists C > 0$  such that  $T_{\text{ave}} \leq C\rho$ , where  $C\rho = \mu + \delta$  with a positive constant  $\delta$ . To derive the probability of the event  $\|\bar{\mathbf{x}} - \mathbf{x}_\star\|_2 \leq \rho$  for a specific  $\mathbf{x}_\star \in \mathcal{T}$ , we apply the

general Hoeffding's inequality [?, Theorem 2.6.2] to the event  $T_{\text{ave}} \leq C\rho$  as follows:

$$\Pr(T_{\text{ave}} \geq C\rho) \leq e^{-\frac{c_1(C\rho-\mu)^2}{K}m'}. \quad (87)$$

Consider the following lemma:

Lemma 9. Define  $T_{\text{ave}}(\mathbf{x}_\star)$  as in (??). Then for any  $\mathbf{x}_\star, \bar{\mathbf{x}} \in \mathcal{T}$  we have

$$T_{\text{ave}}(\mathbf{x}_\star) \leq \frac{1}{m'} \sum_{\ell=1}^m \|\boldsymbol{\tau}^{(\ell)}\|_1 + c_{\mathbf{A}} (\|\mathbf{x}_\star - \bar{\mathbf{x}}\|_2 + 1), \quad (88)$$

where  $c_{\mathbf{A}} = \frac{1}{n} \sum_{j=1}^n \|\mathbf{a}_j\|_2$ .

Proof: For any  $j \in [n], \ell \in [m]$  and any  $\mathbf{x}_\star, \bar{\mathbf{x}} \in \mathcal{T}$  we have

$$\begin{aligned} \left| \mathbf{a}_j(\mathbf{x}_\star - \bar{\mathbf{x}}) - \tau_j^{(\ell)} \right| &\leq \left| \tau_j^{(\ell)} \right| + |\mathbf{a}_j(\mathbf{x}_\star - \bar{\mathbf{x}})| \\ &\leq \left| \tau_j^{(\ell)} \right| + \|\mathbf{a}_j\|_2 \|\mathbf{x}_\star - \bar{\mathbf{x}}\|_2, \end{aligned} \quad (89)$$

where the last step is derived based on the Cauchy–Schwarz inequality. By averaging the left and the right-hand sides of (??) over all  $j \in [n], \ell \in [m]$  we have

$$\begin{aligned} \frac{1}{m'} \sum_{j,\ell=1}^{m'} \left| \mathbf{a}_j(\mathbf{x}_\star - \bar{\mathbf{x}}) - \tau_j^{(\ell)} \right| &\leq \frac{1}{m'} \sum_{\ell=1}^m \|\boldsymbol{\tau}^{(\ell)}\|_1 + \|\mathbf{x}_\star - \bar{\mathbf{x}}\|_2 \left( \frac{1}{n} \sum_{j=1}^n \|\mathbf{a}_j\|_2 \right) \\ &= \frac{1}{m'} \sum_{\ell=1}^m \|\boldsymbol{\tau}^{(\ell)}\|_1 + c_{\mathbf{A}} \|\mathbf{x}_\star - \bar{\mathbf{x}}\|_2, \end{aligned} \quad (90)$$

where  $T_{\text{ave}}(\mathbf{x}_\star - \bar{\mathbf{x}}) = \frac{1}{m'} \sum_{j,\ell=1}^{m'} \left| \mathbf{a}_j(\mathbf{x}_\star - \bar{\mathbf{x}}) - \tau_j^{(\ell)} \right|$ . Note that  $T_{\text{ave}}(\mathbf{x}_\star - \bar{\mathbf{x}})$  and  $T_{\text{ave}}(\mathbf{x}_\star)$  are related because

$$\begin{aligned} \left| \mathbf{a}_j(\mathbf{x}_\star - \bar{\mathbf{x}}) - \tau_j^{(\ell)} \right| &= \left| \mathbf{a}_j\mathbf{x}_\star - \tau_j^{(\ell)} - \mathbf{a}_j\bar{\mathbf{x}} \right| \\ &\geq \left| \mathbf{a}_j\mathbf{x}_\star - \tau_j^{(\ell)} \right| - |\mathbf{a}_j\bar{\mathbf{x}}| \\ &\geq \left| \mathbf{a}_j\mathbf{x}_\star - \tau_j^{(\ell)} \right| - \|\mathbf{a}_j\|_2 \|\bar{\mathbf{x}}\|_2. \end{aligned} \quad (91)$$

Averaging both sides of (??) over all  $j \in [n], \ell \in [m]$  leads to

$$\frac{1}{m'} \sum_{j,\ell=1}^{m'} \left| \mathbf{a}_j(\mathbf{x}_\star - \bar{\mathbf{x}}) - \tau_j^{(\ell)} \right| \geq \frac{1}{m'} \sum_{j,\ell=1}^{m'} \left| \mathbf{a}_j\mathbf{x}_\star - \tau_j^{(\ell)} \right| - c_{\mathbf{A}} \|\bar{\mathbf{x}}\|_2, \quad (92)$$

which informs  $T_{\text{ave}}(\mathbf{x}_\star) \leq T_{\text{ave}}(\mathbf{x}_\star - \bar{\mathbf{x}}) + c_{\mathbf{A}} \|\bar{\mathbf{x}}\|_2$ . Combining this result with (??) completes the proof.  $\blacksquare$

Based on Lemma ??, to ensure that the event  $T_{\text{ave}} \leq C\rho = \mu + \delta$  implies  $\|\bar{\mathbf{x}} - \mathbf{x}_\star\|_2 \leq \rho$  with a failure probability at most  $e^{-\frac{c_1\delta^2}{K}m'}$ , we should have

$$\delta \ll \frac{1}{m'} \sum_{\ell=1}^m \|\boldsymbol{\tau}^{(\ell)}\|_1 - \mu + c_{\mathbf{A}} (\rho + 1). \quad (93)$$

Now to include all possible  $\mathbf{x}_\star, \bar{\mathbf{x}} \in \mathcal{T}$  that satisfy the uniform perfect reconstruction criterion with  $\bar{\mathbf{x}} \in \mathcal{B}_\rho(\mathbf{x}_\star)$ , we consider a  $\rho$ -net  $\{\bar{\mathbf{x}}_1, \dots, \bar{\mathbf{x}}_K\}$  for the set  $\mathcal{T}$  which means that, for any  $\mathbf{x}_\star \in \mathcal{T}$ , one can find  $k \in [1 : K]$  such that  $\|\mathbf{x}_\star - \bar{\mathbf{x}}_k\|_2 \leq \rho$ . We can take  $K \leq \left(1 + \frac{3}{\rho}\right)^d$  [?] and employ the union bound to obtain the general case of (??) as follows:

$$\begin{aligned}
& \Pr(T_{\text{ave}} \geq C\rho \text{ for some } k \in [1 : K]) \\
& \leq K e^{\frac{-c_1(C\rho-\mu)^2}{K} m'} \\
& \leq \left(1 + \frac{3}{\rho}\right)^d e^{\frac{-c_1(C\rho-\mu)^2}{K} m'} \\
& \leq e^{d \log\left(1 + \frac{3}{\rho}\right) - \frac{c_1(C\rho-\mu)^2}{K} m'} \\
& \leq e^{\frac{3d}{\rho} - \frac{c_1(C\rho-\mu)^2}{K} m'}.
\end{aligned} \tag{94}$$

Considering the bound on  $\delta$  in (??), to achieve a minimum probability of  $1 - \eta$ , it is sufficient to ensure that the upper bound of (??) is lower than  $\eta$  which results in

$$m' \geq \frac{K}{c_1} \delta^{-2} \left( \frac{3d}{\rho} + \log \left( \frac{1}{\eta} \right) \right), \tag{95}$$

Therefore, the constant  $C_1$  in Theorem ?? is  $C_1 = \frac{K}{c_1}$ .

## Appendix D

### Proof of Corollary ??

To prove Corollary ??, we alternatively define the distances in (??) as

$$d_j^{(\ell)}(\mathbf{x}_\star, \tau_j^{(\ell)}) = \left| \mathbf{a}_j \mathbf{x}_\star - \tau_j^{(\ell)} \right|, \quad j \in [n], \ell \in [m]. \tag{96}$$

Note that the definition presented in (??) is exactly the same with the one in (??) owing to the fact that each  $r_j^{(\ell)}$  is either  $+1$  or  $-1$ . The dct coefficients  $\{f_j = \mathbf{a}_j \mathbf{x}_\star\}$  are defined as

$$f_j = \sqrt{\frac{2}{d}} \sum_i^d \Lambda_i \cos \left( \frac{\pi j}{2d} (2i + 1) \right) (\mathbf{x}_\star)_i, \quad j \in [n], \tag{97}$$

where  $\Lambda_i$  is

$$\Lambda_i = \begin{cases} \frac{1}{\sqrt{2}} & \text{for } i = 0, \\ 1 & \text{otherwise.} \end{cases} \tag{98}$$

Note that when the distances defined in (??) are bounded random variables,  $0 \leq d_j^{(\ell)}(\mathbf{x}_\star, \tau_j^{(\ell)}) \leq b$ , we will have  $\Pr(T_{\text{ave}} \geq C\rho) \leq e^{\frac{-2(C\rho-\mu)^2}{b^2} m'}$  [?, Theorem 2]. Therefore, to prove the Corollary ??, we only need to obtain the parameters  $\mu$  and  $b$ . It is easy to verify that

$$b = \max \left\{ \left| \sup_{j \in [n]} f_j + \tilde{b} \right|, \left| \inf_{j \in [n]} f_j - \tilde{b} \right| \right\}. \tag{99}$$



To obtain the first term in (??), we need to compute  $\sup_{j \in [n]} f_j$ . By the assumption  $\|\mathbf{x}_\star\|_2 \leq 1$ , we have  $\|\mathbf{x}_\star\|_1 \leq \sqrt{d}\|\mathbf{x}_\star\|_2 \leq \sqrt{d}$ . Therefore,  $\sup_{j \in [n]} f_j$  can be computed as

$$\begin{aligned} f_j &= \sqrt{\frac{2}{d}} \sum_i^d \Lambda_i \cos\left(\frac{\pi j}{2d}(2i+1)\right) (\mathbf{x}_\star)_i \\ &\leq \sqrt{\frac{2}{d}} \sum_i^d \Lambda_i (\mathbf{x}_\star)_i \leq \sqrt{\frac{2}{d}} \sum_i^d (\mathbf{x}_\star)_i \\ &\leq \sqrt{\frac{2}{d}} \sum_i^d |(\mathbf{x}_\star)_i| = \sqrt{\frac{2}{d}} \|\mathbf{x}_\star\|_1 \leq \sqrt{2}. \end{aligned} \quad (100)$$

Therefore,  $\sup_{j \in [n]} f_j = \sqrt{2}$ . Similarly, we can obtain  $\inf_{j \in [n]} f_j = -\sqrt{2}$ . Based on (??), we have  $b = \sqrt{2} + \tilde{b}$ . For simplicity, denote  $d_j^{(\ell)}(\mathbf{x}_\star, \tau_j^{(\ell)})$  by  $d_j^{(\ell)}$ . To obtain  $\mu$ , we should compute

$$\begin{aligned} \mu &= \mathbb{E} \left\{ \frac{1}{m'} \sum_{j,\ell=1}^{m'} d_j^{(\ell)} \right\} = \mathbb{E} \left\{ \frac{1}{m'} \sum_{j,\ell=1}^{m'} |f_j - \tau_j^{(\ell)}| \right\} \\ &= \mathbb{E} \left\{ \frac{1}{n} \sum_{j=1}^n |f_j - \tau_j^{(\ell)}| \right\}. \end{aligned} \quad (101)$$

Based on (??), we should consider three scenarios: (i)  $-\tilde{b} \leq f_j \leq \tilde{b}$ , (ii)  $f_j \geq \tilde{b}$ , and (iii)  $f_j \leq -\tilde{b}$ . For simplicity, denote  $\tau_j^{(\ell)} = \tau$ . Under scenario (i),  $\mathbb{E}\{|f_j - \tau|\}$  can be computed as:

$$\begin{aligned} \mathbb{E}\{|f_j - \tau|\} &= \frac{1}{2\tilde{b}} \int_{-\tilde{b}}^{\tilde{b}} |f_j - \tau| d\tau \\ &= \frac{1}{2\tilde{b}} \left[ \int_{-\tilde{b}}^{f_j} f_j - \tau d\tau + \int_{f_j}^{\tilde{b}} \tau - f_j d\tau \right] = \frac{1}{2\tilde{b}} (\tilde{b}^2 + f_j^2). \end{aligned} \quad (102)$$

Under scenario (ii), we have

$$\mathbb{E}\{|f_j - \tau|\} = \frac{1}{2\tilde{b}} \int_{-\tilde{b}}^{\tilde{b}} |f_j - \tau| d\tau = \frac{1}{2\tilde{b}} \int_{-\tilde{b}}^{\tilde{b}} f_j - \tau d\tau = f_j, \quad (103)$$

and in scenario (iii)

$$\mathbb{E}\{|f_j - \tau|\} = \frac{1}{2\tilde{b}} \int_{-\tilde{b}}^{\tilde{b}} |f_j - \tau| d\tau = \frac{1}{2\tilde{b}} \int_{-\tilde{b}}^{\tilde{b}} \tau - f_j d\tau = -f_j. \quad (104)$$

Combining (??), (??) and (??) with (??) leads to (??) which completes the proof.

## Appendix E

### Proof of Corollary ??

To prove Corollary ??, assume  $\|\mathbf{x}_\star\|_2^2 = M' \leq 1$ . Then, it can be easily derived that  $\mathbf{a}_j \mathbf{x}_\star \sim \mathcal{N}(0, M')$ . Due to the independency of  $\mathbf{a}_j \mathbf{x}_\star$  and  $\tau_j^{(\ell)}$ ,  $(\mathbf{a}_j \mathbf{x}_\star - \tau_j^{(\ell)})$  is distributed as  $(\mathbf{a}_j \mathbf{x}_\star - \tau_j^{(\ell)}) \sim \mathcal{N}(0, \alpha^2)$ , where  $\alpha^2 = M' + \sigma^2$ . Consequently, each distance in (??) follows the folded normal distribution characterized by a location parameter zero and a scale parameter  $\alpha^2$ . To complete the proof of Corollary ??, our remaining task is to determine the parameters  $\mu$  and  $K$  based on the proof of Theorem ?? for each distance stated in (??). It can be easily confirmed that, for each distance mentioned in (??) with the folded normal distribution as previously described, the value of  $\mu$  satisfies

$$\mu = \alpha \sqrt{\frac{2}{\pi}} = \sqrt{\frac{2}{\pi} (\sigma^2 + M')} \leq \sqrt{\frac{2}{\pi} (\sigma^2 + 1)}. \quad (105)$$

To find  $K$ , we need to obtain the sub-Gaussian norm of the folded normal random variable. To achieve this, we begin by expressing the sub-Gaussian norm of the Gaussian random variable  $X$  characterized by  $X \sim \mathcal{N}(0, \alpha^2)$ :

$$\|X\|_{\psi_2} = \inf \left\{ t > 0 : \mathbb{E} \left\{ e^{X^2/t^2} \right\} \leq 2 \right\}, \quad (106)$$

which leads to

$$\begin{aligned} \|X\|_{\psi_2} &= \inf \left\{ t > 0 : \frac{1}{\sqrt{2\pi\alpha^2}} \int_{-\infty}^{\infty} e^{x^2/t^2} e^{-x^2/(2\alpha^2)} dx \leq 2 \right\} \\ &= \inf \left\{ t > 0 : \sqrt{\frac{2}{\pi\alpha^2}} \int_0^{\infty} e^{x^2/t^2} e^{-x^2/(2\alpha^2)} dx \leq 2 \right\}. \end{aligned} \quad (107)$$

In [?, Page 28], it is presented that solving (??) leads to  $\|X\|_{\psi_2} \leq c_2 \alpha$  for a positive value  $c_2$ . To obtain the value of  $c_2$  that satisfies  $\|X\|_{\psi_2} = c_2 \alpha$ , we should find  $\inf c_2 > 0$  such that

$$\frac{1}{\sqrt{2\pi\alpha^2}} \int_{-\infty}^{\infty} e^{x^2/c_2^2 \alpha^2} e^{-x^2/(2\alpha^2)} dx \leq 2, \quad (108)$$

or equivalently

$$\frac{1}{\sqrt{2\pi\alpha^2}} \int_{-\infty}^{\infty} e^{-x^2/(2\alpha^2)[1-2/c_2^2]} dx \leq 2. \quad (109)$$

Set  $\beta^2 = 1 - 2/c_2^2$  and  $u = \beta x$ , we can rewrite (??) as

$$\frac{1}{\beta} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\alpha^2}} e^{-u^2/(2\alpha^2)} du \leq 2, \quad (110)$$

which results in  $c_2 \geq \sqrt{\frac{8}{3}}$ . Therefore, we have  $\|X\|_{\psi_2} = \sqrt{\frac{8}{3}} \alpha$ . The probability density function (PDF) of the folded normal random variable  $|X|$  with a location parameter zero

and a scale parameter  $\alpha^2$  can be presented as  $f_X(x) = \sqrt{\frac{2}{\pi\alpha^2}}e^{-x^2/(2\alpha^2)}$  for  $x \geq 0$ . Then, based on the definition of the sub-Gaussian norm, we have

$$\|X\|_{\psi_2} = \inf \left\{ t > 0 : \sqrt{\frac{2}{\pi\alpha^2}} \int_0^\infty e^{x^2/t^2} e^{-x^2/(2\alpha^2)} dx \leq 2 \right\}, \quad (111)$$

which is the same as the result we have obtained in (??). As a result, the sub-Gaussian norm of the folded normal random variable  $|X|$  is  $\|X\|_{\psi_2} = \sqrt{\frac{8}{3}}\alpha$ . By setting  $C_1 = \frac{8}{3c_1}$  and combining this result with (??), the proof is now complete.

## Appendix F

### Proof of Theorem ??

For simplicity denote  $d_j^{(\ell)s}(\mathbf{x}_\star, \tau_j^{(\ell)})$  by  $d_j^{(\ell)s}$ . Define the average of such distances as

$$T_{\text{ave}}^s = \frac{1}{m'} \sum_{j,\ell=1}^{m'} d_j^{(\ell)s}. \quad (112)$$

Assume that  $\mathbb{E}\{T_{\text{ave}}^s\} = \mu_s$  and  $\left\|d_j^{(\ell)s}\right\|_{\psi_2}^2 \leq K_s$ . Fix an index set  $S$  of size  $s$  and identify the space  $\Sigma_S = \{\mathbf{x} \in \mathbb{R}^n : \text{supp}(\mathbf{x}) \subseteq S, \|\mathbf{x}\|_2 \leq 1\}$  with  $\mathbb{R}^s$ . Then our goal is to obtain the required number of hyperplanes (one-bit data) to achieve the uniform perfect reconstruction criterion with  $\mathbf{x}' \in \mathcal{B}_\rho(\mathbf{x}_\star)$  considering all  $\mathbf{x}_\star, \mathbf{x}' \in \Sigma_S$ . Based on this criterion and the conclusion made in Lemma ??, to see the connection between  $\|\mathbf{x}' - \mathbf{x}_\star\|_2 \leq \rho_s$  and  $T_{\text{ave}}^s$ , there exists a set  $\mathcal{G}_s$  with the minimum cardinality  $\text{card}(\mathcal{G}_s) \sim s$  which contains  $\text{card}(\mathcal{G}_s)$  minimum distances  $d_j^{(\ell)s}$  such that  $\frac{1}{\text{card}(\mathcal{G}_s)} \sum_{d_j^{(\ell)s} \in \mathcal{G}_s} d_j^{(\ell)s} \sim \rho_s$ . Existing such a set  $\mathcal{G}_s$  with high probability implies  $\|\mathbf{x}' - \mathbf{x}_\star\|_2 \leq \rho_s$ . For the complement of the set  $\mathcal{G}_s$  denoted by  $\bar{\mathcal{G}}_s$ , we have  $d_j^{(\ell)s} \geq \rho_s$  for all  $d_j^{(\ell)s} \in \bar{\mathcal{G}}_s$  which leads to  $\frac{1}{\text{card}(\bar{\mathcal{G}}_s)} \sum_{d_j^{(\ell)s} \in \bar{\mathcal{G}}_s} d_j^{(\ell)s} \geq \rho_s$ . Therefore, we can conclude that  $T_{\text{ave}}^s \geq \rho_s$ . To meet the condition  $\|\mathbf{x}' - \mathbf{x}_\star\|_2 \leq \rho_s$ ,  $\exists C > 0$  such that  $T_{\text{ave}}^s \leq C\rho_s$ , where  $C\rho_s = \mu_s + \delta$  with a positive constant  $\delta$ . To derive the probability of the event  $\|\mathbf{x}' - \mathbf{x}_\star\|_2 \leq \rho_s$  for a specific  $\mathbf{x}_\star \in \Sigma_S$ , we apply the general Hoeffding's inequality [?, Theorem 2.6.2] to the event  $T_{\text{ave}}^s \leq C\rho_s$  as follows:

$$\Pr(T_{\text{ave}}^s \geq C\rho_s \geq C\|\mathbf{x}' - \mathbf{x}_\star\|_2) \leq e^{\frac{-c_1(C\rho_s - \mu_s)^2}{K_s}m'}. \quad (113)$$

Based on Lemma ??, to ensure that the event  $T_{\text{ave}}^s \leq C\rho_s = \mu_s + \delta$  implies  $\|\bar{\mathbf{x}} - \mathbf{x}_\star\|_2 \leq \rho_s$  with a failure probability at most  $e^{\frac{-c_1\delta^2}{K_s}m'}$ , we should have

$$\delta \ll \frac{1}{m'} \sum_{\ell=1}^m \|\boldsymbol{\tau}^{(\ell)}\|_1 - \mu_s + c_{\mathbf{A}_s}(\rho_s + 1). \quad (114)$$

Now to include all possible  $\mathbf{x}_*, \mathbf{x}' \in \Sigma_S$  that satisfy the uniform perfect reconstruction criterion, we consider a  $\rho_s$ -net  $\{\mathbf{x}'_1, \dots, \mathbf{x}'_K\}$  for the set  $\Sigma_S$  which means that, for any  $\mathbf{x}_* \in \Sigma_S$ , one can find  $k \in [1 : K]$  such that  $\|\mathbf{x}_* - \mathbf{x}'_k\|_2 \leq \rho_s$ . We can take  $K \leq \left(1 + \frac{3}{\rho_s}\right)^s$  and employ the union bound to derive the general case of (??) as follows:

$$\begin{aligned}
& \Pr(T_{\text{ave}}^s \geq C\rho_s \text{ for some } k \in [1 : K]) \\
& \leq K e^{\frac{-c_1(C\rho_s - \mu_s)^2}{K_s} m'} \\
& \leq \left(1 + \frac{3}{\rho}\right)^s e^{\frac{-c_1(C\rho_s - \mu_s)^2}{K_s} m'} \\
& \leq e^{s \log\left(1 + \frac{3}{\rho}\right) - \frac{c_1(C\rho_s - \mu_s)^2}{K_s} m'} \\
& \leq e^{\frac{3s}{\rho_s} - \frac{c_1(C\rho_s - \mu_s)^2}{K_s} m'}.
\end{aligned} \tag{115}$$

The probability mentioned earlier corresponds to a single possibility of a fixed index set  $S$ . However, to account for all possible  $s$ -sparse solutions, we unfix the index set  $S$  and employ the union bound and express the general case of (??) as follows:

$$\begin{aligned}
& \Pr(T_{\text{ave}}^s \geq C\rho_s \text{ for some } S \subseteq [1 : n] \text{ with } \text{card}(S) = s) \\
& = \binom{d}{s} e^{\frac{3s}{\rho_s} - \frac{c_1(C\rho_s - \mu_s)^2}{K_s} m'} \\
& \leq e^{s \log\left(\frac{ed}{s}\right) + \frac{3}{\rho_s} s - \frac{c_1(C\rho_s - \mu_s)^2}{K_s} m'} \\
& = e^{s \left(\log\left(\frac{ed}{s}\right) + \frac{3}{\rho_s}\right) - \frac{c_1(C\rho_s - \mu_s)^2}{K_s} m'},
\end{aligned} \tag{116}$$

where we have used the inequality  $\binom{d}{s} \leq \left(\frac{ed}{s}\right)^s$ . Considering the bound on  $\delta$  in (??), to achieve a minimum probability of  $1 - \eta$ , it is sufficient to ensure that the upper bound of (??) is lower than  $\eta$  which results in

$$m' \geq \frac{K_s}{c_1} \delta^{-2} \left( \log\left(\frac{1}{\eta}\right) + s \left( \log\left(\frac{ed}{s}\right) + \frac{3}{\rho_s} \right) \right). \tag{117}$$

Therefore, the constant  $\delta_s$  in Theorem ?? is  $\delta_s = \frac{K_s}{c_1} \delta^{-2}$ .

## Appendix G

### Proof of Theorem ??

Define the set  $\mathcal{K}_{n_1, r}$  as

$$\mathcal{K}_{n_1, r} = \left\{ \mathbf{X} \in \mathbb{R}^{n_1 \times n_1} \mid \text{rank}(\mathbf{X}) \leq r, \|\mathbf{X}\|_F \leq 1 \right\}. \tag{118}$$

Assume  $\mathbf{X}_\star \in \mathcal{K}_{n_1, r}$ . Based on our formulations in Section ??, define the distance  $d_j^{(\ell)r}$  as

$$d_j^{(\ell)r} = \left| \text{Tr}(\mathbf{A}_j^\top \mathbf{X}_\star) - \tau_j^{(\ell)} \right|, \quad j \in [n], \quad \ell \in [m], \quad (119)$$

where the average of such distances is formulated as

$$T_{\text{ave}}^r = \frac{1}{m'} \sum_{j, \ell=1}^{m'} d_j^{(\ell)r}. \quad (120)$$

Assume that  $\mathbb{E}\{T_{\text{ave}}^r\} = \mu_r$  and  $\left\|d_j^{(\ell)r}\right\|_{\psi_2}^2 \leq K_r$ . Our goal is to obtain the required number of hyperplanes to achieve the uniform perfect reconstruction criterion with  $\mathbf{X}' \in \mathcal{B}_\rho(\text{vec}(\mathbf{X}_\star))$  considering all  $\mathbf{X}_\star, \mathbf{X}' \in \mathcal{K}_{n_1, r}$ . Based on this criterion and the conclusion made in Lemma ??, we establish the connection between  $\|\mathbf{X}' - \mathbf{X}_\star\|_{\text{F}} \leq \rho_r$  and  $T_{\text{ave}}^r$  by identifying a set  $\mathcal{G}_r$  with a minimum cardinality  $\text{card}(\mathcal{G}_r) \sim n_1 r$ , containing  $\text{card}(\mathcal{G}_r)$  minimum distances  $d_j^{(\ell)r}$  such that  $\frac{1}{\text{card}(\mathcal{G}_r)} \sum_{d_j^{(\ell)r} \in \mathcal{G}_r} d_j^{(\ell)r} \sim \rho_r$ . Existing such a set  $\mathcal{G}_r$  with high probability implies  $\|\mathbf{X}' - \mathbf{X}_\star\|_{\text{F}} \leq \rho_r$ . For the complement of the set  $\mathcal{G}_r$  denoted by  $\bar{\mathcal{G}}_r$ , we have  $d_j^{(\ell)r} \geq \rho_r$  for all  $d_j^{(\ell)r} \in \bar{\mathcal{G}}_r$  which leads to  $\frac{1}{\text{card}(\bar{\mathcal{G}}_r)} \sum_{d_j^{(\ell)r} \in \bar{\mathcal{G}}_r} d_j^{(\ell)r} \geq \rho_r$ . Consequently, we can conclude that  $T_{\text{ave}}^r \geq \rho_r$ . To meet  $\|\mathbf{X}' - \mathbf{X}_\star\|_{\text{F}} \leq \rho_r$ ,  $\exists C > 0$  such that  $T_{\text{ave}}^r \leq C\rho_r$ , where  $C\rho_r = \mu_r + \delta$  with a positive constant  $\delta$ . Note that in the probabilistic sense, to derive the probability of the event  $\|\mathbf{X}' - \mathbf{X}_\star\|_{\text{F}} \leq \rho_r$  for a specific  $\mathbf{X}_\star \in \mathcal{K}_{n_1, r}$  we can utilize the probability of  $T_{\text{ave}}^r \leq C\rho_r$ . Therefore, by the general Hoeffding's inequality [?, Theorem 2.6.2], we have

$$\Pr(T_{\text{ave}}^r \geq C\rho_r \geq C\|\mathbf{X}' - \mathbf{X}_\star\|_{\text{F}}) \leq e^{-\frac{c_1(C\rho_r - \mu_r)^2}{K_r} m'}. \quad (121)$$

Based on Lemma ??, to ensure that the event  $T_{\text{ave}}^r \leq C\rho_r = \mu_r + \delta$  implies  $\|\mathbf{X}' - \mathbf{X}_\star\|_{\text{F}} \leq \rho_r$  with a failure probability at most  $e^{-\frac{c_1 \delta^2}{K_r} m'}$ , we should have

$$\delta \ll \frac{1}{m'} \sum_{\ell=1}^m \|\boldsymbol{\tau}^{(\ell)}\|_1 - \mu_r + c_{\mathbf{V}}(\rho_r + 1), \quad (122)$$

where  $\mathbf{V}$  is defined in Lemma ??. Now to include all possible  $\mathbf{X}_\star, \mathbf{X}' \in \mathcal{K}_{n_1, r}$  that satisfy the uniform perfect reconstruction criterion, we consider a  $\rho_r$ -net  $\{\mathbf{X}'_1, \dots, \mathbf{X}'_K\}$  for the set  $\mathcal{K}_{n_1, r}$  which means that, for any  $\mathbf{X}_\star \in \mathcal{K}_{n_1, r}$ , one can find  $k \in [1 : K]$  such that  $\|\mathbf{X}_\star - \mathbf{X}'_k\|_{\text{F}} \leq \rho_r$ .

According to [?, Lemma 3.1], we can take  $K \leq \left(1 + \frac{6}{\rho_r}\right)^{(2n_1+1)r}$ . Then by employing the union bound, the general case of (??) can be expressed as

$$\begin{aligned}
& \Pr(T_{\text{ave}}^r \geq C\rho_r \text{ for some } k \in [1 : K]) \\
& \leq K e^{\frac{-c_1(C\rho_r - \mu_r)^2}{K_r} m'} \\
& \leq \left(1 + \frac{6}{\rho_r}\right)^{(2n_1+1)r} e^{\frac{-c_1(C\rho_r - \mu_r)^2}{K_r} m'} \\
& \leq e^{\frac{18n_1 r}{\rho_r} - \frac{c_1(C\rho_r - \mu_r)^2}{K_r} m'}.
\end{aligned} \tag{123}$$

Considering the bound on  $\delta$  in (??), to achieve a minimum probability of  $1 - \eta$ , it is sufficient to ensure that the upper bound of (??) is lower than  $\eta$  which results in

$$m' \geq \frac{K_r}{c_1} \delta^{-2} \left( \frac{18n_1 r}{\rho_r} + \log \left( \frac{1}{\eta} \right) \right). \tag{124}$$

Therefore, the constant  $\delta_r$  in Theorem ?? is  $\delta_r = \frac{K_r}{c_1} \delta^{-2}$ .

## Appendix H

### Proof of Lemma ??

Consider the following lemma:

Lemma 10. [?, Section 2.5.2] Let  $X_1, X_2, \dots, X_n$  be a sequence of sub-Gaussian random variables with  $K = \max_i \|X_i\|_{\psi_2}$ . Then for  $n \geq 2$  we have

$$\mathbb{E} \left\{ \max_{i \leq n} |X_i| \right\} \leq CK \sqrt{\log n}, \tag{125}$$

where  $C$  is an absolute constant.

Denote the uniform samples of  $\mathbf{y}$  by  $\{y_j\}_{j=1}^n$ . The DR of the high-resolution measurements is defined as  $\text{DR}_{\mathbf{y}} = \|\mathbf{y}\|_{\infty}$ . Based on Lemma ??, the expected value of the random variable  $\text{DR}_{\mathbf{y}}$  is bounded by

$$\begin{aligned}
\mathbb{E} \{ \text{DR}_{\mathbf{y}} = \|\mathbf{y}\|_{\infty} \} &= \mathbb{E} \{ \|\mathbf{A}\mathbf{x}\|_{\infty} \} \\
&= \mathbb{E} \left\{ \max_{j \leq n} |\mathbf{a}_j \mathbf{x}| \right\} \\
&\leq CK \|\mathbf{x}\|_2 \sqrt{\log n} \\
&\leq CKR \sqrt{\log n},
\end{aligned} \tag{126}$$

where  $K = \max_j \|\mathbf{a}_j \mathbf{x}\|_{\psi_2}$ . Similar to this, we can obtain the bound on the DR of Gaussian dither as  $\mathbb{E} \{ \max_{j \leq n} |\tau_j^g| \} \leq C \sqrt{\frac{8}{3}} \sigma \sqrt{\log n}$ . For the DR of the Uniform dither we have

$\mathbb{E} \left\{ \max_{j \leq n} |\tau_j^u| \right\} \leq C \frac{\tilde{b}}{\log(2)} \sqrt{\log n}$  [?, Section 2.5.2]. Therefore, by setting  $\sigma = \mathcal{O}(\sqrt{\frac{3}{8}}KR)$  in the Gaussian dither and  $\tilde{b} = \mathcal{O}(\log(2)KR)$  in the Uniform dither, we can guarantee that the DR of high-resolution measurements  $\mathbf{y}$  can be covered with a high probability.

## Appendix I

### Proof of Theorem ??

Assume that  $\{\sigma_{i\mathbf{P}}\}$  and  $\{\sigma_{i\mathbf{A}}\}$  denote the singular values of the matrices  $\mathbf{P}$  and  $\mathbf{A}$ , respectively. To obtain the singular values of  $\mathbf{P}$ , the matrix  $\mathbf{W} = \mathbf{P}^\top \mathbf{P}$  is computed as

$$\begin{aligned} \mathbf{W} &= \mathbf{P}^\top \mathbf{P} \\ &= \left[ \mathbf{A}^\top \boldsymbol{\Omega}^{(1)} \mid \dots \mid \mathbf{A}^\top \boldsymbol{\Omega}^{(m)} \right] \left[ \mathbf{A}^\top \boldsymbol{\Omega}^{(1)} \mid \dots \mid \mathbf{A}^\top \boldsymbol{\Omega}^{(m)} \right]^\top \\ &= \mathbf{A}^\top \boldsymbol{\Omega}^{(1)} \boldsymbol{\Omega}^{(1)} \mathbf{A} + \dots + \mathbf{A}^\top \boldsymbol{\Omega}^{(m)} \boldsymbol{\Omega}^{(m)} \mathbf{A} \\ &= m \mathbf{A}^\top \mathbf{I} \mathbf{A} = m \mathbf{A}^\top \mathbf{A}, \end{aligned} \quad (127)$$

which means that the singular values of  $\mathbf{P}$  are  $\{\sigma_{i\mathbf{P}}\} = \sqrt{m} \{\sigma_{i\mathbf{A}}\}$ . Also, the Frobenius norm of  $\mathbf{P}$  is obtained as

$$\begin{aligned} \|\mathbf{P}\|_F^2 &= \text{Tr}(\mathbf{P}^\top \mathbf{P}) \\ &= \text{Tr}(m \mathbf{A}^\top \mathbf{A}) = m \|\mathbf{A}\|_F^2. \end{aligned} \quad (128)$$

Plugging in (??) and (??) in the definition of the scaled condition number provided in Section ??, one can conclude that  $\kappa(\mathbf{P}) = \kappa(\mathbf{A})$ .

## Appendix J

### Proof of Corollary ??

Based on the convergence rate of ORKA in the sample abundance scenario (??), the upper recovery bound of ORKA  $\mathbb{E} \left\{ \|\mathbf{x}_{i+1} - \mathbf{x}_\star\|_2^2 \right\} \leq \epsilon_0$  can be met when  $q^i \omega_0 + \rho^2 \leq \epsilon_0$ , where  $q = 1 - \kappa^{-2}(\mathbf{A})$ . This leads to

$$i \geq \frac{\log(\epsilon_0 - \rho^2)}{\log(1 - \kappa^{-2}(\mathbf{A}))}. \quad (129)$$

The right-hand side of (??) can be approximated by [?]:

$$\frac{\log(\epsilon_0 - \rho^2)}{\log(1 - \kappa^{-2}(\mathbf{A}))} \approx \kappa^2(\mathbf{A}) \log \left( \frac{1}{\epsilon_0 - \rho^2} \right). \quad (130)$$

For a well conditioned sensing matrix  $\mathbf{A}$ , i.e.  $\varrho(\mathbf{A}) = \mathcal{O}(1)$ , we have  $\kappa^2(\mathbf{A}) = \mathcal{O}(d)$ . Therefore, the required number of iterations for ORKA to meet the aforementioned upper recovery bound is  $i = \mathcal{O} \left( d \log \left( \frac{1}{\epsilon_0 - \rho^2} \right) \right)$ .

## Appendix K

### Proof of Proposition ??

We begin the proof by presenting the following lemma:

Lemma 11. Let  $\mathcal{H}_j = \{\mathbf{x} : \mathbf{c}_j \mathbf{x} \geq b_j - n_j\}$  be the solution spaces of the noisy linear inequalities. Let  $\bar{\mathcal{H}}_j = \{\mathbf{x} : \mathbf{c}_j \mathbf{x} \geq b_j\}$  be the solution spaces of the system run by the noisy RKA. Then  $\bar{\mathcal{H}}_j = \{\mathbf{x} + \alpha_j \mathbf{c}_j^H : \mathbf{x} \in \mathcal{H}_j\}$ , where  $\alpha_j = \frac{n_j}{\|\mathbf{c}_j\|_2^2}$ .

Proof: Assume  $\mathbf{x} \in \mathcal{H}_j$ , then we can write

$$\mathbf{c}_j (\mathbf{x} + \alpha_j \mathbf{c}_j^H) \geq b_j - n_j + \alpha_j \|\mathbf{c}_j\|_2^2. \quad (131)$$

By plugging  $\alpha_j = \frac{n_j}{\|\mathbf{c}_j\|_2^2}$  in (??), we can obtain

$$\mathbf{c}_j (\mathbf{x} + \alpha_j \mathbf{c}_j^H) \geq b_j, \quad (132)$$

which implies  $\mathbf{x} + \alpha_j \mathbf{c}_j^H \in \bar{\mathcal{H}}_j$ . ■

Assume  $\bar{\mathbf{x}}_i$  denotes the  $i$ -th iterate of the noisy RKA run with  $\mathbf{C}\mathbf{x} \succeq \mathbf{b}$ . Denote  $\hat{\mathbf{x}} \in \cap_{j=1}^m \mathcal{H}_j$ . Based on Lemma ??, we can write

$$\bar{\mathbf{x}}_i - \hat{\mathbf{x}} = \mathbf{x}_i - \hat{\mathbf{x}} + \alpha_j \mathbf{c}_j^H. \quad (133)$$

where  $\mathbf{x}_i \in \mathcal{H}_j$ . For  $\mathbf{x}_i \in \mathcal{H}_j$ , we have  $\mathbf{c}_j \mathbf{x}_i \geq b_j - n_j$ . Then, assuming  $n_j \leq 0$ , we still have  $\mathbf{c}_j \mathbf{x}_i \geq b_j$  which informs  $\mathcal{H}_j \subseteq \bar{\mathcal{H}}_j$  and there is no inequality flip in the  $i$ -th iteration of the noisy RKA. As a result, we modify  $\alpha_j$  in (??) to  $(\alpha_j)^+$  to take into account our above discussion. Taking the norm-2 of (??) leads to

$$\begin{aligned} \|\bar{\mathbf{x}}_i - \hat{\mathbf{x}}\|_2^2 &= \|\mathbf{x}_i - \hat{\mathbf{x}} + (\alpha_j)^+ \mathbf{c}_j^H\|_2^2 \\ &\leq \|\mathbf{x}_i - \hat{\mathbf{x}}\|_2^2 + \|(\alpha_j)^+ \mathbf{c}_j^H\|_2^2 \\ &= \|\mathbf{x}_i - \hat{\mathbf{x}}\|_2^2 + \frac{((n_j)^+)^2}{\|\mathbf{c}_j\|_2^2}. \end{aligned} \quad (134)$$

Denote the scaled condition number of the matrix  $\mathbf{C}$  by  $\kappa(\mathbf{C}) = \sqrt{R}$ . Then by considering the convergence rate of the RKA in (??) and drawing inspiration from [?, Lemma 2.2] in the context of linear system of inequalities, we have

$$\begin{aligned} \mathbb{E} \{\|\bar{\mathbf{x}}_i - \hat{\mathbf{x}}\|_2^2\} &\leq \left(1 - \frac{1}{R}\right) \|\bar{\mathbf{x}}_{i-1} - \hat{\mathbf{x}}\|_2^2 + \frac{((n_j)^+)^2}{\|\mathbf{c}_j\|_2^2} \\ &\leq \left(1 - \frac{1}{R}\right) \|\bar{\mathbf{x}}_{i-1} - \hat{\mathbf{x}}\|_2^2 + \max_j \gamma_j, \end{aligned} \quad (135)$$



where  $\gamma_j$  is defined in Proposition ??, and the expectation is conditioned upon the choice of the random selections in the first  $i - 1$  iterations. Then applying this recursive relation iteratively and taking full expectation, we have

$$\begin{aligned} \mathbb{E} \{ \|\bar{\mathbf{x}}_i - \hat{\mathbf{x}}\|_2^2 \} &\leq \left(1 - \frac{1}{R}\right)^i \|\mathbf{x}_0 - \hat{\mathbf{x}}\|_2^2 \\ &\quad + \sum_{k=0}^i \left(1 - \frac{1}{R}\right)^k \max_j \gamma_j \\ &\leq \left(1 - \frac{1}{R}\right)^i \|\mathbf{x}_0 - \hat{\mathbf{x}}\|_2^2 + R \max_j \gamma_j, \end{aligned} \quad (136)$$

which completes the proof.

## Appendix L

### Convergence analysis of ST-ORKA, SVP-ORKA, and HT-ORKA

In this section, we prove the convergence, and derive the convergence rates of ST-ORKA, SVP-ORKA, and HT-ORKA.

#### A. Convergence Analysis for ST-ORKA

As discussed in [?], if  $\Omega_c$  is a closed and convex set in any Hilbert space  $\mathcal{H}$ , then the projection operator  $P : \mathcal{H} \rightarrow \Omega_c$  is non-expansive, i.e., for any vectors  $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{H}$ ,  $\|P(\mathbf{x}_1) - P(\mathbf{x}_2)\|_2 \leq \|\mathbf{x}_1 - \mathbf{x}_2\|_2$ . Due to the fact that the ST operator projects the signal to  $\|\mathbf{x}\|_1 \leq \epsilon$  which is a closed and convex set, we can write that for any vectors  $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{H}$ ,  $\|S_\kappa(\mathbf{x}_1) - S_\kappa(\mathbf{x}_2)\|_2 \leq \|\mathbf{x}_1 - \mathbf{x}_2\|_2$ .

Define  $\mathbf{e}_i = \mathbf{x}_i - \hat{\mathbf{x}}$ ,  $t_j^{(\ell)} = r_j^{(\ell)} \tau_j^{(\ell)}$  and  $\mathbf{p}_j^{(\ell)} = r_j^{(\ell)} \mathbf{a}_j$ , we derive the convergence rate of ST-ORKA as

$$\begin{aligned} \|\mathbf{e}_{i+1}\|_2^2 &= \|S_\kappa(\mathbf{z}_{i+1}) - S_\kappa(\hat{\mathbf{x}})\|_2^2 \\ &\leq \|\mathbf{z}_{i+1} - \hat{\mathbf{x}}\|_2^2 \\ &= \left\| \mathbf{e}_i + \frac{\left(t_j^{(\ell)} - \mathbf{p}_j^{(\ell)} \mathbf{x}_i\right)^+}{\|\mathbf{p}_j^{(\ell)}\|_2^2} \mathbf{p}_j^{(\ell)\text{H}} \right\|_2^2 \\ &= \|\mathbf{e}_i\|_2^2 + \frac{\left(\left(t_j^{(\ell)} - \mathbf{p}_j^{(\ell)} \mathbf{x}_i\right)^+\right)^2}{\|\mathbf{p}_j^{(\ell)}\|_2^2} + \frac{2 \left(t_j^{(\ell)} - \mathbf{p}_j^{(\ell)} \mathbf{x}_i\right)^+ \mathbf{p}_j^{(\ell)} \mathbf{e}_i}{\|\mathbf{p}_j^{(\ell)}\|_2^2}. \end{aligned} \quad (137)$$

Since  $\mathbf{p}_j^{(\ell)} \widehat{\mathbf{x}} \geq t_j^{(\ell)}$ , we have  $\mathbf{p}_j^{(\ell)} \mathbf{e}_i = \mathbf{p}_j^{(\ell)} (\mathbf{x}_i - \widehat{\mathbf{x}}) \leq \mathbf{p}_j^{(\ell)} \mathbf{x}_i - t_j^{(\ell)}$ . Therefore, one can rewrite (??) as

$$\begin{aligned}
\|\mathbf{e}_{i+1}\|_2^2 &\leq \|\mathbf{e}_i\|_2^2 + \frac{\left(\left(t_j^{(\ell)} - \mathbf{p}_j^{(\ell)} \mathbf{x}_i\right)^+\right)^2}{\|\mathbf{p}_j^{(\ell)}\|_2^2} + \frac{2\left(t_j^{(\ell)} - \mathbf{p}_j^{(\ell)} \mathbf{x}_i\right)^+ \mathbf{p}_j^{(\ell)}}{\|\mathbf{p}_j^{(\ell)}\|_2^2} \mathbf{e}_i \\
&\leq \|\mathbf{e}_i\|_2^2 + \frac{\left(\left(t_j^{(\ell)} - \mathbf{p}_j^{(\ell)} \mathbf{x}_i\right)^+\right)^2}{\|\mathbf{p}_j^{(\ell)}\|_2^2} - 2 \frac{\left(t_j^{(\ell)} - \mathbf{p}_j^{(\ell)} \mathbf{x}_i\right)^+ \left(t_j^{(\ell)} - \mathbf{p}_j^{(\ell)} \mathbf{x}_i\right)}{\|\mathbf{p}_j^{(\ell)}\|_2^2} \\
&= \|\mathbf{e}_i\|_2^2 - \frac{\left(\left(t_j^{(\ell)} - \mathbf{p}_j^{(\ell)} \mathbf{x}_i\right)^+\right)^2}{\|\mathbf{p}_j^{(\ell)}\|_2^2}.
\end{aligned} \tag{138}$$

Define  $\mathbf{t} = \left[ \mathbf{t}_1^\top \mid \dots \mid \mathbf{t}_m^\top \right]^\top$ , where  $\mathbf{t}_\ell = \left[ t_j^{(\ell)} \right]_{j=1}^n$  for  $\ell \in [m]$ . Taking the expectation from both sides of (??) results in

$$\begin{aligned}
\mathbb{E} \{ \|\mathbf{e}_{i+1}\|_2^2 \} &\leq \|\mathbf{e}_i\|_2^2 - \sum_{j,\ell=1}^{m'} \frac{\|\mathbf{p}_j^{(\ell)}\|_2^2}{\|\mathbf{P}\|_F^2} \frac{\left(\left(t_j^{(\ell)} - \mathbf{p}_j^{(\ell)} \mathbf{x}_i\right)^+\right)^2}{\|\mathbf{p}_j^{(\ell)}\|_2^2} \\
&= \|\mathbf{e}_i\|_2^2 - \frac{\|(\mathbf{t} - \mathbf{P} \mathbf{x}_i)^+\|_2^2}{\|\mathbf{P}\|_F^2}.
\end{aligned} \tag{139}$$

Based on the Hoffman bound [?, Theorem 4.2], we have  $\|\mathbf{e}_i\|_2^2 \leq L_1 \|(\mathbf{t} - \mathbf{P} \mathbf{x}_i)^+\|_2^2$ , where  $L_1 = 1/\sigma_{\min}^2(\mathbf{P})$ . Combining this result with (??) leads to

$$\begin{aligned}
\mathbb{E} \{ \|\mathbf{e}_{i+1}\|_2^2 \} &\leq \left( 1 - \frac{\sigma_{\min}^2(\mathbf{P})}{\|\mathbf{P}\|_F^2} \right) \|\mathbf{e}_i\|_2^2 \\
&= \left( 1 - \frac{1}{\kappa^2(\mathbf{P})} \right) \|\mathbf{e}_i\|_2^2.
\end{aligned} \tag{140}$$

According to Theorem ??, we know  $\kappa(\mathbf{A}) = \kappa(\mathbf{P})$ . Thus, after  $i$  iterations we have

$$\mathbb{E} \{ \|\mathbf{e}_{i+1}\|_2^2 \} \leq \left( 1 - \frac{1}{\kappa^2(\mathbf{A})} \right)^i \|\mathbf{e}_0\|_2^2. \tag{141}$$

It is evident that the rest of the proof follows the same logic as the proof of Proposition ??. Note that the convergence rate (??) can be derived for any operator follows  $P : \mathcal{H} \rightarrow \Omega_c$ . This underscores the versatility of the algorithm and its ability to accommodate various types of operators.

### B. Convergence Analysis for SVP-ORKA

To prove the convergence of SVP-ORKA, consider an operator function  $\mathcal{G}_f$  applied to a matrix  $\mathbf{X}$  with rank  $r'$  as follows:

$$\mathcal{G}_f(\mathbf{X}) = \sum_{k=1}^{r'} f(\sigma_k) \mathbf{u}_k \mathbf{v}_k^H, \quad (142)$$

where  $\{\sigma_k, \mathbf{u}_k, \mathbf{v}_k\}_{k=1}^{r'}$  are singular values of  $\mathbf{X}$  and its corresponding singular vectors, and  $f$  is a  $L$ -Lipschitz continuous projector function. As comprehensively discussed in [?] and [?, Theorem 4.2], the following relation holds for two matrices  $\mathbf{X}_1$  and  $\mathbf{X}_2$  belonging to the Hilbert space  $\mathcal{H}$ :

$$\|\mathcal{G}_f(\mathbf{X}_1) - \mathcal{G}_f(\mathbf{X}_2)\|_F \leq L \|\mathbf{X}_1 - \mathbf{X}_2\|_F. \quad (143)$$

In SVP-ORKA,  $f$  is an operator which only chooses the  $r$ -largest singular values. It is straightforward to verify that such  $f$  satisfies (??) with  $L = 1$ . Therefore, one can conclude

$$\|P_r(\mathbf{X}_1) - P_r(\mathbf{X}_2)\|_F \leq \|\mathbf{X}_1 - \mathbf{X}_2\|_F, \quad \forall \mathbf{X}_1, \mathbf{X}_2 \in \mathcal{H}. \quad (144)$$

Since  $\|\mathbf{X}_1 - \mathbf{X}_2\|_F = \|\text{vec}(\mathbf{X}_1) - \text{vec}(\mathbf{X}_2)\|_2$ , the convergence proof of SVP-ORKA is identical to that of ST-ORKA from this point forward. It is worth noting that using the SVT operator instead of SVP can lead to the same convergence proof for finding the solution in  $\mathcal{P}_1^{(M)}$ . This is because in SVT operator, the function  $f(\cdot)$  is  $f(\mathbf{x}) = (\mathbf{x} - \boldsymbol{\tau})^+$ , where  $\boldsymbol{\tau}$  is a predefined threshold. As discussed for the ST operator in Appendix ??, this function also satisfies the Lipschitz continuity.

### C. Convergence Analysis for HT-ORKA

Define  $\mathbf{e}_i = \mathbf{x}_i - \hat{\mathbf{x}}$ ,  $t_j^{(\ell)} = r_j^{(\ell)} \tau_j^{(\ell)}$ , and  $\mathbf{p}_j^{(\ell)} = r_j^{(\ell)} \mathbf{a}_j$ . Since  $\mathcal{T}_s(\cdot)$  is an operator that selects the best  $s$ -sparse approximation of the solution at each iteration, we can determine the convergence rate of HT-ORKA as follows:

$$\begin{aligned} \|\mathbf{e}_{i+1}\|_2^2 &= \|\mathcal{T}_s(\mathbf{z}_{i+1}) - \hat{\mathbf{x}}\|_2^2 \\ &\leq \|\mathcal{T}_s(\mathbf{z}_{i+1}) - \hat{\mathbf{x}}\|_2^2 + \|\mathbf{z}_{i+1} - \hat{\mathbf{x}}\|_2^2 \\ &\leq 2 \|\mathbf{z}_{i+1} - \hat{\mathbf{x}}\|_2^2. \end{aligned} \quad (145)$$

As presented earlier, the term  $\|\mathbf{z}_{i+1} - \hat{\mathbf{x}}\|_2^2$  can be bounded as

$$\|\mathbf{z}_{i+1} - \hat{\mathbf{x}}\|_2^2 \leq \|\mathbf{e}_i\|_2^2 - \frac{\left( \left( t_j^{(\ell)} - \mathbf{p}_j^{(\ell)} \mathbf{x}_i \right)^+ \right)^2}{\|\mathbf{p}_j^{(\ell)}\|_2^2}, \quad (146)$$

and

$$\mathbb{E} \{ \|\mathbf{z}_{i+1} - \hat{\mathbf{x}}\|_2^2 \} \leq \|\mathbf{e}_i\|_2^2 - \frac{\|(\mathbf{t} - \mathbf{P}\mathbf{x}_i)^+\|_2^2}{\|\mathbf{P}\|_F^2}. \quad (147)$$

According to the Hoffman bound [?, Theorem 4.2] and (??), we have

$$\mathbb{E} \{ \|\mathbf{e}_{i+1}\|_2^2 \} \leq 2 \left( 1 - \frac{1}{\kappa^2(\mathbf{A})} \right)^i \|\mathbf{e}_0\|_2^2. \quad (148)$$

It is straightforward to verify that a similar roadmap to the proof of Proposition ?? may be followed.