

1-BIT COMPRESSED SENSING OF POSITIVE SEMI-DEFINITE MATRICES VIA RANK-1 MEASUREMENT MATRICES

Xiyuan Wang*, Kun Wang†, Zhongshan Zhang*, and Keping Long*

*School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing, China

†National Computer Network Emergency Response Technical Team/Coordination Center of China, Beijing, China

ABSTRACT

In this paper, we investigate the problem of recovering positive semi-definite (PSD) matrix from 1-bit sensing. The measurement matrix is rank-1 and constructed by the outer product of a pair of vectors, whose entries are independent and identically distributed (i.i.d.) Gaussian variables. The recovery problem is solved in closed form through a convex programming. Our analysis reveals that the solution is biased in general. However, in case of error-free measurement, we find that for rank- r PSD matrix with bounded condition number, the bias decreases with an order of $\mathcal{O}(1/r)$. Therefore, an approximate recovery is still possible. Numerical experiments are conducted to verify our analysis.

Index Terms— 1-bit compressed sensing, positive semi-definite matrix recovery, signal quantization, rank-1 measurement matrix

1. INTRODUCTION

In contrast with the conventional compressed sensing (CS) [1,2], which relies on real-value measurements to recover the signal, the 1-bit CS method [3] aims at reconstructing the original signal from sequence of bits which are obtained by quantizing the measured values. 1-bit CS is particularly attractive to applications where low-cost and low-precision analog-to-digital converter (ADC) of high speed is required. Therefore, the research on 1-bit CS has grown rapidly in recent years, and has developed from its burgeoning field of error-free quantization [3–6] to robust 1-bit CS [7–10]. In addition, the idea of 1-bit CS has also been extended to areas such as 1-bit matrix completion [11], wideband spectral estimation [12, 13], and 1-bit subspace learning [14], etc.

In this paper, we generalize 1-bit CS to the problem of positive semi-definite matrix (PSD) recovery, which can find applications in covariance matrix sketching [15], compressed spectral sensing [12, 13], etc. In our measurement model, the PSD matrix is linearly measured by rank-1 matrix, which is constructed by the outer product of two random vectors with independent and identically distributed (i.i.d.) Gaussian entries, and 1-bit quantization is applied to the resulting sample values. Our work is connected with covariance

matrix recovery from compressed sampling, which is based on real-value measurements and has drawn many studies such as [15–17]. However, research on PSD matrix recovery in the context of 1-bit CS is still very rare. Among the scarce works, the recent work in [14] is closely related to ours. In fact, we show that the proposed rank-1 measurement is equivalent to the measurement model in [14], where the difference of two quadratic measurements is quantized. However, the rank-1 measurement allows more tractable analysis, which will be shown in our paper. Moreover, the goal of [14] is to estimate the principle subspace, which is quite different from ours. Another work related to ours is the 1-bit wideband power spectral estimation method proposed in [12], which is called frugal sensing in [12]. In [12], by comparing the sample values from quadratic measurements against some threshold, the resulting signs are used to recover the power spectrum. This problem can be reduced to recovering a PSD Toeplitz matrix. However, like [14], the use of quadratic measurement makes the analysis of performance difficult. Although numerical experiment provides empirical evidence for the effectiveness of the method, a rigorous error analysis is still missing. Besides, in [12], the threshold value should be carefully designed to guarantee good performance, whereas no quantization threshold other than zero is needed in our work.

Built upon the proposed rank-1 measurement and inspired by the method in [10], we solve the estimated PSD matrix in closed form through a convex programming. As in [14], the estimated PSD matrix possesses the identical eigenvectors as the original PSD matrix. But due to the bias in the eigenvalues, which is also observed in [14], the estimator is biased. Utilizing the rank-1 structure of the measurement matrix, we are able to quantify the bias. Compared with the bounds on the estimated eigenvalues in [14], our results are more useful in analyzing the recovery error in our case. As the key finding of our paper, we reveal that in the case of error-free 1-bit quantization, for a $n \times n$ PSD matrix with rank r and bounded condition number, the recovery error under m 1-bit measurements is bounded by $\mathcal{O}(\sqrt{nr \log(n)/m}) + \mathcal{O}(1/r)$ with high probability. Whereas the first term in the bound diminishes with $\mathcal{O}(1/\sqrt{m})$, the second term is the bias term and decreases with $\mathcal{O}(1/r)$. As a result, in spite of the bias, an approximate recovery can be expected if r is properly increased. In the numerical experiments, we observe that small recovery error can be achieved even for PSD matrix with medium rank. Note that although obtained under the assumption of error-free measurements, our results can provide insight into the more general problem of 1-bit PSD matrix recovery.

Notations: \mathbf{I}_n is $n \times n$ identity matrix, and $\mathbb{R}^{m \times n}$ is the set of $m \times n$ real matrix. $\text{diag}\{x_1, \dots, x_n\}$ is the diagonal matrix with diagonal entries x_1, \dots, x_n . For vector, $\|\cdot\|$ is the ℓ_2 norm. For matrix, $\|\cdot\|$ is the spectral norm, and $\|\cdot\|_F$ is the Frobenius norm.

This work was supported by the key project of the National Natural Science Foundation of China (No. 61431001), Program for New Century Excellent Talents in University (NECT-12-0774), the open research fund of National Mobile Communications Research Laboratory Southeast University (No.2013D12), Fundamental Research Funds for the Central Universities (FRF-BD-15-012A), Key Laboratory of Cognitive Radio and Information Processing, Ministry of Education (Guilin University of Electronic Technology), and the Foundation of Beijing Engineering and Technology Center for Convergence Networks and Ubiquitous Services. The corresponding author is Dr. Zhongshan Zhang (email: zhangzs@ustb.edu.cn).

2. 1-BIT SENSING VIA RANK-1 MEASUREMENTS

In this section, we present the measurement model and the recovery problem for 1-bit CS sensing of PSD matrix. We denote the PSD matrix to be sensed as $\Sigma \in \mathbb{R}^{n \times n}$, which could be the sample covariance matrix as in [14]. Suppose that Σ has eigenvalue decomposition $\Sigma = \mathbf{Q}\Lambda\mathbf{Q}^T$, where $\mathbf{Q} \in \mathbb{R}^{n \times n}$ is orthogonal, and diagonal matrix $\Lambda = \text{diag}\{\lambda_1, \dots, \lambda_n\}$ with diagonal elements λ_i ($i = 1, \dots, n$) sorted in descending order. Denote the rank of Σ as r . Then, the eigenvalues of Σ satisfy $\lambda_1 \geq \dots \geq \lambda_r > \lambda_{r+1} = \dots = \lambda_n = 0$. The condition number of Σ is defined as $\kappa = \lambda_1/\lambda_r \leq \infty$. Since the information of matrix norm is lost in 1-bit sensing, we adopt the assumption of $\|\Sigma\|_F = 1$, which gives $\sum_{i=1}^r \lambda_i^2 = 1$.

Denote the k th ($k = 1, \dots, m$) rank-1 measurement matrix as $\mathbf{W}_k = \mathbf{a}_k \mathbf{b}_k^T$, where $\mathbf{a}_k, \mathbf{b}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_n)$ are independent. As in [8], $y_k \in \{+1, -1\}$ in bit sequence $\mathbf{y} = [y_1, \dots, y_m]^T$ is drawn randomly with distribution satisfying

$$\mathbb{E}[y_k | \mathbf{W}_k, \Sigma] = \theta(\langle \mathbf{W}_k, \Sigma \rangle) = \theta(\mathbf{a}_k^T \Sigma \mathbf{b}_k), \quad (1)$$

where $\theta(z) \in [-1, 1]$ is some nonlinear function, which may be unknown. Similar to [8], we also define the function

$$\lambda(\sigma^2) = \mathbb{E}_{g \sim \mathcal{N}(0, \sigma^2)} [\theta(g)g]. \quad (2)$$

In (2), since the Gaussian variable g has variance σ^2 , $\lambda(\sigma^2)$ is a function depending on σ^2 . This is different from [8], where g is a standard Gaussian variable and λ is constant. For noiseless 1-bit quantization, i.e., $y_k = \text{sign}(\mathbf{a}_k^T \Sigma \mathbf{b}_k)$, we have $\theta(z) = \text{sign}(z)$ and $\lambda(\sigma^2) = \sigma \sqrt{2/\pi}$.

In [14], the measurement matrix is constructed by $\tilde{\mathbf{W}}_k = (\tilde{\mathbf{a}}_k \tilde{\mathbf{a}}_k^T - \tilde{\mathbf{b}}_k \tilde{\mathbf{b}}_k^T)/2$,¹ where $\tilde{\mathbf{a}}_k, \tilde{\mathbf{b}}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_n)$ are independent. Suppose that the 1-bit measurement with $\tilde{\mathbf{W}}_k$ follows the same model in (1). Namely, $\mathbb{E}[y_k | \tilde{\mathbf{W}}_k, \Sigma] = \theta(\langle \tilde{\mathbf{W}}_k, \Sigma \rangle) = \theta((\tilde{\mathbf{a}}_k^T \Sigma \tilde{\mathbf{a}}_k - \tilde{\mathbf{b}}_k^T \Sigma \tilde{\mathbf{b}}_k)/2)$. If

$$\tilde{\mathbf{a}}_k = \frac{1}{\sqrt{2}} (\mathbf{a}_k + \mathbf{b}_k), \tilde{\mathbf{b}}_k = \frac{1}{\sqrt{2}} (\mathbf{a}_k - \mathbf{b}_k), \quad (3)$$

we have $\tilde{\mathbf{W}}_k = ((\mathbf{a}_k + \mathbf{b}_k)(\mathbf{a}_k + \mathbf{b}_k)^T - (\mathbf{a}_k - \mathbf{b}_k)(\mathbf{a}_k - \mathbf{b}_k)^T)/4 = (\mathbf{W}_k + \mathbf{W}_k^T)/2$ and $\langle \tilde{\mathbf{W}}_k, \Sigma \rangle = \langle \mathbf{W}_k, \Sigma \rangle$. Just notice that the transform in (3) is orthogonal and will preserve the independence of $\tilde{\mathbf{a}}_k$ and $\tilde{\mathbf{b}}_k$. Therefore, the measurement procedure in [14] is equivalent to the measurement with the rank-1 matrix in our paper as long as $\tilde{\mathbf{a}}_k$ and $\tilde{\mathbf{b}}_k$ are related to \mathbf{a}_k and \mathbf{b}_k by the transform in (3). We will show in Section 3 that compared with its equivalence in [14], the rank-1 measurement matrix is more tractable when it comes to analyzing the recovery error.

Denote $\mathbf{S}_m = \frac{1}{2m} \sum_{k=1}^m y_k (\mathbf{W}_k + \mathbf{W}_k^T)$ which has eigenvalue decomposition as $\mathbf{S}_m = \mathbf{U}\mathbf{T}\mathbf{U}^T$ with orthogonal matrix $\mathbf{U} \in \mathbb{R}^{n \times n}$ and diagonal matrix \mathbf{T} . Similar to the convex problem in (6) of [10], which is for recovering sparse signal from 1-bit measurements, we use the following convex programming to recovery Σ from its 1-bit measurements:

$$\max_{\mathbf{X}} -\text{tr}(\mathbf{X}\mathbf{S}_m) + \gamma \text{tr}(\mathbf{X}), \text{ s.t. } \mathbf{X} \succeq 0, \|\mathbf{X}\|_F \leq 1, \quad (4)$$

whose solution is given in closed form as

$$\hat{\mathbf{X}} = \begin{cases} \frac{1}{\|\mathcal{D}_\gamma(\mathbf{S}_m)\|_F} \mathcal{D}_\gamma(\mathbf{S}_m), & \text{if } \|\mathbf{S}_m\| > \gamma, \\ 0, & \text{otherwise,} \end{cases} \quad (5)$$

¹Here we add a scaling factor of 1/2 in $\tilde{\mathbf{W}}_k$, which is slightly different from the original definition in [14].

where $\gamma > 0$ is the regularization parameter, and $\mathcal{D}_\gamma(\mathbf{S}_m) = \mathbf{U}(\mathbf{T} - \gamma \mathbf{I}_n)^+ \mathbf{U}^T$ is the singular value thresholding operator defined in [18].

3. RECOVERY ERROR ANALYSIS

In this section, we derive the recover error of the estimator $\hat{\mathbf{X}}$. We start our analysis with the following lemma regarding $\mathbb{E}[y_k \mathbf{W}_k]$.

Lemma 1. *Let the mean of $y_k \mathbf{W}_k$ be $\mathbf{S} = \mathbb{E}[y_k \mathbf{W}_k]$. Then, \mathbf{S} has eigenvalue decomposition as $\mathbf{S} = \mathbf{Q}\mathbf{D}\mathbf{Q}^T$, where the diagonal matrix $\mathbf{D} = \text{diag}\{\sigma_1, \dots, \sigma_n\}$ with $\sigma_i = \lambda_i \delta_i$ for $i = 1, \dots, r$ and $\sigma_i = 0$ for $i = r+1, \dots, n$. The scalar δ_i ($i = 1, \dots, r$) is given by*

$$\delta_i = \mathbb{E} \left[\frac{\lambda(\mathbf{z}^T \Lambda^2 \mathbf{z})}{\mathbf{z}^T \Lambda^2 \mathbf{z}} z_i^2 \right], \quad (6)$$

where the random vector $\mathbf{z} = [z_1, \dots, z_n]^T \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_n)$. In particular, for error-free 1-bit quantization,

$$\delta_i = \sqrt{\frac{2}{\pi}} \mathbb{E} \left[\frac{z_i^2}{\sqrt{\mathbf{z}^T \Lambda^2 \mathbf{z}}} \right]. \quad (7)$$

Proof. Notice that

$$\begin{aligned} \mathbf{S} &= \mathbb{E}[y_k \mathbf{W}_k] = \mathbb{E}[\mathbf{W}_k \mathbb{E}[y_k | \mathbf{W}_k, \Sigma]] = \mathbb{E}[\theta(\mathbf{a}_k^T \Sigma \mathbf{b}_k) \mathbf{a}_k \mathbf{b}_k^T] \\ &= \mathbb{E}[\theta(\sigma \mathbf{a}_k^T \mathbf{u}_1) (\mathbf{a}_k^T \mathbf{u}_1 \mathbf{u}_1 + \mathbf{a}_k^T \mathbf{u}_2 \mathbf{u}_2) \mathbf{b}_k^T], \end{aligned}$$

where $\sigma = \|\Sigma \mathbf{b}_k\|$, and unit vectors $\mathbf{u}_1 = \Sigma \mathbf{b}_k / \sigma$ and \mathbf{u}_2 satisfy $\mathbf{u}_1^T \mathbf{u}_2 = 0$. Conditioned upon \mathbf{b}_k , random variables $\mathbf{a}_k^T \mathbf{u}_1, \mathbf{a}_k^T \mathbf{u}_2 \sim \mathcal{N}(0, 1)$ are independent. So we have

$$\mathbb{E}[\theta(\sigma \mathbf{a}_k^T \mathbf{u}_1) (\mathbf{a}_k^T \mathbf{u}_1 \mathbf{u}_1 + \mathbf{a}_k^T \mathbf{u}_2 \mathbf{u}_2) | \mathbf{b}_k] = \frac{\lambda(\sigma^2)}{\sigma^2} \Sigma \mathbf{b}_k \mathbf{b}_k^T,$$

which yields

$$\mathbf{S} = \mathbb{E} \left[\frac{\lambda(\sigma^2)}{\sigma^2} \Sigma \mathbf{b}_k \mathbf{b}_k^T \right] = \mathbf{Q}\Lambda\mathbf{D}\mathbf{Q}^T, \quad (8)$$

where matrix $\mathbf{D} = \mathbb{E}[\lambda(\mathbf{z}^T \Lambda^2 \mathbf{z}) \mathbf{z} \mathbf{z}^T / \mathbf{z}^T \Lambda^2 \mathbf{z}]$, and $\mathbf{z} = \mathbf{Q}^T \mathbf{b}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_n)$. As $\mathbb{E}[\lambda(\mathbf{z}^T \Lambda^2 \mathbf{z}) z_i z_j / \mathbf{z}^T \Lambda^2 \mathbf{z}] = -\mathbb{E}[\lambda(\mathbf{z}^T \Lambda^2 \mathbf{z}) z_i z_j / \mathbf{z}^T \Lambda^2 \mathbf{z}]$ for $i, j = 1, \dots, n$ and $i \neq j$, which follows from the symmetry of the distribution $\mathcal{N}(\mathbf{0}, \mathbf{I}_n)$, we have that the off-diagonal elements of \mathbf{D} are zero, and therefore \mathbf{D} is diagonal. So, \mathbf{S} has the expression in (8) as its eigenvalue decomposition and the result in the lemma is readily obtained. \square

As is discussed in Section 2, the equivalence between measurement matrices \mathbf{W}_k and $\tilde{\mathbf{W}}_k$ gives identical empirical mean \mathbf{S}_m . Thus, the bound on the deviation of \mathbf{S}_m from $\mathbf{S} = \mathbb{E}[\mathbf{S}_m]$ in Lemma 2 of [14] can be readily applied here, and we simply reproduce the result below for reference.

Lemma 2. *Let $\delta \in [0, 1]$. With probability $1 - \delta$, we have that*

$$\|\mathbf{S}_m - \mathbf{S}\| \leq \sqrt{\frac{C_1 n}{m} \log \frac{2n}{\delta}}, \quad (9)$$

where $C_1 > 0$ is some constant.

Before proceeding with our analysis of the recovery error, we give the following proposition, which bounds the eigenvalues of \mathbf{S} for the noiseless quantization and is thus the counterpart of Lemma 1 in [14]. Derived from the exact expression of eigenvalue, the bounds here are much tighter.

Proposition 1. *In the case of the error-free 1-bit quantization, the non-zero eigenvalues of \mathbf{S} are bounded by*

$$\frac{2}{\pi} \geq \sigma_i \geq \frac{2}{\sqrt{\pi\kappa r}} \frac{\Gamma(\frac{r+1}{2})}{\Gamma(\frac{r}{2})}, i = 1, \dots, r. \quad (10)$$

Proof. Evidently, from (7),

$$\sigma_i \leq \sqrt{\frac{2}{\pi}} \mathbb{E} \left[\frac{\lambda_i z_i^2}{\lambda_i |z_i|} \right] = \sqrt{\frac{2}{\pi}} \mathbb{E}[|z_i|] = \frac{2}{\pi}, \quad (11)$$

which gives the upper bound in (10). As to the lower bound,

$$\sigma_i \geq \frac{2}{\pi} \mathbb{E} \left[\frac{\lambda_i z_i^2}{\lambda_i \|\mathbf{z}_r\|} \right] \geq \sqrt{\frac{2}{\pi}} \mathbb{E} \left[\frac{z_i^2}{\kappa \|\mathbf{z}_r\|} \right], \quad (12)$$

where $\mathbf{z}_r = [z_1, \dots, z_r]^T \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_r)$. Notice that $\mathbb{E}[z_1^2/\|\mathbf{z}_r\|] = \dots = \mathbb{E}[z_r^2/\|\mathbf{z}_r\|]$, which is from the symmetry of the distribution of \mathbf{z}_r . Then, we have that $r\mathbb{E}[z_i^2/\|\mathbf{z}_r\|] = \sum_{i=1}^r \mathbb{E}[z_i^2/\|\mathbf{z}_r\|] = \mathbb{E}[\|\mathbf{z}_r\|] = \sqrt{2}\Gamma((r+1)/2)/\Gamma(r/2)$. The expectation of $\|\mathbf{z}_r\|$, which is distributed according to the χ distribution with r degrees of freedom, follows from the moment result in [19, p.452]. Plugging the result into (12) yields the lower bound in (10). \square

Remark 1. *It is easy to check that in (10) the upper bound is tight in the case of $r = 1$, and the lower bound is tight when all non-zero eigenvalues of $\mathbf{\Sigma}$ are equal and $\kappa = 1$. As a byproduct of Proposition 1, from the limit of the ratio of two gamma functions [20, p.257, 6.1.46], σ_i ($i = 1, \dots, r$) can be lower bounded asymptotically by*

$$\lim_{r \rightarrow \infty} \frac{2}{\sqrt{\pi\kappa r}} \frac{\Gamma(\frac{r+1}{2})}{\Gamma(\frac{r}{2})} = \frac{\sqrt{2}}{\kappa\sqrt{\pi}} \frac{1}{\sqrt{r}}. \quad (13)$$

To perform the error analysis, we first multiply the \mathbf{S} by its Frobenius norm to get $\bar{\mathbf{S}} = \mathbf{S}/\|\mathbf{S}\|_F$. According to Lemma 1, matrices $\bar{\mathbf{S}}$ and $\mathbf{\Sigma}$ share the identical eigenspace. Moreover, it can be observed from (7) that when $r = 1$ or $r > 1$ and all non-zero eigenvalues of $\mathbf{\Sigma}$ are equal, $\bar{\mathbf{S}}$ and $\mathbf{\Sigma}$ have the same eigenvalues as well. Except for the aforementioned cases, the eigenvalues of $\bar{\mathbf{S}}$ and $\mathbf{\Sigma}$ are not the same. Consequently, an error term in the difference of $\bar{\mathbf{S}}$ and $\mathbf{\Sigma}$ arises from the bias in the eigenvalues of $\bar{\mathbf{S}}$. It is complicated to quantify such error and give an analysis of recovery error for general nonlinear function $\theta(z)$. As such, in this paper, we settle with the simple case of error-free 1-bit quantization, i.e., $\theta(z) = \text{sign}(z)$. As to the error between $\bar{\mathbf{S}}$ and $\mathbf{\Sigma}$, we have the following lemma.

Lemma 3. *In the case of error-free 1-bit quantization, denote $\rho = 1/\|\mathbf{S}\|_F$. If $r > 5\kappa^5$, the error between $\bar{\mathbf{S}}$ and $\mathbf{\Sigma}$ is bounded by*

$$\|\bar{\mathbf{S}} - \mathbf{\Sigma}\|_F = \sqrt{\sum_{i=1}^r \lambda_i^2 (\rho\delta_i - 1)^2} \leq \frac{C_2 \kappa^{10}}{r}, \quad (14)$$

with $C_2 > 0$ as some constant.

Resorting to the lemma below, which bounds the deviation of δ_i from 1, Lemma 3 can be proved by straightforward derivation. We omit the proof due to space limitation.

Lemma 4. *Assuming error-free 1-bit quantization, we have the following bound on δ_i ($i = 1, \dots, r$)*

$$|\delta_i - 1| \leq \frac{5\kappa^5}{r}. \quad (15)$$

Proof. Denote $u = \mathbf{z}^T \mathbf{\Lambda}^2 \mathbf{z} - 1$. Since $\sum_{i=1}^r \lambda_i^2 = 1$ and z_i s are i.i.d. standard Gaussian variables, by simple algebra, we can derive that $\mathbb{E}[uz_i^2] = \sum_{j=1}^r \lambda_j^2 \mathbb{E}[z_i^2(z_j^2 - 1)] = \lambda_i^2 (\mathbb{E}[z_i^4] - \mathbb{E}[z_i^2]) = 2\lambda_i^2$ and $\mathbb{E}[u^2] = \sum_{i,j=1}^r \lambda_i^2 \lambda_j^2 \mathbb{E}[(z_i^2 - 1)(z_j^2 - 1)] = \sum_{i=1}^r \lambda_i^4 \mathbb{E}[(z_i^2 - 1)^2] = 2 \sum_{i=1}^r \lambda_i^4$. So, for $i = 1, \dots, r$, we have

$$\begin{aligned} |\delta_i - 1| &= \left| \mathbb{E} \left[\frac{z_i^2}{\sqrt{1+u}} - 1 + \frac{1}{2}uz_i^2 - \frac{1}{2}uz_i^2 \right] \right| \\ &= \left| \mathbb{E} \left[z_i^2 \left(\frac{1}{\sqrt{1+u}} - 1 + \frac{1}{2}u \right) \right] - \frac{1}{2} \mathbb{E}[uz_i^2] \right| \\ &\leq \mathbb{E}[z_i^2] \mathbb{E} \left[\left| \frac{1}{\sqrt{1+u}} - 1 + \frac{1}{2}u \right| \right] + \lambda_i^2. \end{aligned}$$

In addition, it can be derived that

$$\begin{aligned} \frac{1}{\sqrt{1+u}} - 1 + \frac{1}{2}u &= \frac{u^2}{2\sqrt{1+u}} \left(\frac{1}{1+\sqrt{1+u}} + \frac{1}{(1+\sqrt{u+1})^2} \right) \\ &\leq \frac{u^2}{\sqrt{1+u}}, \end{aligned}$$

and $r\lambda_1^2/\kappa^2 \leq \lambda_1^2(\sum_{i=2}^r (\lambda_i^2/\lambda_1^2) + 1) = 1 = \sum_{i=1}^r \lambda_i^2 = \lambda_r^2(\sum_{i=1}^{r-1} (\lambda_i^2/\lambda_r^2) + 1) \leq r\kappa^2\lambda_r^2$, which yields bounds on λ_i^2 as $1/(\kappa^2 r) \leq \lambda_i^2 \leq \kappa^2/r$. Then, aided by the results above, we can further bound $|\delta_i - 1|$ as

$$\begin{aligned} |\delta_i - 1| &\leq \mathbb{E} \left[\frac{u^2}{\sqrt{1+u}} \right] + \lambda_i^2 \leq \mathbb{E}[u^2] \mathbb{E} \left[\frac{1}{\sqrt{1+u}} \right] + \frac{\kappa^2}{r} \\ &\stackrel{(a)}{\leq} \left(2 \sum_{i=1}^r \lambda_i^4 \right) \mathbb{E} \left[\frac{\kappa\sqrt{r}}{\|\mathbf{z}_r\|} \right] + \frac{\kappa^2}{r} \stackrel{(b)}{\leq} \frac{2\kappa^5}{r} \sqrt{\frac{r}{2}} \frac{\Gamma(\frac{r-1}{2})}{\Gamma(\frac{r}{2})} + \frac{\kappa^2}{r} \\ &\stackrel{(c)}{\leq} \frac{2\kappa^5}{r} \frac{\sqrt{\frac{r}{2}}}{\sqrt{\frac{r}{2} - \frac{3}{4}}} + \frac{\kappa^2}{r} \stackrel{(d)}{\leq} \frac{4\kappa^5}{r} + \frac{\kappa^2}{r} \leq \frac{5\kappa^5}{r}, \end{aligned}$$

where (a) is from the results of $\mathbb{E}[u^2]$ and $\mathbf{z}^T \mathbf{\Lambda}^2 \mathbf{z} \geq \kappa^2 \|\mathbf{z}\|^2/r$, (b) is derived from the upper bound of λ_i^2 and directly calculating $\mathbb{E}[1/\|\mathbf{z}_r\|]$ based on the probability density function (pdf) of χ^2 distribution [19], (c) is from applying the upper bound of gamma function ratio in [21, (2.13)], and (d) follows from the monotonic decreasing of function $\sqrt{x/(x-3/4)}$. \square

With Lemmas 1-4, we are now in a position to present our main result on the recovery error of the solution $\hat{\mathbf{X}}$ in (5) as Theorem 1.

Theorem 1. *In the case of error-free 1-bit quantization, let*

$$\gamma = 2\sqrt{\frac{C_1 n \log(2n/\delta)}{m}} \quad (16)$$

in (4). Then, with probability at least $1 - \delta$, we have for $r > 5\kappa^5$ the following bound on recovery error:

$$\|\hat{\mathbf{X}} - \mathbf{\Sigma}\|_F \leq 3\sqrt{\frac{C_1 nr \log(2n/\delta)}{m}} + \frac{C_2 \kappa^{10}}{r}. \quad (17)$$

Proof. Using triangular inequality, we have $\|\hat{\mathbf{X}} - \mathbf{\Sigma}\|_F \leq \|\hat{\mathbf{X}} - \bar{\mathbf{S}}\|_F + \|\bar{\mathbf{S}} - \mathbf{\Sigma}\|_F$. The bound of the second error term, which is the bias in $\hat{\mathbf{X}}$, follows directly from Lemma 3. The first term in the error can be bounded based on the idea of the proof for Theorem 1 in [10], with the aid of Lemma 2. To accommodate the proof in [10] to our case, the only modification needed is to perform the decomposition of tangent and normal spaces at $\mathbf{\Sigma}$ as is in [22]. Therefore, we omit the derivation here since it is largely a repetition of the proof in [10]. \square

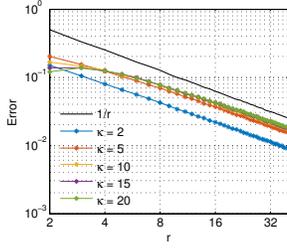


Fig. 1: Average error between $\bar{\mathbf{S}}$ and Σ versus r .

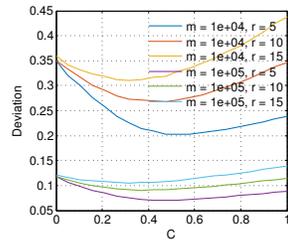
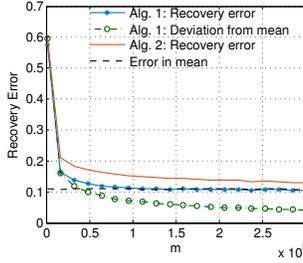
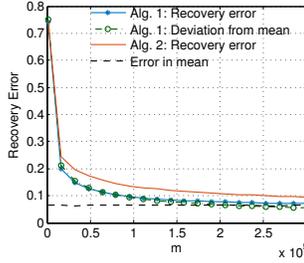


Fig. 2: Average deviation of $\hat{\mathbf{X}}$ from $\bar{\mathbf{S}}$ versus C .



(a) $r = 5$



(b) $r = 10$

Fig. 3: Recovery error versus m .

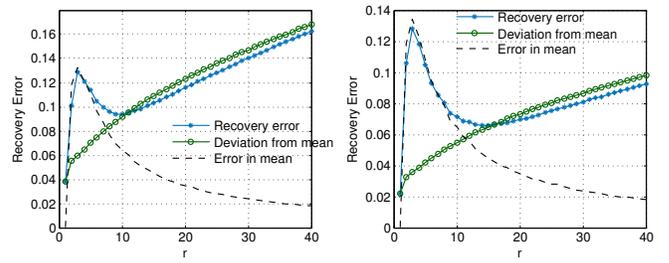
Remark 2. Theorem 1 states that the recovery error in the error-free case is of the order $\mathcal{O}(nr \log(n)/m) + \mathcal{O}(1/r)$ for the PSD matrix with bounded condition number. In the expression of the recovery error, it can be seen that with increased r , the second term $\mathcal{O}(1/r)$, which is the estimation bias, can be reduced, whereas the first error term will grow as $\mathcal{O}(\sqrt{r})$. Thus, we can anticipate that at some r the recovery error reaches a minimum if m is fixed and large enough. At first glance, the existence of the bias may seem to be a pessimistic result. However, as will be shown in our simulations, an approximate recovery with acceptable accuracy can be obtained for moderate r .

4. SIMULATIONS

In this section, we conduct simulation experiments to verify our analysis. All simulation results are from averaging over 1000 trials, and the measurement in the experiment is noiseless. For Σ , n is set as 50, the eigenvectors are generated by orthogonalizing the column vectors of $\mathbf{X} \in \mathbb{R}^{50 \times r}$, whose entries are i.i.d. standard Gaussian variables, and except for the extreme eigenvalues, which are set as κ and 1, other eigenvalues are drawn uniformly on the interval $[1, \kappa]$. The eigenvalues of Σ are then normalized so that $\|\Sigma\|_F = 1$. For the eigenvalues of $\bar{\mathbf{S}}$, Monte Carlo method is used to calculate the integrals in (7).

As a justification of Lemma 3, the curves of average error $\mathbb{E}[\|\bar{\mathbf{S}} - \Sigma\|_F]$ versus r with various κ s are plotted in Fig. 1. In the figure, r varies from 2 to 40, and the curve of $1/r$ is also plotted as comparison. As is predicted by Lemma 3, for large r the error decays with the order of $\mathcal{O}(1/r)$ and increases with κ . But as shown in the figure, the increase with κ tends to be bounded, which indicates that the bound in Lemma 3 can be further improved.

To determine the regularization γ in (4), we adopt a similar approach as in [10] to set $\gamma = C\sqrt{n \log(n)/m}$. The curves of error $\|\hat{\mathbf{X}} - \bar{\mathbf{S}}\|_F$ versus $C \in [0, 1]$ under several combinations of m and r are presented in Fig. 2. From Fig. 2, the best performance is obtained



(a) $m = 1 \times 10^5$

(b) $m = 3 \times 10^5$

Fig. 4: Recovery error versus r .

around $C = 0.5$, which is used in the remaining experiments.

Figures 3a and 3b show the performance of recovery error $\|\hat{\mathbf{X}} - \Sigma\|_F$ versus $m \in [1000, 300000]$. In Figures 3a and 3b, $r = 5$ and 10, respectively, and $\kappa = 20$. We denote the estimator in Section 2 as Algorithm 1 (Alg. 1). As a comparison, we also solve the problem

$$\min_{\mathbf{X}} -\text{tr}(\mathbf{X}\mathbf{S}_m), \text{ s.t. } \mathbf{X} \succeq 0, \|\mathbf{X}\|_F = 1, \|\mathbf{X}\|_* \leq \sqrt{r}, \quad (18)$$

which is an extension of the problem in (III.2) of [8] and is referred to as Algorithm 2 (Alg. 2). Apart from recovery errors of Algorithms 1 and 2, the deviation of $\hat{\mathbf{X}}$ from $\bar{\mathbf{S}}$ (Deviation from mean), i.e., $\|\hat{\mathbf{X}} - \bar{\mathbf{S}}\|_F$, and the error $\|\bar{\mathbf{S}} - \Sigma\|_F$ (Error in mean) are depicted as well. In the figures, the recovery errors of both algorithms are quite close and get smaller with increasing m . Also, according to Lemma 3, the error between Σ and $\bar{\mathbf{S}}$ is irreducible by increasing m , so the recovery error is bounded by the bias in the estimator, which dominates the performance for large m . As is expected by Theorem 1, the bias in the estimator in Fig. 3b for $r = 10$ is smaller than that for $r = 5$ in Fig. 3a. However, from both figures, it can be observed that the bias is not large even for moderate r . As a result, satisfactory recovery can still be achieved provided that the sample size is adequate.

Figs. 4a and 4b show the curves of recovery error versus $r \in [1, 40]$ for Algorithm 1. In Figs. 4a and 4b, $m = 1 \times 10^5$ and 3×10^5 , respectively, and $\kappa = 20$. In the figures, for $r = 1$, there is no error between $\bar{\mathbf{S}}$ and Σ , which can be deduced from Lemma 1, so the recovery error is small. For $r \geq 3$, as the deviation $\|\hat{\mathbf{X}} - \bar{\mathbf{S}}\|_F$ grows with r and the error $\|\bar{\mathbf{S}} - \Sigma\|_F$ decays with r , the recovery error reaches a minimum at some r , which confirms our discussion in Remark 2 for Theorem 1. Interestingly, there is a peak of recovery error at $r = 3$. One explanation for this is that the large κ makes the $r = 2$ case approximate the $r = 1$ case and leads to a smaller recovery error than the $r = 3$ case. As in Figs. 3a and 3b, it can also be observed that increasing m is helpful in reducing the deviation of $\hat{\mathbf{X}}$ from $\bar{\mathbf{S}}$ but not the error between $\bar{\mathbf{S}}$ and Σ .

5. CONCLUSIONS

We have studied the problem of recovering PSD matrix from 1-bit sensing with rank-1 measurement matrices. We solved the problem in closed form, and for the case of error-free measurement, we derived the recovery error bound and found in the error a bias term which decreases with the rank of PSD matrix. The analysis results were verified by simulation experiments. We believe that in the field of 1-bit CS of PSD matrix, there are still many unsolved problems. Important topics for future research includes PSD matrix recovery from noisy 1-bit measurement, 1-bit sensing of PSD Toeplitz matrix and other structured PSD matrix, etc.

6. REFERENCES

- [1] D. L. Donoho, “Compressed sensing,” *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
- [2] E. J. Candès and T. Tao, “Near-optimal signal recovery from random projections: Universal encoding strategies?,” *IEEE Trans. Inform. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.
- [3] P. T. Boufounos and R. Baraniuk, “1-bit compressive sensing,” in *Proc. 42nd Annu. Conf. Inf. Sci. Syst. (CISS’2008)*, Princeton, NJ, May 2008, pp. 16–21.
- [4] P. T. Boufounos, “Greedy sparse signal reconstruction from sign measurements,” in *Proc. Asilomar Conf. Signals Syst. Comput.*, Asilomar, CA, Nov. 2009, pp. 1305–1309.
- [5] J. N. Laska, Z. Wen, W. Yin, and R. G. Baraniuk, “Trust, but verify: Fast and accurate signal recovery from 1-bit compressive measurements,” *IEEE Trans. Signal Processing*, vol. 59, no. 11, pp. 5289–5301, Nov. 2011.
- [6] Y. Plan and R. Vershynin, “One-bit compressed sensing by linear programming,” *Comm. Pure Appl. Math.*, vol. 66, no. 8, pp. 1275–1297, Aug. 2013.
- [7] M. Yan, Y. Yang, and S. Osher, “Robust 1-bit compressive sensing using adaptive outlier pursuit,” *IEEE Trans. Signal Processing*, vol. 60, no. 7, pp. 3868–3875, July 2012.
- [8] Y. Plan and R. Vershynin, “Robust 1-bit compressed sensing and sparse logistic regression: A convex programming approach,” *IEEE Trans. Inform. Theory*, vol. 59, no. 1, pp. 482–494, Jan. 2013.
- [9] L. Jacques, J. N. Laska, P. T. Boufounos, and R. G. Baraniuk, “Robust 1-bit compressive sensing via binary stable embeddings of sparse vectors,” *IEEE Trans. Inform. Theory*, vol. 59, no. 4, pp. 2082–2102, Apr. 2013.
- [10] L. Zhang, J. Yi, and R. Jin, “Efficient algorithms for robust one-bit compressive sensing,” in *Proc. 31st Int. Conf. on Machine Learning (ICML’14)*, Beijing, China, June, 21–26 2014, pp. 820–828.
- [11] M. A. Davenport, Y. Plan, E. van den Berg, and M. Wootters, “1-bit matrix completion,” *Information and Inference*, 2014.
- [12] O. Mehanna and N. D. Sidiropoulos, “Frugal sensing: Wideband power spectrum sensing from few bits,” *IEEE Trans. Signal Processing*, vol. 61, no. 10, pp. 2693–2703, May 2013.
- [13] A. Konar, N. D. Sidiropoulos, and O. Mehanna, “Parametric frugal sensing of power spectra for moving average models,” *IEEE Trans. Signal Processing*, vol. 63, no. 5, pp. 1073–1085, Mar. 2015.
- [14] Y. Chi, “Subspace learning from bits,” *arXiv preprint arXiv:1407.6288*, July 2014.
- [15] Y. Chen, Y. Chi, and A. J. Goldsmith, “Exact and stable covariance estimation from quadratic sampling via convex programming,” *IEEE Trans. Inform. Theory*, vol. 61, no. 7, pp. 4034–4059, July 2015.
- [16] D. Romero and G. Leus, “Compressive covariance sampling,” in *Proc. IEEE Inform. Theory and Applicat. Workshop (ITA)*, San Diego, CA, Feb. 10–15 2013, pp. 1–8.
- [17] H. Qiao and P. Pal, “Generalized nested sampling for compression and exact recovery of symmetric Toeplitz matrices,” in *Proc. IEEE Global Conf. Signal and Inform. Process. (GlobalSIP)*, Atlanta, GA, Dec. 3–5 2014, pp. 443–447.
- [18] J.-F. Cai, E. J. Candès, and Z. Shen, “A singular value thresholding algorithm for matrix completion,” *SIAM J. Optim.*, vol. 20, no. 4, pp. 1956–1982, Apr. 2010.
- [19] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Continuous Univariate Distributions*, vol. 1, John Wiley & Sons, Inc., New York, 2nd edition, 1994.
- [20] M. Abramowitz and I. A. Stegun, Eds., *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Applied Mathematics Series 55. National Bureau of Standards, Washington, DC, 1970.
- [21] F. Qi, “Bounds for the ratio of two Gamma functions,” *Journal of Inequalities and Applications*, vol. 2010, no. 1, pp. 493058, 2010.
- [22] E. Candès and B. Recht, “Exact matrix completion via convex optimization,” *Foundations of Computational Mathematics*, vol. 9, no. 6, pp. 717–772, 2009.