ONE-BIT COMPRESSED SENSING WITH PARTIAL GAUSSIAN CIRCULANT MATRICES

SJOERD DIRKSEN*, HANS CHRISTIAN JUNG*, AND HOLGER RAUHUT*

ABSTRACT. In this paper we consider memoryless one-bit compressed sensing with randomly subsampled Gaussian circulant matrices. We show that in a small sparsity regime and for small enough accuracy δ, $m \sim \delta^{-4} s \log(N/s\delta)$ measurements suffice to reconstruct the direction of any s-sparse vector up to accuracy δ via an efficient program. We derive this result by proving that partial Gaussian circulant matrices satisfy an ℓ_1/ℓ_2 RIP-property. Under a slightly worse dependence on δ, we establish stability with respect to approximate sparsity, as well as full vector recovery results.

1. INTRODUCTION

In the past decade, compressed sensing has established itself as a new paradigm in signal processing. It predicts that one can reconstruct signals from a small number of linear measurements using efficient algorithms, by exploiting that many real-world signals possess a sparse representation. In the traditional compressed sensing literature, it is typically assumed that one can reconstruct a signal based on its analog linear measurements. In a realistic sensing scenario, measurements need to be quantized to a finite number of bits before they can be transmitted, stored, and processed. Formally, this means that one needs to reconstruct a sparse signal x based on *non-linear* measurements of the form $y = Q(Ax)$, where $Q : \mathbb{R}^m \to \mathcal{A}^m$ is a quantizer and $\mathcal{A}$ denotes a finite quantization alphabet.

In this paper, we study the measurement model

$$(1) \qquad y = \text{sign}(Ax + \tau),$$

where $A \in \mathbb{R}^{m \times N}$, $m \ll N$, sign is the signum function applied element-wise and $\tau \in \mathbb{R}^m$ is a (possibly random) vector consisting of thresholds. Thus, every linear measurement is quantized to a single bit in a memoryless fashion, i.e., each measurement is quantized independently. This quantizer is attractive from a practical point of view, as it can be implemented using an energy-efficient comparator to a fixed voltage level (if $\tau_i = c$ for all i) combined with dithering (if τ is random). In the case $\tau = 0$, this model was coined *one-bit compressed sensing* by Boufounos and Baraniuk [4]. Taking all thresholds equal to zero has the disadvantage that the energy $\|x\|_2^2$ of the original signal is lost during quantization and one can only hope to recover the direction of the signal. More recent works [15, 2] have shown that by using random thresholds, it is under appropriate circumstances possible to completely reconstruct the signal (up to any prescribed precision).

Until now, recovery results for the one-bit compressed sensing model (1) dealt almost exclusively with a Gaussian measurement matrix A. The only exception seems to be [1], which deals with subgaussian matrices. The goal of this paper is to derive reconstruction guarantees in the case that A is a randomly subsampled

Gaussian circulant matrix. This compressed sensing model is important for several real-world applications, including SAR radar imaging, Fourier optical imaging and channel estimation (see e.g. [24] and the references therein). Our work seems to be the first to give rigorous reconstruction guarantees for memoryless one-bit compressed sensing involving a structured random matrix.

Our results concern guarantees for uniform recovery under a *small sparsity assumption*. Concretely, for a desired accuracy parameter $0 < \varepsilon \leq 1$, we assume that the sparsity s is small enough, i.e.,

$$s \lesssim \sqrt{\varepsilon N / \log(N)}.$$

we suppose that the (expected) number of measurements satisfy

$$(2) \qquad m \gtrsim \begin{cases} \varepsilon^{-1} s \log(eN/(s\varepsilon)) & \text{if } 0 < \varepsilon \leq \left(\log^2(s) \log(N)\right)^{-1} \\ \varepsilon^{-1/2} s \log(s) \log^{3/2}(N) & \text{if } \left(\log^2(s) \log(N)\right)^{-1} < \varepsilon \leq 1. \end{cases}$$

Let us first phrase our results for $\tau = 0$. We consider two different recovery methods to reconstruct x, namely via a single hard thresholding step

$$(\text{HT}) \qquad\qquad x_{\text{HT}}^{\#} = H_s(A^* \text{sign}(Ax))$$

and via the program

$$(\text{LP}) \qquad \min_{z \in \mathbb{R}^n} \|z\|_1 \qquad \text{s.t.} \qquad \text{sign}(Az) = \text{sign}(Ax) \qquad \text{and} \qquad \|Az\|_1 = 1.$$

As the first constraint is equivalent to $(Az)_i \, \text{sign}((Ax)_i) \geq 0$ for $i = 1, \ldots, m$ and due to the second constraint it can be written as $\sum_{i=1}^{n} \text{sign}((Ax)_i)(Az)_i = 1$, it follows that (LP) is a linear program.

Our first result shows that under (2) the following holds with high probability: for any s-sparse x with $\|x\|_2 = 1$ the hard thresholding reconstruction $x_{\text{HT}}^{\#}$ satisfies $\|x - x_{\text{HT}}^{\#}\|_2 \leq \varepsilon^{1/4}$. Moreover, under slightly stronger conditions (see Theorem 4.2 with $\delta = \varepsilon^{1/4}$) for any vector satisfying $\|x\|_1 \leq \sqrt{s}$ and $\|x\|_2 = 1$, any solution $x_{\text{LP}}^{\#}$ satisfies $\|x - x_{\text{LP}}^{\#}\|_2 \leq \varepsilon^{1/8}$. As a consequence, we can reconstruct the direction $x/\|x\|_2$ of any s-sparse (resp. effectively sparse) signal via an efficient program.

Our second result gives guarantees for the full recovery of effectively sparse vectors, provided that an upper bound R on their energy is known. We suppose that τ is a vector of independent, $\mathcal{N}(0, \pi R^2 / 2)$-distributed random variables. If a condition similar to (2) is satisfied (see Theorem 4.2), then the following holds with high probability: for any $x \in \mathbb{R}^N$ with $\|x\|_1 \leq \sqrt{s}\|x\|_2$ and $\|x\|_2 \leq R$, the solution $x_{\text{CP}}^{\#}$ to the second-order cone program

$$(\text{CP}) \qquad \min_{z \in \mathbb{R}^N} \|z\|_1 \qquad \text{s.t.} \qquad \text{sign}(Az + \tau) = \text{sign}(Ax + \tau), \ \|z\|_2 \leq R$$

satisfies $\|x - x_{\text{CP}}^{\#}\|_2 \leq R\varepsilon^{1/8}$.

Our analysis relies on an observation of Foucart [8], who showed that it is sufficient for the matrix A to satisfy an ℓ_1/ℓ_2-*RIP property* to guarantee successful uniform recovery via (HT) and (LP). In the same vein, we show that the program (CP) is guaranteed to succeed under an ℓ_1/ℓ_2-RIP property for a modification of A. We prove the required RIP-properties in Theorem 5.2. The final section of the paper discusses two additional consequences of these RIP-results. In Corollary 6.1 we follow the work [6] to derive a recovery guarantee for (unquantized) outlier-robust compressed sensing with Gaussian circulant matrices. In Theorem 6.2 we use a recent result from [12] to derive an improved guarantee for recovery from uniform scalar quantized Gaussian circulant measurements.

2. Related work

Standard compressive sensing with partial circulant matrices. In standard (unquantized) compressive sensing, the task is to recover an (approximately) sparse vector $x \in \mathbb{R}^N$ from measurements $y = Ax$, where $A \in \mathbb{R}^{m \times N}$ with $m \ll N$. A number of reconstruction algorithms have been introduced, most notably ℓ_1-minimization which computes the minimizer of

$$\min_{z \in \mathbb{R}^N} \|z\|_1 \quad \text{subject to } Az = Ax.$$

The $(\ell_2$-)restricted isometry property is a classical way of analyzing the performance of compressive sensing [9]. The restricted isometry constant δ_s is defined as the smallest constant δ such that

$$(3) \qquad (1 - \delta)\|x\|_2^2 \leq \|Ax\|_2^2 \leq (1 + \delta)\|x\|_2^2 \quad \text{for all } s\text{-sparse } x \in \mathbb{R}^N.$$

If $\delta_{2s} < 1/\sqrt{2}$ then all s-sparse signals can be reconstructed via ℓ_1-minimization exactly, see e.g. [5, 9]. Stability under noise and sparsity defect can be shown as well and similar guarantees also hold for other reconstruction algorithms [9]. It is well-known that Gaussian random matrices satisfy $\delta_s \leq \delta$ with probability at least $1 - \eta$ if $m \gtrsim \delta^{-2}(s \log(eN/s) + \log(1/\eta))$ [9, Chapter 9].

The situation that A is a subsampled random circulant matrix (see below for a formal definition) has been analyzed in several contributions [24, 22, 23, 16, 14, 18]. The best available result states [16] that a properly normalized (deterministically) subsampled random circulant matrix (generated by a Gaussian random vector) satisfies $\delta_s \leq \delta$ with probability at least $1 - \eta$ if

$$m \gtrsim \delta^{-2} s (\log^2(s) \log^2(N) + \log(1/\eta)).$$

The original contribution [24] by Romberg uses random subsampling of a circulant matrix and requires slightly more logarithmic factors, but is able to treat sparsity with respect to an arbitrary basis. In the case of randomly subsampled random convolutions and sparsity with respect to the standard basis, stable and robust s-sparse recovery via ℓ_1-minimization could recently be shown via the null space property [9] in [18] in a small sparsity regime $s \lesssim \sqrt{N/\log(N)}$ under the optimal condition

$$(4) \qquad\qquad\qquad m \gtrsim s \log(eN/s).$$

We note that the proof in [18] provides the lower RIP-bound in (3) and may be extended to show the lower ℓ_1/ℓ_2 RIP bound in (8) below under condition (4).

Non-uniform recovery results have been shown in [22, 23, 14] which require only $m \gtrsim s \log(N)$ measurements for exact recovery from (deterministically) subsampled random convolutions via ℓ_1-minimization.

One-bit compressive sensing with Gaussian measurements, $\tau = 0$. The majority of the known signal reconstruction results in one-bit compressed sensing are restricted to standard Gaussian measurement matrices. Let us first consider the results in the case $\tau = 0$. It was shown in [13, Theorem 2] that if A is $m \times N$ Gaussian and $m \gtrsim \delta^{-1} s \log(N/\delta)$ then, with high probability, any s-sparse x, x' with $\|x\|_2 = \|x'\|_2 = 1$ and $\text{sign}(Ax) = \text{sign}(Ax')$ satisfy $\|x - x'\|_2 \leq \delta$. In particular, this shows that one can approximate x up to error δ by the solution of the non-convex program

$$\min \|z\|_0 \qquad \text{s.t.} \qquad \text{sign}(Ax) = \text{sign}(Az), \ \|z\|_2 = 1.$$

This result is near-optimal in the following sense: any reconstruction $x^\#$ based on $\text{sign}(Ax)$ satisfies $\|x^\# - x\|_2 \gtrsim s/(m + s^{3/2})$ [13, Theorem 1]. That is, the dependence of m on δ can in general not be improved. It was shown in [10, Theorem 7] that this optimal error dependence can be obtained using a polynomial time

algorithm if the measurement matrix is modified. Specifically, they showed that if $m \gtrsim \delta^{-1} m' \log(m'/\delta)$ if $A = A_2 A_1$, where A_2 is $m \times m'$ Gaussian and A_1 is any $m' \times N$ matrix with RIP constant bounded by $1/6$ (so one can take $m' \sim s \log(N/s)$ if A_1 is Gaussian), then with high probability one can recover any s-sparse x with unit norm up to error δ from $\text{sign}(Ax)$ using an efficient algorithm. To recover efficiently from Gaussian one-bit measurements, Plan and Vershynin [20] proposed the reconstruction program in (LP). They showed that using $m \gtrsim \delta^{-1} s \log^2(N/s)$ Gaussian measurements one can recover every x with $\|x\|_1 \leq \sqrt{s}$ and $\|x\|_2 = 1$ via (LP) with reconstruction error $\delta^{1/5}$. In [21] they introduced a different convex program and showed that if $m \gtrsim \delta^{-1} s \log(N/s)$, then one can achieve a reconstruction error $\delta^{1/6}$ even if there is (adversarial) quantization noise present.

Thresholds. It was recently shown that one can recover full signals (instead of just their directions) by incorporating appropriate thresholds. In [15] it was shown that by taking Gaussian thresholds τ_i one can recover energy information by slightly modifying the linear program (LP). A similar observation for recovery using the program (CP) was made in [2]. The paper [15] also proposed a method to estimate $\|x\|_2$ using a single deterministic threshold $\tau_i = \tau$ that works well if one has some prior knowledge of the energy range.

Subgaussian measurements. The results described above are all restricted to Gaussian measurements. It seems that [1] is currently the only work on memoryless one-bit compressed sensing for non-Gaussian matrices. Even though one-bit compressed sensing can fail in general for subgaussian matrices, it is shown in [1] that some non-uniform recovery results from [21] can be extended to subgaussian matrices if the signal to be recovered is not too sparse (meaning that $\|x\|_\infty$ is small) or if the measurement vectors are close to Gaussian in terms of the total variation distance.

Uniform scalar quantization. Some recovery results for circulant matrices are essentially known for a different memoryless quantization scheme. Consider the uniform scalar quantizer $Q_\delta : \mathbb{R}^m \to (\delta\mathbb{Z} + \delta/2)^m$ defined by $Q_\delta(z) = \left(\delta\lfloor z_i/\delta \rfloor + \delta/2\right)_{i=1}^m$. As we point out in Appendix A, if A consists of $m \gtrsim s \log^2 s \log^2 N$ deterministic samples of a subgaussian circulant matrix, then it follows from [16] that with high probability one can recover any s-sparse vector up to a reconstruction error δ from its quantized measurements $Q_\delta(Ax)$. By using random subsampling and imposing a small sparsity assumption similar to ours, this number of measurements can be decreased to $m \gtrsim s \log(N/s)$ [18]. In these results, the recovery error does not improve beyond the resolution δ of the quantizer, even if one takes more measurements. In Theorem 6.2 we show that for a randomly subsampled Gaussian circulant matrix it is possible to achieve a reconstruction error decay beyond the quantization resolution, provided that one introduces an appropriate dithering in the quantizer.

Adaptive quantization methods. The results discussed above all concern *memoryless* quantization schemes, meaning that each measurement is quantized independently. By quantizing adaptively based on previous measurements, one can improve the reconstruction error decay rate. In [2] it was shown for Gaussian measurement matrices that by using adaptive thresholds, one can even achieve an (optimal) exponential decay in terms of the number of measurements. Very recently, it was shown that one can efficiently recover a signal from randomly subsampled subgaussian partial circulant measurements that have been quantized using a popular scheme called sigma-delta quantization [7]. In particular, [7, Theorem 5] proves that based on $m \sim s \log^2 s \log^2 N$ one-bit sigma-delta quantized measurements, one can use a convex program to find an approximant of the signal that exhibits

polynomial reconstruction error decay. Although adaptive methods such as sigma-delta quantization can achieve a better error decay than memoryless quantization schemes, they require a more complicated hardware architecture and higher energy consumption in operation than the memoryless quantizer studied in this work.

3. Notation

We use Id_N to denote the $N \times N$ identity matrix. If $A \in \mathbb{R}^{m \times N}$ and $B \in \mathbb{R}^{m \times M}$, then $[A\ B] \in \mathbb{R}^{m \times (N+M)}$ is the matrix obtained by concatenating A and B. For $x \in \mathbb{R}^N$ and $y \in \mathbb{R}^M$ let $[x, y] \in \mathbb{R}^{N+M}$ be vector obtained by appending y at the end of x. If $I \subset [N]$, then $x_I \in \mathbb{R}^{|I|}$ is the vector obtained by restricting x to its entries in I. We let $R_I : \mathbb{R}^N \to \mathbb{R}^{|I|}$, $R_I(x) = x_I$ denote the restriction operator. Further, $\|x\|_p$ denotes the usual ℓ_p-norm of a vector x, $\|A\|_{\ell_2 \to \ell_2}$ the spectral norm of a matrix A and $\|A\|_F$ its Frobenius norm. For an event E, 1_E is the characteristic function of E.

We let $\Sigma_{s,N}$ denote the set of all s-sparse vectors with unit norm. We say that $x \in \mathbb{R}^N$ is s-*effectively sparse* if $\|x\|_1 \leq \sqrt{s}\|x\|_2$. We let $\Sigma_{s,N}^{\mathrm{eff}}$ denote the set of all s-effectively sparse vectors. Clearly, if x is s-sparse, then it is s-effectively sparse. We let H_s denote the hard thresholding operator, which sets all coefficients of a vector except the s largest ones (in absolute value) to 0.

For any $x \in \mathbb{R}^N$ we let $\Gamma_x \in \mathbb{R}^{N \times N}$ and $D_x \in \mathbb{R}^N$ be the circulant matrix and diagonal matrix, respectively, generated by x. That is,

$$D_x = \begin{bmatrix} x_1 & 0 & \cdots & 0 & 0 \\ 0 & x_1 & 0 & & 0 \\ 0 & 0 & x_2 & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & & 0 & x_N \end{bmatrix}, \Gamma_x = \begin{bmatrix} x_N & x_1 & x_2 & \cdots & x_{N-2} & x_{N-1} \\ x_{N-1} & x_N & x_1 & \cdots & x_{N-3} & x_{N-2} \\ x_{N-2} & x_{N-1} & x_N & \cdots & x_{N-4} & x_{N-3} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_1 & x_2 & x_3 & \cdots & x_{N-1} & x_N \end{bmatrix}.$$

We study the following linear measurement matrix. We consider a vector θ of i.i.d. random selectors with mean m/N and let $I = \{i \in [N] : \theta_i = 1\}$. Let g be an N-dimensional standard Gaussian vector that is independent of θ. We define the *randomly subsampled Gaussian circulant matrix* by $A = R_I \Gamma_g$. Note that $\mathbb{E}|I| = m$, so m corresponds to the expected number of measurements in this model.

4. Recovery via $\mathrm{RIP}_{1,2}$-properties

Let us start by stating our main recovery result for vectors with small sparsity located on the unit sphere.

Theorem 4.1. *Let* $0 < \delta, \eta \leq 1$ *and* $s \in [N]$ *such that*

$$(5) \qquad s \lesssim \min\{\sqrt{\delta^2 N/\log(N)}, \delta^2 N/\log(1/\eta)\}.$$

If $0 < \delta \leq \big(\log(s)\sqrt{\log(N)}\big)^{-1}$ *suppose that*

$$(6) \qquad m \gtrsim \delta^{-2} s \log(eN/(s\delta\eta)).$$

If $\big(\log(s)\sqrt{\log(N)}\big)^{-1} < \delta \leq 1$ *suppose that*

$$(7) \quad m \gtrsim \delta^{-1} s \max\left\{ \log(s)\log^{3/2}(N), \frac{\log(1/\eta)}{\log(s)\sqrt{\log(N)}}, \frac{\log(1/\eta)\log(s)\sqrt{\log(N)}}{s} \right\}.$$

Let $A = R_I \Gamma_g$. *Then, with probability at least* $1 - \eta$, *for every* $x \in \mathbb{R}^N$ *with* $\|x\|_0 \leq s$ *and* $\|x\|_2 = 1$, *the hard thresholding reconstruction* $x_{HT}^{\#}$ *satisfies* $\|x - x_{HT}^{\#}\|_2 \lesssim \sqrt{\delta}$.

Let us remark that for polynomially scaling probabilities $\eta = N^{-\alpha}$ the second and third term in the maximum in (7) can be bounded by a constant c_α times the first term and then (7) reduces to

$$m \gtrsim \delta^{-1} s \log(s) \log^{3/2}(N),$$

which is implied by the even simpler condition $m \gtrsim \delta^{-1} s \log^{5/2}(N)$. We further note that (5) also gives an implicit condition on δ. In fact, if $\delta \lesssim \sqrt{\log(N)/N}$, then (5) excludes all nontrivial sparsities $s \geq 1$. However, we expect that the requirement (5) of small sparsity is only an artefact of our proof and that recovery can also be expected for larger sparsities under conditions similar to (6) and (7) with possibly more logarithmic factors, see also [18] for an analogue phenomenon for standard compressed sensing. In fact, our proof relies on the $\text{RIP}_{1,2}$ and [18] provides at least the lower $\text{RIP}_{1,2}$ bound also for larger sparsities.

Let us now state our main result for the LP-reconstruction which unfortunately requires worse scaling in δ than (6) and (7).

Theorem 4.2. *Let $0 < \delta, \eta \leq 1$. If $0 < \delta \leq (\log^2(s) \log(N))^{-1/4}$ assume that*

$$s \lesssim \min\{\sqrt{\delta^4 N / \log(N)}, \delta^2 N / \log(1/\eta)\},$$
$$m \gtrsim \delta^{-4} s \log(eN/s),$$

and if $(\log^2(s) \log(N))^{-1/4} < \delta \leq 1$ assume that

$$s \lesssim \min\left\{\delta^{4/3} \sqrt{N / \log^2(N)}, \delta^2 N / \log(1/\eta)\right\}$$

$$m \gtrsim \delta^{-4/3} s \max\left\{\log^{4/3}(s) \log^{5/3}(N), \frac{\log(1/\eta)}{\log^{3/2}(s) \log^{1/3}(N)}, \frac{\log^{2/3}(s) \log^{1/3}(N)}{\delta^{4/3} s}\right\}.$$

Then the following holds with probability exceeding $1 - \eta$: for every $x \in \mathbb{R}^N$ with $\|x\|_1 \leq s$ and $\|x\|_2 = 1$, the LP-reconstruction $x_{LP}^{\#}$ satisfies $\|x - x_{LP}^{\#}\|_2 \lesssim \sqrt{\delta}$.

We will prove Theorem 4.1 by using a recent observation of Foucart [8]. He showed that one can accurately recover signals from one-bit measurements if the measurement matrix satisfies an appropriate RIP-type property. Let us say that a matrix A *satisfies* $\text{RIP}_{1,2}(s, \delta)$ if

$$(8) \qquad (1 - \delta)\|x\|_2 \leq \|Ax\|_1 \leq (1 + \delta)\|x\|_2, \qquad \text{for all } x \in \Sigma_{s,N}$$

and A *satisfies* $\text{RIP}_{1,2}^{eff}(s, \delta)$ if

$$(9) \qquad (1 - \delta)\|x\|_2 \leq \|Ax\|_1 \leq (1 + \delta)\|x\|_2, \qquad \text{for all } x \in \Sigma_{s,N}^{\text{eff}}.$$

Lemma 4.3. [8, Theorem 8] *Suppose that A satisfies $\text{RIP}_{1,2}(2s, \delta)$. Then, for every $x \in \mathbb{R}^N$ with $\|x\|_0 \leq s$ and $\|x\|_2 = 1$, the hard thresholding reconstruction $x_{HT}^{\#}$ satisfies $\|x - x_{HT}^{\#}\|_2 \leq 2\sqrt{5\delta}$.*

Let $\delta \leq 1/5$. Suppose that A satisfies $\text{RIP}_{1,2}^{eff}(9s, \delta)$. Then, for every $x \in \mathbb{R}^N$ with $\|x\|_1 \leq s$ and $\|x\|_2 = 1$, the LP-reconstruction $x_{LP}^{\#}$ satisfies $\|x - x_{LP}^{\#}\|_2 \leq 2\sqrt{5\delta}$.

Remark 4.4. *It is in general not possible to extend Theorem 4.1 to subgaussian circulant matrices. Indeed, suppose that the measurement matrix is a (rescaled, subsampled) Bernoulli circulant matrix, the threshold vector τ in (1) is zero and consider, for $0 < \lambda < 1$, the normalized 2-sparse vectors*

$$(10) \quad x_{+\lambda} = (1 + \lambda^2)^{-1/2}(1, \lambda, 0, \ldots, 0), \qquad x_{-\lambda} = (1 + \lambda^2)^{-1/2}(1, -\lambda, 0, \ldots, 0).$$

Then

$$\text{sign}(\langle (\Gamma_\varepsilon)_i, x_{+\lambda} \rangle) = \text{sign}(\varepsilon_{n+1-i} + \lambda \varepsilon_{n+2-i})$$

$$= \mathrm{sign}(\varepsilon_{n+1-i}) = \mathrm{sign}(\varepsilon_{n+1-i} - \lambda\varepsilon_{n+2-i}) = \mathrm{sign}(\langle(\Gamma_\varepsilon)_i, x_{-\lambda}\rangle)$$

That is, $x_{+\lambda}$ and $x_{-\lambda}$ produce identical one-bit measurements.

As a consequence, Lemma 4.3 implies that a subsampled Bernoulli circulant matrix cannot satisfy the $RIP_{1,2}(4,\delta)$ property for small values of δ, regardless of how we subsample and scale the matrix. Indeed, suppose that $A = \alpha R_I \Gamma_\varepsilon$ satisfies this property for a suitable $I \subset [N]$ and scaling factor α. Since $\mathrm{sign}(Ax_{+\lambda}) = \mathrm{sign}(Ax_{-\lambda})$, we find using Lemma 4.3 and the triangle inequality

$$\begin{aligned}
\frac{2\lambda}{(1+\lambda^2)^{1/2}} &= \|x_{+\lambda} - x_{-\lambda}\|_2 \\
&\leq \|x_{+\lambda} - H_s(A^* \mathrm{sign}(Ax_{+\lambda}))\|_2 + \|H_s(A^* \mathrm{sign}(Ax_{-\lambda})) - x_{-\lambda}\|_2 \\
&\leq 4\sqrt{5}\delta.
\end{aligned}$$

By taking $\lambda \to 1$ we find $\delta \geq 1/40$.

However, by excluding extremely sparse vectors via a suitable ℓ_∞-norm bound, it might be possible to work around this counter example. In fact, in the case of unstructured subgaussian random matrices, positive recovery results for sparse vectors with such additional constraint were shown in [1].

So far, our recovery results only allow to recover vectors lying on the unit sphere. By incorporating Gaussian dithering in the quantization process we can reconstruct any effectively sparse vector, provided that we have an a-priori upper bound on its energy.

Theorem 4.5. *Let $A = R_I \Gamma_g$ and let $\tau_1, \dots, \tau_m$ be independent $\mathcal{N}(0, \pi R^2/2)$-distributed random variables. Under the assumptions on s, m, n, δ, η of Theorem 4.2 the following holds with probability exceeding $1 - \eta$: for any $x \in \mathbb{R}^N$ with $\|x\|_1 \leq \sqrt{s}\|x\|_2$ and $\|x\|_2 \leq R$, any solution $x_{CP}^{\#}$ to the second-order cone program (CP) satisfies $\|x - x_{CP}^{\#}\|_2 \leq R\sqrt{\delta}$.*

To prove this result, we let $C \in \mathbb{R}^{m \times (N+1)}$ and consider the following abstract version of (CP):

$$(11) \qquad \min_{z \in \mathbb{R}^N} \|z\|_1 \qquad \text{s.t.} \qquad \mathrm{sign}(C[z, R]) = \mathrm{sign}(C[x, R]), \ \|z\|_2 \leq R.$$

It is straightforward to verify that (CP) is obtained by taking $C = \frac{1}{m}\sqrt{\frac{\pi}{2}}B$, where

$$(12) \qquad B := R_I[\Gamma_g \ h] = R_I \begin{bmatrix} g_N & g_1 & g_2 & \cdots & g_{N-2} & g_{N-1} & h_1 \\ g_{N-1} & g_N & g_1 & \cdots & g_{N-3} & g_{N-2} & h_2 \\ g_{N-2} & g_{N-1} & g_N & \cdots & g_{N-4} & g_{N-3} & h_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ g_1 & g_2 & g_3 & \cdots & g_{N-1} & g_N & h_N \end{bmatrix},$$

and h is a standard Gaussian vector that is independent of θ and g.

Lemma 4.6. *Let $\delta < 1/5$. Suppose that C satisfies $RIP_{1,2}^{eff}(36(\sqrt{s}+1)^2, \delta)$. Then, for any $x \in \mathbb{R}^N$ satisfying $\|x\|_1 \leq \sqrt{s}\|x\|_2$ and $\|x\|_2 \leq R$, any solution $x^{\#}$ to (11) satisfies*

$$\|x - x^{\#}\|_2 \leq 2R\sqrt{\delta}.$$

The following proof is based on arguments in [8, Section 8.4] and [2, Corollary 9].

Proof. In the proof of [2, Corollary 9] it shown that

$$(13) \qquad \|u - v\|_2 \leq 2\left\|\frac{[u,1]}{\|[u,1]\|_2} - \frac{[v,1]}{\|[v,1]\|_2}\right\|_2$$

for any two vectors $u, v \in B_{\ell_2^N}$. Let $x \in \Sigma_{s,N}^{\mathrm{eff}} \cap RB_{\ell_2^N}$, $x^\#$ be any solution to (11) and write

$$\bar{x} = [x, R]/\|[x, R]\|_2, \qquad \bar{x}^\# = [x^\#, R]/\|[x^\#, R]\|_2.$$

Since $x/R, x^\#/R \in B_{\ell_2^N}$, (13) implies that

$$\|x - x^\#\|_2 \le 2R\|\bar{x} - \bar{x}^\#\|_2.$$

By the parallellogram identity,

$$\text{(14)} \qquad \left\|\frac{\bar{x} - \bar{x}^\#}{2}\right\|_2^2 = \frac{\|\bar{x}\|_2^2 + \|\bar{x}^\#\|_2^2}{2} - \left\|\frac{\bar{x} + \bar{x}^\#}{2}\right\|_2^2.$$

Let us observe that $[x, R]$ and $[x^\#, R]$ are $(\sqrt{s} + 1)^2$-effectively sparse. Indeed, by optimality of $x^\#$ for (11) and s-effective sparsity of x,

$$\|[x^\#, R]\|_1 \le \|[x, R]\|_1 \le \sqrt{s}\|x\|_2 + R \le R(\sqrt{s} + 1)$$

and $\|[x, R]\|_2, \|[x^\#, R]\|_2 \ge R$. We claim that

$$\text{(15)} \qquad z := \frac{\bar{x} + \bar{x}^\#}{2} \in \Sigma_{36(\sqrt{s}+1)^2, N+1}^{\mathrm{eff}}.$$

Once this is shown, we can use $\mathrm{sign}(C\bar{x}) = \mathrm{sign}(C\bar{x}^\#)$ and the $\mathrm{RIP}_{1,2}^{\mathrm{eff}}(36(\sqrt{s} + 1)^2, \delta)$-property of C to find

$$\text{(16)} \qquad \left\|\frac{\bar{x} + \bar{x}^\#}{2}\right\|_2 \ge \frac{1}{1 + \delta}\left\|C\left(\frac{\bar{x} + \bar{x}^\#}{2}\right)\right\|_1 = \frac{\|C\bar{x}\|_1 + \|C\bar{x}^\#\|_1}{2(1 + \delta)} \ge \frac{(1 - \delta)}{(1 + \delta)}.$$

Hence, (14) implies

$$\left\|\frac{\bar{x} - \bar{x}^\#}{2}\right\|_2^2 \le 1 - \frac{(1 - \delta)^2}{(1 + \delta)^2} \le \frac{4\delta}{(1 + \delta)^2}.$$

Let us now prove (15). Since $[x, R]$ and $[x^\#, R]$ are $(\sqrt{s} + 1)^2$-effectively sparse,

$$\|z\|_1 \le \frac{1}{2}\left\|\frac{[x, R]}{\|[x, R]\|_2}\right\|_1 + \frac{1}{2}\left\|\frac{[x^\#, R]}{\|[x^\#, R]\|_2}\right\|_1 \le \sqrt{s} + 1.$$

It remains to bound $\|z\|_2$ from below. In (16) we already observed that

$$\text{(17)} \qquad \|Cz\|_1 = \frac{1}{2}\left\|\frac{C[x, R]}{\|[x, R]\|_2}\right\|_1 + \frac{1}{2}\left\|\frac{C[x^\#, R]}{\|[x^\#, R]\|_2}\right\|_1 \ge (1 - \delta).$$

Set $t = 8s + 8 \ge 4(\sqrt{s} + 1)^2$. Let T_0 be the index set corresponding to the t largest entries of z, T_1 be the set corresponding to the next t largest entries of z, and so on. Then, for all $k \ge 1$,

$$\|z_{T_k}\|_2 \le \sqrt{t}\|z_{T_k}\|_\infty \le \|z_{T_{k-1}}\|_1/\sqrt{t}.$$

Since C satisfies $\mathrm{RIP}_{1,2}^{\mathrm{eff}}(36(\sqrt{s} + 1)^2, \delta)$, it satisfies $\mathrm{RIP}_{1,2}(t, \delta)$ and hence

$$\|Cz\|_1 \le \sum_{k \ge 0}\|Cz_{T_k}\|_1 \le (1 + \delta)\Big(\|z_{T_0}\|_2 + \sum_{k \ge 1}\|z_{T_k}\|_2\Big)$$

$$\le (1 + \delta)\|z\|_2 + \frac{(1 + \delta)}{\sqrt{t}}\|z\|_1 \le (1 + \delta)\|z\|_2 + \frac{(1 + \delta)}{\sqrt{t}}(\sqrt{s} + 1).$$

$$\text{(18)} \qquad \le (1 + \delta)\|z\|_2 + \frac{1}{2}(1 + \delta).$$

Since $\delta \le 1/5$, (17) and (18) together yield

$$\|z\|_2 \ge \frac{(1 - \delta) - \frac{1}{2}(1 + \delta)}{(1 + \delta)} = \frac{\frac{1}{2} - \frac{3}{2}\delta}{1 + \delta} \ge \frac{1}{6}.$$

$\square$

5. PROOF OF THE $\mathrm{RIP}_{1,2}$-PROPERTIES

We will now prove the RIP-properties required in Lemmas 4.3 and 4.6. For our analysis we recall two standard concentration inequalities. The Hanson-Wright inequality [11] states that if g is a standard Gaussian vector and $B \in \mathbb{R}^{N \times N}$, then

$$(19) \qquad \mathbb{P}(|g^T B g - \mathbb{E} g^T B g| \geq t) \leq \exp\left(-c \min\left\{\frac{t^2}{\|B\|_F^2}, \frac{t}{\|B\|_{\ell_2 \to \ell_2}}\right\}\right),$$

where c is an absolute constant and $\|B\|_F$ and $\|B\|_{\ell_2 \to \ell_2}$ are the Frobenius and operator norms of B, respectively. We refer to [25] for a modern proof. In addition, we will use a well-known concentration inequality for suprema of Gaussian processes (see e.g. [3, Theorem 5.8]). If $T \subset \mathbb{R}^N$, then

$$(20) \qquad \mathbb{P}\left(\left|\sup_{x \in T}\langle x, g\rangle - \mathbb{E}\sup_{x \in T}\langle x, g\rangle\right| \geq t\right) \leq 2e^{-t^2/2\sigma^2},$$

where $\sigma^2 = \sup_{x \in T}\|x\|_2^2$. We will use the following observation.

Lemma 5.1. *Suppose that $y \in \mathbb{R}^N$ satisfies $\|y\|_1 \leq \sqrt{s}$ and $\|y\|_2 = 1$. Let g be a standard Gaussian vector. For any $t > 0$,*

$$(21) \qquad \mathbb{P}\left(\left|\frac{1}{N}\|\Gamma_g y\|_2^2 - 1\right| \geq t\right) \leq 2e^{-cN \min\{t^2, t\}/s}$$

and

$$(22) \qquad \mathbb{P}\left(\left|\frac{1}{N}\sqrt{\frac{\pi}{2}}\|\Gamma_g y\|_1 - 1\right| \geq t\right) \leq 2e^{-Nt^2/\pi s}.$$

If h is a standard Gaussian vector and $y \in \mathbb{R}^{N+1}$ satisfies $\|y\|_1 \leq \sqrt{s}$ and $\|y\|_2 = 1$, then

$$(23) \qquad \mathbb{P}\left(\left|\frac{1}{N}\sqrt{\frac{\pi}{2}}\|[\Gamma_g \ h]y\|_1 - 1\right| \geq t\right) \leq 2e^{-Nt^2/\pi s}.$$

Proof. Note that $\Gamma_g y = \Gamma_y g$. By the Hanson-Wright inequality (19)

$$\mathbb{P}\left(|\|\Gamma_y g\|_2^2 - N\|y\|_2^2| \geq Nt\right) \leq \exp\left(-c\min\left\{\frac{t^2 N^2}{\|\Gamma_y^* \Gamma_y\|_F^2}, \frac{tN}{\|\Gamma_y^* \Gamma_y\|_{\ell^2 \to \ell^2}}\cdot\right\}\right) .$$

Recall that the convolution satisfies $\|\Gamma_y z\|_2 = \|y * z\|_2 \leq \|y\|_1 \|z\|_2 \leq \sqrt{s}\|z\|_2$ for all $z \in \mathbb{R}^N$, which implies

$$(24) \qquad \|\Gamma_y^* \Gamma_y\|_{\ell_2 \to \ell_2} = \|\Gamma_y\|_{\ell_2 \to \ell_2}^2 \leq s,$$

$$\|\Gamma_y^* \Gamma_y\|_F \leq \|\Gamma_y\|_{\ell_2 \to \ell_2}\|\Gamma_y\|_F \leq \sqrt{s}\sqrt{N}\|y\|_2 = \sqrt{sN}.$$

To prove (22) observe that

$$\mathbb{E}\left(\frac{1}{N}\sqrt{\frac{\pi}{2}}\|\Gamma_g y\|_1\right) = \sqrt{\frac{\pi}{2}}\mathbb{E}|\langle g, y\rangle| = \|y\|_2 = 1$$

and

$$\|\Gamma_g y\|_1 = \sup_{z \in B_{\ell_\infty^N}}\langle z, \Gamma_g y\rangle = \sup_{z \in B_{\ell_\infty^N}}\langle z, \Gamma_y g\rangle = \sup_{z \in B_{\ell_\infty^N}}\langle \Gamma_y^* z, g\rangle.$$

Hence, by the concentration inequality (20) for suprema of Gaussian processes,

$$\mathbb{P}\left(\left|\frac{1}{N}\sqrt{\frac{\pi}{2}}\|\Gamma_g y\|_1 - 1\right| \geq t\right) \leq 2e^{-t^2/2\sigma_y^2},$$

where

$$\sigma_y^2 = \frac{1}{N^2}\frac{\pi}{2}\sup_{z \in B_{\ell_\infty^N}}\|\Gamma_y^* z\|_2^2.$$

For any $z \in B_{\ell_\infty^N}$, we obtain using (24)

$$\|\Gamma_y^* z\|_2^2 \leq \|\Gamma_y\|_{\ell_2 \to \ell_2}^2 \|z\|_2^2 \leq sN.$$

The proof of (23) is similar. By writing $y = [y_{[N]}, y_{N+1}]$ we find

$$[\Gamma_g \ h]y = \Gamma_g y_{[N]} + y_{N+1} h = [\Gamma_{y_{[N]}} \ y_{N+1} \mathrm{Id}_N][g, h].$$

It follows that

$$\|[\Gamma_g \ h]y\|_1 = \sup_{z \in B_{\ell_\infty^N}} \langle z, [\Gamma_{y_{[N]}} \ y_{N+1} \mathrm{Id}_N][g, h]\rangle = \sup_{z \in B_{\ell_\infty^N}} \langle [\Gamma_{y_{[N]}} \ y_{N+1} \mathrm{Id}_N]^* z, [g, h]\rangle.$$

Moreover,

$$\mathbb{E}\Big(\frac{1}{N}\sqrt{\frac{\pi}{2}}\|[\Gamma_g \ h]y\|_1\Big) = \|y\|_2 = 1.$$

The concentration inequality (20) for suprema of Gaussian processes now implies

$$\mathbb{P}\Big(\Big|\frac{1}{N}\sqrt{\frac{\pi}{2}}\|\Gamma_g y\|_1 - 1\Big| \geq t\Big) \leq 2e^{-t^2/2\sigma_y^2},$$

where

$$\sigma_y^2 = \frac{1}{N^2}\frac{\pi}{2}\sup_{z \in B_{\ell_\infty^N}} \|[\Gamma_{y_{[N]}} \ y_{N+1} \mathrm{Id}_N]^* z\|_2^2.$$

For any $z \in B_{\ell_\infty^N}$, we obtain using (24) and the Cauchy-Schwarz inequality,

$$\|[\Gamma_{y_{[N]}} \ y_{N+1} \mathrm{Id}_N]^* z\|_2^2 = \|\Gamma_{y_{[N]}}^* z\|_2^2 + y_{N+1}^2 \ \|z\|_2^2$$
$$\leq (\|\Gamma_{y_{[N]}}\|_{\ell_2 \to \ell_2}^2 + y_{N+1}^2)\|z\|_2^2 \leq (s\|y_{[N]}\|_2^2 + y_{N+1}^2)N \leq sN.$$

This completes the proof. $\square$

By Lemmas 4.3 and 4.6, our main recovery results in Theorems 4.1 and 4.5 are implied by the following theorem. As before, θ consists of i.i.d. random selectors with mean m/N, $I = \{i \in [N] \ : \ \theta_i = 1\}$ and g, h are independent n-dimensional standard Gaussian vectors that are independent of θ.

Theorem 5.2. *Fix $\delta > 0$. Let $A = R_I \Gamma_g$ be a randomly subsampled Gaussian circulant matrix and let $B = R_I[\Gamma_g \ h]$. Under the assumptions on s, m, N, δ, η of Theorem 4.1, $\frac{1}{m}\sqrt{\frac{\pi}{2}}A$ satisfies $RIP_{1,2}(s, \delta)$ with probability at least $1-\eta$. Moreover, under the assumptions of Theorem 4.2 $\frac{1}{m}\sqrt{\frac{\pi}{2}}A$ and $\frac{1}{m}\sqrt{\frac{\pi}{2}}B$ satisfy $RIP_{1,2}^{eff}(s, \delta)$ with probability at least $1 - \eta$.*

Proof. Let $\kappa > 0$ be a number to be chosen later. Let $\mathcal{N}_{\delta/(1+\kappa)} \subset \Sigma_{s,N}$ be a minimal $\delta/(1 + \kappa)$-net for $\Sigma_{s,N}$ with respect to the Euclidean norm. It is well known, see e.g. [9, Proposition C.3], that

$$\log |\mathcal{N}_{\delta/(1+\kappa)}| \leq s \log\left(\frac{3(1 + \kappa)eN}{s\delta}\right).$$

Fix $x \in \Sigma_{s,N}$ and let $y \in \mathcal{N}_{\delta/(1+\kappa)}$ be such that $\|x - y\|_2 \leq \delta/(1 + \kappa)$. We consider the events

$$E_I = \Big\{\frac{m}{2} \leq |I| \leq \frac{3m}{2}\Big\}$$

$$E_{\mathrm{RIP}} = \Big\{\forall \ z \in \Sigma_{2s,N} \ : \ \frac{1}{\sqrt{m}}\|Az\|_2 \leq 1 + \kappa\Big\}$$

(25)

$$E_{\Gamma,\ell_1} = \Big\{\forall y \in \mathcal{N}_{\delta/(1+\kappa)} \ : \ \Big|\frac{1}{N}\sqrt{\frac{\pi}{2}}\|\Gamma_g y\|_1 - 1\Big| \leq \delta\Big\}$$

$$E = \Big\{\forall y \in \mathcal{N}_{\delta/(1+\kappa)} \ : \ \Big|\frac{1}{m}\sqrt{\frac{\pi}{2}}\|Ay\|_1 - \frac{1}{N}\sqrt{\frac{\pi}{2}}\|\Gamma_g y\|_1\Big| \leq 2\delta\Big\}$$

respectively. Under E_I and E_{RIP},

$$\left| \frac{1}{m}\sqrt{\frac{\pi}{2}}\|Ax\|_1 - \frac{1}{m}\sqrt{\frac{\pi}{2}}\|Ay\|_1 \right| \leq \frac{\delta}{1+\kappa}\frac{1}{m}\sqrt{\frac{\pi}{2}}\left\|A\Big(\frac{x-y}{\|x-y\|_2}\Big)\right\|_1$$

$$\leq \frac{\delta}{1+\kappa}\frac{|I|}{m}\sup_{z\in\Sigma_{2s,N}}\frac{1}{\sqrt{|I|}}\sqrt{\frac{\pi}{2}}\|Az\|_2 \leq \sqrt{\frac{9\pi}{8}}\delta.$$

Therefore, if the events in (25) hold simultaneously, then by the triangle inequality

$$\left| \frac{1}{m}\sqrt{\frac{\pi}{2}}\|Ax\|_1 - 1 \right|$$

$$\leq \left| \frac{1}{m}\sqrt{\frac{\pi}{2}}\|Ax\|_1 - \frac{1}{m}\sqrt{\frac{\pi}{2}}\|Ay\|_1 \right| + \left| \frac{1}{m}\sqrt{\frac{\pi}{2}}\|Ay\|_1 - \frac{1}{N}\sqrt{\frac{\pi}{2}}\|\Gamma_g y\|_1 \right|$$

$$+ \left| \frac{1}{N}\sqrt{\frac{\pi}{2}}\|\Gamma_g y\|_1 - 1 \right| \leq (\sqrt{9\pi/8}+3)\delta.$$

Hence, it remains to show that the events in (25) hold with probability at least $1-\eta$. The Chernoff bound immediately yields $\mathbb{P}(E_I^c) \leq e^{-cm}$. By Theorem B.1, under the event E_I, if

$$(26) \qquad\qquad m \gtrsim \kappa^{-2}s(\log^2(s)\log^2(N)+\log(1/\eta))$$

then

$$\mathbb{P}_g(E_{\mathrm{RIP}}^c) = \mathbb{P}_g\Big(\exists z\in\Sigma_{2s,N} \ : \ \frac{1}{\sqrt{|I|}}\|R_I\Gamma_g z\|_2 \geq 1+\kappa\Big) \leq \eta$$

Moreover, by (22) and a union bound,

$$\mathbb{P}\Big(\sup_{y\in\mathcal{N}_{\delta/(1+\kappa)}}\left|\frac{1}{N}\sqrt{\frac{\pi}{2}}\|\Gamma_g y\|_1 - 1\right| \geq \delta\Big) \leq |N_{\delta/(1+\kappa)}|2e^{-\delta^2 N/\pi s}$$

$$\leq 2e^{s\log(3e(1+\kappa)N/(s\delta))-\delta^2 N/\pi s},$$

so $\mathbb{P}(E_{\Gamma,\ell_1}^c) \leq 2\eta$ if

$$(27) \qquad\qquad N \geq \pi\delta^{-2}s^2\log(3eN(1+\kappa)/s\delta) + \pi\delta^{-2}s\log(1/\eta).$$

Thus, it remains to show that $\mathbb{P}(E^c) \leq \eta$. To prove this, we consider

$$X_y = \frac{1}{m}\sqrt{\frac{\pi}{2}}\|D_\theta\Gamma_g y\|_1 - \frac{1}{N}\sqrt{\frac{\pi}{2}}\|\Gamma_g y\|_1$$

and $X_y' = \frac{1}{m}\sqrt{\frac{\pi}{2}}\|D_{\theta'}\Gamma_g y\|_1 - \frac{1}{N}\sqrt{\frac{\pi}{2}}\|\Gamma_g y\|_1$ for $y\in\mathcal{N}_\delta$, where θ' is an independent copy of θ. By symmetrization, see Lemma C.3,

$$\mathbb{P}_\theta\Big(\sup_{y\in\mathcal{N}_{\delta/(1+\kappa)}}|X_y| \geq 2\delta\Big) \leq \mathbb{P}_\theta\Big(\sup_{y\in\mathcal{N}_{\delta/(1+\kappa)}}|X_y - X_y'| \geq \delta\Big) + \sup_{y\in\mathcal{N}_{\delta/(1+\kappa)}}\mathbb{P}_\theta(|X_y| \geq \delta)$$

and so

$$(28) \qquad \mathbb{P}_{\theta,g}\Big(\sup_{y\in\mathcal{N}_{\delta/(1+\kappa)}}|X_y| \geq 2\delta\Big) \leq \mathbb{P}_{\theta,g}\Big(\sup_{y\in\mathcal{N}_{\delta/(1+\kappa)}}|X_y - X_y'| \geq \delta\Big)$$

$$+ \mathbb{E}_g\sup_{y\in\mathcal{N}_{\delta/(1+\kappa)}}\mathbb{P}_\theta(|X_y| \geq \delta).$$

To bound the first term on the right hand side, observe that $X_y - X_y'$ and

$$\frac{1}{m}\sqrt{\frac{\pi}{2}}\sum_{i=1}^N \varepsilon_i(\theta_i - \theta_i')|\langle a_i, y\rangle|,$$

where ε is a Rademacher vector, i.e., a vector of independent random signs, that is independent of g, θ and θ', are identically distributed and so

$$\mathbb{P}_{\theta,\theta',g}\Big(\sup_{y\in\mathcal{N}_{\delta/(1+\kappa)}}|X_y - X'_y| \geq \delta\Big)$$

$$= \mathbb{P}_{\varepsilon,\theta,\theta',g}\Big(\frac{1}{m}\sqrt{\frac{\pi}{2}}\sup_{y\in\mathcal{N}_{\delta/(1+\kappa)}}\Big|\sum_{i=1}^{N}\varepsilon_i(\theta_i - \theta'_i)|\langle a_i, y\rangle|\Big| \geq \delta\Big)$$

$$\leq 2\mathbb{P}_{\varepsilon,\theta,g}\Big(\frac{1}{m}\sqrt{\frac{\pi}{2}}\sup_{y\in\mathcal{N}_{\delta/(1+\kappa)}}\Big|\sum_{i=1}^{N}\varepsilon_i\theta_i|\langle a_i, y\rangle|\Big| \geq \delta/2\Big)$$

$$\leq 2\mathbb{P}_{\varepsilon,\theta,g}\Big(\frac{1}{m}\sqrt{\frac{\pi}{2}}\sup_{y\in\mathcal{N}_{\delta/(1+\kappa)}}\Big|\sum_{i=1}^{N}\varepsilon_i\theta_i 1_{E_{\mathrm{RIP}}}|\langle a_i, y\rangle|\Big| \geq \delta/2\Big) + \mathbb{P}_{\theta,g}(E^c_{\mathrm{RIP}}).$$

By Hoeffding's inequality and assuming E_I,

$$\mathbb{P}_\varepsilon\Big(\frac{1}{m}\sqrt{\frac{\pi}{2}}\Big|\sum_{i=1}^{N}\varepsilon_i\theta_i 1_{E_{\mathrm{RIP}}}|\langle a_i, y\rangle|\Big| \geq \delta/2\Big) \leq 2\exp\Big(-\frac{m^2\delta^2}{2\pi\sum_{i=1}^{N}\theta_i 1_{E_{\mathrm{RIP}}}|\langle a_i, y\rangle|^2}\Big)$$

$$= 2\exp\Big(-\frac{m\delta^2}{2\pi 1_{E_{\mathrm{RIP}}}\frac{1}{m}\|Ay\|_2^2}\Big) \leq 2e^{-\frac{m\delta^2}{2\pi(1+\kappa)^2}}$$

Hence a union bound yields

$$\mathbb{P}_{\varepsilon,\theta,g}\Big(\frac{1}{m}\sqrt{\frac{\pi}{2}}\sup_{y\in\mathcal{N}_{\delta/(1+\kappa)}}\Big|\sum_{i=1}^{N}\varepsilon_i\theta_i 1_{E_{\mathrm{RIP}}}|\langle a_i, y\rangle|\Big| \geq \delta/2\Big) \leq 2|\mathcal{N}_{\delta/(1+\kappa)}|e^{-\frac{m\delta^2}{2\pi(1+\kappa)^2}} \leq \eta$$

provided that

$$(29) \qquad m \gtrsim \frac{(1+\kappa)^2}{\delta^2}\left(s\log\left(\frac{3e(1+\kappa)N}{\delta s}\right) + \log(1/\eta)\right).$$

To bound the second term on the right hand side of (28), consider the event

$$E_{\Gamma,\ell_2} = \Big\{\forall y \in \mathcal{N}_{\delta/(1+\kappa)} \ : \ \frac{1}{\sqrt{N}}\|\Gamma_g y\|_2 \leq 2\Big\}.$$

By (21) and a union bound,

$$\mathbb{P}_g(E^c_{\Gamma,\ell_2}) \leq 2|\mathcal{N}_{\delta/(1+\kappa)}|e^{-cN/s} \leq \eta$$

under the condition $N \gtrsim \delta^{-2}s^2\log\left(\frac{3e(1+\kappa)N}{s\delta}\right) + s\log(1/\eta)$, which is weaker than (27). This shows that

$$\mathbb{E}_g\sup_{y\in\mathcal{N}_{\delta/(1+\kappa)}}\mathbb{P}_\theta(|X_y| \geq \delta) \leq \mathbb{E}_g\sup_{y\in\mathcal{N}_{\delta/(1+\kappa)}}\mathbb{P}_\theta(|X_y 1_{E_{\Gamma,\ell_2}}| \geq \delta) + \eta.$$

Now recall the following facts. If X is a random variable, X' is an independent copy, and $\mathrm{med}(X)$ is a median of X, then for any $\delta > 0$ (see Lemma C.2)

$$\mathbb{P}(|X - \mathrm{med}(X)| \geq \delta) \leq 2\mathbb{P}(|X - X'| \geq \delta)$$

and

$$|\mathrm{med}(X) - \mathbb{E}X| \leq (\mathbb{E}(X - \mathbb{E}X)^2)^{1/2}.$$

Combining these, we find

$$\mathbb{P}(|X - \mathbb{E}X| \geq \delta) \leq 2\mathbb{P}(|X - X'| \geq \delta - (\mathbb{E}(X - \mathbb{E}X)^2)^{1/2}).$$

We apply these inequalities with $X = X_y 1_{E_{\Gamma,\ell_2}}$ and $X' = X'_y 1_{E_{\Gamma,\ell_2}}$. Note that $\mathbb{E}_\theta X = 0$. By symmetrization, see e.g. [17, Lemma 6.3] or [9, Lemma 8.4],

$$(\mathbb{E}_\theta(X - \mathbb{E}_\theta X)^2)^{1/2} = \Big(\mathbb{E}_\theta\Big|\frac{1}{m}\sqrt{\frac{\pi}{2}}\|Ay\|_1 - \mathbb{E}_\theta\Big(\frac{1}{m}\sqrt{\frac{\pi}{2}}\|Ay\|_1\Big)\Big|^2\Big)^{1/2}1_{E_{\Gamma,\ell_2}}$$

$$\leq \frac{2}{m}\sqrt{\frac{\pi}{2}}\Big(\mathbb{E}_{\theta,\varepsilon}\Big|\sum_{i=1}^{N}\varepsilon_i\theta_i|\langle a_i,y\rangle|\Big|^2\Big)^{1/2}1_{E_{\Gamma,\ell_2}}$$

$$=\frac{2}{m}\sqrt{\frac{\pi}{2}}\Big(\mathbb{E}_\theta\sum_{i=1}^{N}\theta_i|\langle a_i,y\rangle|^2\Big)^{1/2}1_{E_{\Gamma,\ell_2}}=\frac{1}{\sqrt{m}}\sqrt{2\pi}\frac{1}{\sqrt{N}}\|\Gamma_g y\|_2 1_{E_{\Gamma,\ell_2}}$$

$$\leq\frac{2\sqrt{2\pi}}{\sqrt{m}}\leq\delta/2$$

if $m\geq 32\pi\delta^{-2}$ the latter being a weaker condition than (29). In summary, we find

$$\mathbb{P}_\theta(|X_y 1_{E_{\Gamma,\ell_2}}|\geq\delta)\leq 2\mathbb{P}(|X_y-X_y'|1_{E_{\Gamma,\ell_2}}\geq\delta/2)$$

$$\leq 4\mathbb{P}_{\theta,\varepsilon}\Big(\frac{1}{m}\sqrt{\frac{\pi}{2}}\sum_{i=1}^{N}\varepsilon_i\theta_i|\langle a_i,y\rangle|1_{E_{\Gamma,\ell_2}}\geq\frac{\delta}{4}\Big).$$

Now apply Hoeffding's inequality to obtain

$$\mathbb{P}_\theta(|X_y 1_{E_{\Gamma,\ell_2}}|\geq\delta)\leq 4\mathbb{E}_\theta\exp\Big(\frac{-m^2\frac{2}{\pi}\delta^2}{16\sum_{i=1}^{N}\theta_i|\langle a_i,y\rangle|^2}\Big)\quad=4\mathbb{E}_\theta\exp\Big(\frac{-m\delta^2}{8\pi\frac{1}{m}\|Ay\|_2^2}\Big)$$

$$\leq 4e^{-\frac{m\delta^2}{8\pi(1+\kappa)^2}}+4\mathbb{P}_\theta(E_{\mathrm{RIP}}^c).$$

If $m\geq 8\pi\frac{(1+\kappa)^2}{\delta^2}\log(1/\eta)$ (which is again weaker than (29)), we find

$$\mathbb{E}_g\sup_{y\in N_\delta}\mathbb{P}_\theta(|X_y|\geq\delta)\leq 5\eta+4\mathbb{P}_{\theta,g}(E_{\mathrm{RIP}}^c)\leq 9\eta.$$

We still need to choose $\kappa>0$ and distinguish two cases to this end.

Case 1: Assume that $0<\delta<\big(\log(s)\sqrt{\log(N)}\big)^{-1}$ and choose $\kappa=1$. A nontrivial $s\geq 1$ is only allowed by (5) if $\delta\gtrsim 1/\sqrt{N}$. In this situation we have $\log(3e(1+\kappa)N/(s\delta))\asymp\log(N)$. Then (5) implies (27). Moreover, (6) implies both (29) and (26) for our conditions on the parameters κ and δ. Altogether, we obtain $\mathbb{P}(E^c)\leq 10\eta$ in this case.

Case 2: If $\big(\log(s)\sqrt{\log(N)}\big)^{-1}<\delta\leq 1$ we choose $\kappa=\sqrt{\delta\log(s)}\log^{1/4}(N)>1$. Again a nontrivial $s\geq 1$ implies $\delta\gtrsim 1/\sqrt{N}$ by (5) and also in this case we have $\log(3e(1+\kappa)N/(s\delta))\asymp\log(N)$. Plugging our choice of κ into (29) and (26), we observe that both these conditions are implied by (7) and $\mathbb{P}(E^c)\leq 10\eta$.

The proof of the second statement for A is similar, so we only indicate the necessary changes in the argument. Let us write $C_{s,N}=\{x\in\mathbb{R}^N:\|x\|_1\leq\sqrt{s},\|x\|_2=1\}$. It clearly suffices to show that

$$\sup_{x\in C_{s,N}}\Big|\frac{1}{m}\sqrt{\frac{\pi}{2}}\|Ax\|_1-1\Big|\leq\delta$$

with probability at least $1-\eta$. Let us first recall that $C_{s,N}\subset 2\mathrm{conv}(\Sigma_{s,N})$ [20, Lemma 3.1]. Hence, under E_{RIP},

$$\frac{1}{\sqrt{m}}\sup_{z\in C_{s,N}}\|Az\|_2\leq 2\frac{1}{\sqrt{m}}\sup_{z\in\Sigma_{s,N}}\|Az\|_2\leq 2(1+\kappa).$$

We repeat the above argument with $\mathcal{N}_{\delta/(1+\kappa)}$ replaced by a minimal $\delta/(1+\kappa)$-net of $C_{s,N}$. Using $C_{s,N}\subset 2\mathrm{conv}(\Sigma_{s,N})$ and Sudakov's inequality, Theorem C.1, we find

$$(30)\quad\log|\mathcal{N}_{\delta/(1+\kappa)}|\lesssim\frac{(1+\kappa)^2}{\delta^2}\Big(\mathbb{E}\sup_{x\in C_{s,N}}\langle g,x\rangle\Big)^2\leq\frac{4(1+\kappa)^2}{\delta^2}\Big(\mathbb{E}\sup_{x\in\Sigma_{s,N}}\langle g,x\rangle\Big)^2$$

$$(31)\qquad\qquad\lesssim\frac{4(1+\kappa)^2}{\delta^2}s\log(eN/s),$$

where the final inequality is [21, Lemma 2.3]. By now chasing through the argument above we arrive at the three conditions

$$
\begin{aligned}
N &\gtrsim \frac{(1+\kappa)^2}{\delta^4} s^2 \log(eN/s) + s\frac{1}{\delta^2} \log(1/\eta) \\
m &\gtrsim \kappa^{-2} s(\log^2(s) \log^2(N) + \log(1/\eta)) \\
m &\gtrsim \frac{(1+\kappa)^4}{\delta^4} s \log(eN/s) + \frac{(1+\kappa)^2}{\delta^2} \log(1/\eta).
\end{aligned}
\tag{32}
$$

Again, we distinguish two cases depending on δ and choose κ as

$$
\kappa = \begin{cases} 1 & \text{if } 0 < \delta \leq (\log^2(s) \log(N))^{-1/4}, \\ (\delta^4 \log^2(s) \log(N))^{1/6} & \text{if } (\log^2(s) \log(N))^{-1/4} < \delta \leq 1. \end{cases}
$$

With this we can deduce the statement of the theorem (noting also that $\log(s) \leq \log(N)$).

Finally, let us prove the second statement for B. Let $N_{\delta/(1+\kappa)}$ be a minimal $\delta/(1+\kappa)$-net of $C_{s,N+1}$. By the first part of the proof, it is readily seen that the result will follow once we show that the events

$$
\begin{aligned}
E_{\mathrm{RIP},B} &= \left\{ \forall\, z \in \Sigma_{2s,N+1} \ : \ \frac{1}{\sqrt{m}} \|Bz\|_2 \leq 3 + \kappa \right\} \\
E_{\Gamma,h,\ell_2} &= \left\{ \forall y \in N_\delta \ : \ \frac{1}{\sqrt{N}} \|[\Gamma_g\ h]y\|_2 \leq 4 \right\} \\
E_{\Gamma,h,\ell_1} &= \left\{ \forall y \in N_\delta \ : \ \left| \frac{1}{N} \sqrt{\frac{\pi}{2}} \|[\Gamma_g\ h]y\|_1 - 1 \right| \leq \delta \right\}
\end{aligned}
\tag{33}
$$

hold with probability at least $1 - c\eta$. For E_{Γ,h,ℓ_1} this is immediate from (23) and a union bound. For $E_{\mathrm{RIP},B}$, observe that

$$
\frac{1}{\sqrt{m}} \|Bz\|_2 \leq \frac{1}{\sqrt{m}} \|Az_{[N]}\|_2 + |z_{N+1}| \frac{1}{\sqrt{m}} \sqrt{\frac{\pi}{2}} \|D_\theta h\|_2.
$$

We have already seen that the event $E_I = \{ \frac{m}{2} \leq |I| \leq \frac{3m}{2} \}$ holds with probability $1 - \eta$. Under this event, the Hanson-Wright inequality (19) yields

$$
\mathbb{P}_g\left(\frac{1}{m} \|D_\theta h\|_2^2 \geq 2 \right) \leq \mathbb{P}_g\left(\frac{1}{|I|} \|D_\theta h\|_2^2 \geq \frac{4}{3} \right) \leq e^{-c|I|} \leq e^{-m/2} \leq \eta
$$

for $m \gtrsim \log(1/\eta)$. Under the event E_{RIP} we have

$$
\frac{1}{\sqrt{m}} \sup_{z \in \Sigma_{2s,N}} \|Az_{[N]}\|_2 \leq 1 + \kappa
$$

with probability $1 - \eta$, so that, with probability at least $1 - 2\eta$,

$$
\frac{1}{\sqrt{m}} \|Bz\|_2 \leq (1+\kappa) \|z_{[N]}\|_2 + 2|z_{N+1}| \leq (3+\kappa) \|z\|_2
$$

under the conditions (32). Very similarly, one shows that E_{Γ,h,ℓ_2} holds with probability at least $1 - \eta$ under (32). As before, distinguishing two cases for δ one arrives at the statement of the theorem. □

6. Further applications

Apart from its usefulness for one-bit compressed sensing, the $\mathrm{RIP}_{1,2}$-property is of interest for (unquantized) outlier robust compressed sensing [6] and for compressed sensing involving uniformly scalar quantized measurements [12, 19]. In this section, we briefly sketch the implications of Theorem 5.2 for these two directions.

Corollary 6.1. *Let $A = R_I \Gamma_g$ be a randomly subsampled Gaussian circulant matrix. Let $0 < \eta < 1$ and $s \in [N]$ such that*

$$s \lesssim \min\left\{\sqrt{N/\log^2(N)}, N/\log(1/\eta)\right\}$$

and suppose that

$$(34) \quad m \gtrsim s \max\left\{\log^{4/3}(s)\log^{5/3}(N), \frac{\log(1/\eta)}{\log^{3/2}(s)\log^{1/3}(N)}, \frac{\log^{2/3}(s)\log^{1/3}(N)}{s}\right\}.$$

Then, with probability exceeding $1 - \eta$ the following holds: for any $x \in \mathbb{C}^n$ and $y = Ax + e$, where $\|e\|_1 \leq \varepsilon$, any solution $x^\#$ to

$$\min_{z \in \mathbb{C}^n} \|z\|_1 \quad \text{subject to} \quad \|y - Az\|_1 \leq \varepsilon.$$

satisfies

$$\|x - x^\#\|_2 \lesssim \frac{\sigma_s(x)_1}{\sqrt{s}} + \frac{\varepsilon}{m}.$$

Proof. As is argued in the proof of [6, Theorem III.3], it suffices to show that with probability at least $1 - \eta$,

$$(35) \quad \frac{1}{m}\|Ax\|_1 \geq c\|x\|_2, \quad \text{for all } x \in \Sigma_{s,N}^{\text{eff}}.$$

for a universal constant $c > 0$. Hence the result immediately follows from Theorem 5.2 by choosing $\delta = c$ constant. $\square$

Note that after estimating $\log(s) \leq \log(N)$ and for inverse polynomial probability $\eta = N^{-2}$, say, condition (34) takes the simpler form

$$m \gtrsim s\log^3(N).$$

In addition, we can use Theorem 5.2 to derive the following reconstruction result involving a uniform scalar quantizer that uses additional dithering. Let $Q_\delta : \mathbb{R}^m \to (\delta\mathbb{Z} + \delta/2)^m$ be the uniform scalar quantizer with resolution δ defined by $Q_\delta(z) = \left(\delta\lfloor z_i/\delta\rfloor + \delta/2\right)_{i=1}^m$.

Theorem 6.2. *Let $A = R_I\Gamma_g$ be a randomly subsampled Gaussian circulant matrix. Let τ be a vector of m independent $\mathcal{N}(0, \pi R^2/2)$-distributed random variables. Suppose that u is a vector of m independent random variables that are uniformly distributed on $[0, \delta]$ and are independent of τ. Assume that, for $0 < \eta, \epsilon \leq 1$,*

$$s \lesssim \min\left\{\sqrt{N/\log^2(N)}, N/\log(1/\eta)\right\}$$

$$(36) \quad m \gtrsim s \max\left\{\log^{4/3}(s)\log^{5/3}(N), \frac{\log(1/\eta)}{\log^{3/2}(s)\log^{1/3}(N)}, \frac{\log^{2/3}(s)\log^{1/3}(N)}{s},\right.$$
$$\left. R^2\delta^{-2}\epsilon^{-6}\log(1/\eta), \frac{\log(1/\eta)}{\epsilon^2 s}\right\}.$$

Then, with probability at least $1 - \eta$ the following holds: for any $x \in \mathbb{R}^N$ with $\|x\|_1 \leq \sqrt{s}\|x\|_2$ and $\|x\|_2 \leq R$, any solution $x^\#$ to the program

$$(37) \quad \min\|z\|_1 \quad \text{s.t.} \quad \|z\|_2 \leq R, \ Q_\delta\left(\sqrt{\frac{\pi}{2}}Az + \tau + u\right) = Q_\delta\left(\sqrt{\frac{\pi}{2}}Ax + \tau + u\right)$$

satisfies $\|x^\# - x\|_2 \lesssim \delta\epsilon$.

Note that the program (37) is convex. Indeed, the second condition in (37) is equivalent to

$$\frac{\delta}{2} \leq \Big(\sqrt{\frac{\pi}{2}}Az + \tau + u\Big)_i - Q_\delta\Big(\sqrt{\frac{\pi}{2}}Ax + \tau + u\Big)_i < \frac{\delta}{2}, \qquad i = 1, \ldots, m.$$

Proof. Let $x^\#$ be any solution to (37). Since x is feasible for (37),

$$\|[x^\#, R]\|_1 \leq \|[x, R]\|_1 \leq \sqrt{s}\|x\|_2 + R \leq R(\sqrt{s} + 1).$$

Since $R \leq \|[x^\#, R]\|_2, \|[x, R]\|_2 \leq \sqrt{2}R$, it follows that

$$[x^\#, R], [x, R] \in \Sigma^{\text{eff}}_{(\sqrt{s}+1)^2, N+1} \cap \sqrt{2}R\, B_{\ell_2^{N+1}}.$$

Moreover, the last condition in (37) precisely means that

$$Q_\delta\Big(\sqrt{\frac{\pi}{2}}B[x^\#, R] + u\Big) = Q_\delta\Big(\sqrt{\frac{\pi}{2}}B[x, R] + u\Big),$$

where $B = R_I[\Gamma_g\ h]$, where h is an independent standard Gaussian that is independent of g and θ. We will now show that this implies the reconstruction error bound.

Under (36), Theorem 5.2 implies that

$$\frac{1}{2}\|z\|_2 \leq \frac{1}{m}\sqrt{\frac{\pi}{2}}\|Bz\|_1 \leq \frac{3}{2}\|z\|_2 \qquad \text{for all } z \in \Sigma^{\text{eff}}_{(\sqrt{s}+1)^2, N+1}$$

with probability at least $1 - \eta$. By [12, Proposition 1] we obtain that under this event, $\sqrt{\frac{\pi}{2}}B$ satisfies with probability at least $1 - \eta$ a quantized version of $\text{RIP}_{1,2}$: for some universal constant c and any $z, z' \in \Sigma^{\text{eff}}_{(\sqrt{s}+1)^2, N+1} \cap \sqrt{2}R\, B_{\ell_2^{N+1}}$

$$(38) \quad \frac{1}{2}\|z - z'\|_2 - c\delta\epsilon \leq \Big\|Q_\delta\Big(\sqrt{\frac{\pi}{2}}Bz + u\Big) - Q_\delta\Big(\sqrt{\frac{\pi}{2}}Bz' + u\Big)\Big\|_1 \leq \frac{1}{2}\|z - z'\|_2 + c\delta\epsilon$$

provided that

$$(39) \qquad m \gtrsim \epsilon^{-2}\mathcal{N}(\Sigma^{\text{eff}}_{(\sqrt{s}+1)^2, N+1} \cap \sqrt{2}R\, B_{\ell_2^{N+1}}, \|\cdot\|_2, \delta\epsilon^2) + \epsilon^{-2}\log(1/\eta),$$

where $\mathcal{N}(\cdot, \|\cdot\|_2, t)$ denotes the covering number with respect to the Euclidean norm. Now observe that

$$\Sigma^{\text{eff}}_{(\sqrt{s}+1)^2, N+1} \cap \sqrt{2}R\, B_{\ell_2^{N+1}} \subset \sqrt{2}R\{x \in \mathbb{R}^{N+1} : \|x\|_1 \leq \sqrt{2s}, \|x\|_2 \leq 1\}$$

$$\subset 2\sqrt{2}R\,\text{conv}(\Sigma_{2s, N+1}),$$

where the final inclusion holds by [20, Lemma 3.1]. Therefore, similarly to (30), Sudakov's inequality (Theorem C.1) implies that

$$\log\mathcal{N}(\Sigma^{\text{eff}}_{(\sqrt{s}+1)^2, N+1} \cap \sqrt{2}R\, B_{\ell_2^{N+1}}, \|\cdot\|_2, \delta\epsilon^2) \lesssim R^2\delta^{-2}\epsilon^{-4}s\log(eN/s)$$

and it follows that (39) is satisfied under our assumptions (36). We can now apply (38) with $z = [x, 1]$ and $z' = [x^\#, 1]$ and use $\|z - z'\|_2 = \|x - x^\#\|_2$ to obtain the asserted error bound. □

APPENDIX A. UNIFORM SCALAR QUANTIZATION AND RECOVERY VIA ℓ_∞-CONSTRAINED ℓ_1-MINIMIZATION

Let $Q_\delta : \mathbb{R}^m(\delta\mathbb{Z} + \delta/2)^m$ be the uniform scalar quantizer of resolution δ defined in Section 6. Quantized measurements take the form $y = Q_\delta(Ax)$. A reconstruction procedure delivering $x^\sharp$ should be quantization consistent, i.e., $Q_\delta(Ax^\sharp) = Q_\delta(Ax)$. This motivates to consider an ℓ_1-minimization problem of the following type for the reconstruction. Let

$$B_\delta = \{z \in \mathbb{R}^m : -\delta/2 \leq z_i \leq \delta/2, i = 1, \ldots, m\}$$

and let $x^\sharp$ be the minimizer of

$$(40) \qquad \min \|z\|_1 \text{ subject to } Az - y \in B_\delta.$$

Then $Q_\delta(Ax^\sharp) = y = Q_\delta(Ax)$. Note that taking the closure of B_δ instead of B_δ the above problem becomes an ℓ_∞-constrained ℓ_1-minimization problem.

The following result concerning reconstruction via (40) based on the standard RIP apparently has not been observed before (but see the discussion in [6] for non-optimal results based on ℓ_p versions of the RIP for $p \neq 2$).

Theorem A.1. *Suppose that $A \in \mathbb{R}^{m \times N}$ is such that $\frac{1}{\sqrt{m}}A$ satisfies the ℓ_2-RIP (3) for $\delta_{2s} < 4/\sqrt{41} \approx 0.62$. Then for any $x \in \mathbb{R}^N$ and $y = Q_\delta(Ax)$ a solution $x^\sharp$ of (40) is quantization consistent and satisfies*

$$(41) \qquad \|x - x^\sharp\|_\infty \lesssim \delta + s^{-1/2} \inf_{w \in \mathbb{R}^N, \|w\|_0 \leq s} \|x - w\|_1.$$

Proof. The optimization problem (40) is closely related to the ℓ_∞-constrained ℓ_1-minimization problem

$$(42) \qquad \min \|z\|_1 \text{ subject to } \|Az - y\|_\infty \leq \delta/2.$$

In fact, either a minimizer of (40) exists in which case it is also a minimizer of (42) or no minimizer of (40) exists in which case the theorem is void. (A minimizer of (42) always exists so that it may be preferred in practice. The error bound (41) still holds for (42), but every minimizer of (42) is quantization inconsistent in the case that no minimizer of (40) exists.)

This close relation of (40) and (42) suggests to study versions of the null space property, see e.g. [9, Chapter 4]. By Theorem [9, Theorem 6.13], the bound on the restricted isometry constants of $\frac{1}{\sqrt{m}}$ implies the ℓ_2-robust null space property in the form

$$\|v_S\|_2 \leq \frac{\rho}{\sqrt{s}}\|v_{S^c}\|_1 + \tau \frac{1}{\sqrt{m}}\|Av\|_2 \quad \text{for all } v \in \mathbb{R}^N \text{ and all } S \subset [N], \#S = s,$$

for constants $\rho \in (0,1)$ and $\tau > 0$ that only depend on δ_{2s}. Since $\|Av\|_2 \leq \sqrt{m}\|Av\|_\infty$ this yields

$$\|v_S\|_2 \leq \frac{\rho}{\sqrt{s}}\|v_{S^c}\|_1 + \tau\|Av\|_\infty \quad \text{for all } v \in \mathbb{R}^N \text{ and all } S \subset [N], \#S = s,$$

which is the ℓ_∞-robust null space property of order s. By [9, Theorem 4.12] this implies the error bound for any minimizer x^* of (42)

$$\|x - x^*\|_2 \lesssim \frac{\inf_{w \in \mathbb{R}^N, \|w\|_0 \leq s} \|x - w\|_1}{\sqrt{s}} + \delta.$$

This concludes the proof. $\qquad\square$

By [16], see also Section 2, a partial random circulant matrix $\frac{1}{\sqrt{|I|}}R_I\Gamma_g$ with sub-sampling on a fixed (deterministic) set $I \subset [N]$ generated by a standard Gaussian random vector satisfies $\delta_{2s} \leq 0.6$ with probability at least $1 - \eta$ if

$$|I| \gtrsim s(\log^2(s)\log^2(N) + \log(1/\eta)).$$

Therefore, Theorem A.1 implies stable reconstruction from quantized measurements $y = Q_\delta(R_I\Gamma_g)$ via (40) under this condition on the number of measurements.

APPENDIX B. UPPER RIP BOUND FOR PARTIAL RANDOM CIRCULANT MATRICES

The proof of our main result only requires the upper (ℓ_2)-RIP bound in (3). For (deterministic) subsampling of random circulant matrices we can deduce the following bound on m from [16], valid also for values of $\kappa \geq 1$.

Theorem B.1. *For a fixed (deterministic) subset $I \subset [N]$, let $A = R_I \Gamma_g$ be a draw of a random circulant matrix generated by a standard Gaussian vector g and $\kappa > 0$. If*

$$|I| \gtrsim \kappa^{-2} s(\log^2(s) \log(N) + \log(1/\eta))$$

then $\sup_{x \in \Sigma_{s,N}} \frac{1}{\sqrt{m}} \|Ax\|_2 \leq 1 + \kappa$ with probability at least $1 - \eta$.

Proof. As argued in [16], we can write $\frac{1}{\sqrt{|I|}} Ax = V_x g$ with $V_x = \frac{1}{\sqrt{|I|}} R_I \Gamma_x$. Denote, $\mathcal{A}_{s,N} = \{V_x : x \in \Sigma_{s,N}\}$. It follows then from [16, Theorem 3.5(a)] that for every $p \geq 1$

$$(43) \quad \left(\mathbb{E} \sup_{x \in \Sigma_{s,N}} \|V_x g\|_2^p \right)^{1/p} \lesssim \gamma_2(\mathcal{A}_{s,N}, \|\cdot\|_{\ell_2 \to \ell_2}) + d_F(\mathcal{A}_{s,N}) + \sqrt{p} d_{\ell_2 \to \ell_2}(\mathcal{A}_{s,N}),$$

where $\gamma_2(\mathcal{A}_{s,N}, \|\cdot\|_{\ell_2 \to \ell_2})$ denotes the γ_2-functional of the set $\mathcal{A}_{s,N}$ with respect to the spectral norm, $d_{\ell_2 \to \ell_2}$ and d_F denote the diameter in the spectral and Frobenius norm, respecticely, of the set in the argument, see [16] for details. These parameters have been estimated in [16, Section 4],

$$d_F(\mathcal{A}_{s,N}) = 1, \qquad d_{\ell_2 \to \ell_2}(\mathcal{A}_{s,N}) \leq \sqrt{s/|I|}$$

$$\gamma_2(\mathcal{A}_{s,N}, \|\cdot\|_{\ell_2 \to \ell_2}) \lesssim \sqrt{s/|I|} \log(s) \log(N).$$

Moreover, the moment bound (43) implies a tail bound (see e.g. [16, Prop. 2.6]), so that

$$\mathbb{P} \left(\sup_{x \in \Sigma_{s,N}} \|Ax\|_2 \geq c(1 + \sqrt{s/|I|} \log(s) \log(N)) + t \right) \leq e^{-c \frac{|I| t^2}{s}}.$$

Requiring that the right hand is bounded by η gives the statement of the theorem. □

APPENDIX C. SOME TOOLS FROM PROBABILITY

The Sudakov inequality provides a bound of the covering numbers in terms of the Gaussian widths, see e.g. [17, Theorem 3.18].

Theorem C.1 (Sudakov). *Let $T \subset \mathbb{R}^N$ and $\delta > 0$. Then the covering numbers with respect to the Euclidean norm obey*

$$\log N(T, \|\cdot\|_2, \delta) \lesssim \delta^{-2} \left(\mathbb{E} \sup_{x \in T} \langle x, g \rangle \right)^2,$$

where g is a standard Gaussian random vector in $\mathbb{R}^N$.

Lemma C.2. *Let X denote a random variable and let $\mathrm{med}(X)$ denote a median of X. Suppose that $\mathbb{E}(X - \mathbb{E}X)^2$ is finite and let X' denote an independent copy of X. Then the following inequalities hold true:*

$$\mathbb{P}(|X - \mathrm{med}(X)| \geq \delta) \leq 2\mathbb{P}(|X - X'| \geq \delta)$$

$$|\mathbb{E}X - \mathrm{med}(X)| \leq (\mathbb{E}(X - \mathbb{E}X)^2)^{\frac{1}{2}}$$

The estimates in Lemma C.2 are well known. We provide a proof for convenience.

Proof. To show the first inequality, observe that since X and X' are identically distributed the following holds for every $\delta > 0$

$$\mathbb{P}(X' \leq \mathrm{med}(X))\mathbb{P}(X \geq \mathrm{med}(X) + \delta) \leq \mathbb{P}(X \geq X' + \delta) .$$

Since $\mathbb{P}(X' \leq \mathrm{med}(X)) \geq \frac{1}{2}$ we conclude that $\mathbb{P}(X \geq \mathrm{med}(X)+\delta) \leq 2\mathbb{P}(X \geq X'+\delta)$ implying the estimate

$$\mathbb{P}(X - \mathrm{med}(X) \geq \delta) \leq 2\mathbb{P}(X - X' \geq \delta) .$$

Repeating the argument with $\mathbb{P}(X - \mathrm{med}(X) \leq -\delta)$ and $\mathbb{P}(X' \geq \mathrm{med}(X)) \geq \frac{1}{2}$ we arrive at the estimate

$$\mathbb{P}(X - \mathrm{med}(X) \leq -\delta) \leq 2\mathbb{P}(X - X' \leq -\delta) .$$

Since the events $\{X - X' \leq -\delta\}$, $\{X - X' \geq \delta\}$ and respectively the events $\{X - \mathrm{med}(X) \leq -\delta\}$, $\{X - \mathrm{med}(X) \geq \delta\}$ are disjoint, it follows that

$$\begin{aligned}
\mathbb{P}(|X - \mathrm{med}(X)| \geq \delta) &= \mathbb{P}(X - \mathrm{med}(X) \leq -\delta) + \mathbb{P}(X - \mathrm{med}(X) \geq \delta) \\
&\leq 2\mathbb{P}(X - X' \leq -\delta) + 2\mathbb{P}(X - X' \geq \delta) \\
&= 2\mathbb{P}(|X - X'| \geq \delta) .
\end{aligned}$$

To show the second inequality observe that with $\sigma = (\mathbb{E}(X - \mathbb{E}X)^2)^{1/2}$ Cantelli's inequality yields

$$\mathbb{P}(X \geq \mathbb{E}X + \sigma) \leq \frac{\mathbb{E}(X - \mathbb{E}X)^2}{\sigma^2 + \sigma^2} = \frac{1}{2} .$$

By using Cantelli's inequality with σ replaced by $-\sigma$ we obtain the estimate

$$\mathbb{P}(X \geq \mathbb{E}X - \sigma) \geq 1 - \frac{\mathbb{E}(X - \mathbb{E}X)^2}{\sigma^2 + \sigma^2} = \frac{1}{2} .$$

Therefore $\mathbb{P}(X < \mathbb{E}X - \sigma) \leq \frac{1}{2}$ and $\mathbb{P}(X \geq \mathbb{E}X + \sigma) \leq \frac{1}{2}$, which shows that every median of X must satisfy

$$\mathbb{E}X - \sigma \leq \mathrm{med}(X) \leq \mathbb{E}X + \sigma .$$

This implies the second inequality and completes the proof. $\qquad\square$

The following probability bound related to symmetrization follows in the same way as in [17, eq. (6.3)].

Lemma C.3. *Let (X_t), $t \in T$, be a family of random variables, indexed by a finite or countable set T, and let (X_t') be an independent copy of (X_t). Then, for $x, y > 0$,*

$$\mathbb{P}(\sup_{t \in T} |X_t| \geq x + y) \leq \mathbb{P}(\sup_{t \in T} |X_t - X_t'| \geq x) + \sup_{t \in T} \mathbb{P}(|X_t| \geq y).$$

References

[1] A. Ai, A. Lapanowski, Y. Plan, and R. Vershynin. One-bit compressed sensing with non-Gaussian measurements. *Linear Algebra Appl.*, 441:222–239, 2014.

[2] R. G. Baraniuk, S. Foucart, D. Needell, Y. Plan, and M. Wootters. Exponential decay of reconstruction error from binary measurements of sparse signals. *IEEE Transactions on Information Theory*, 63(6):3368–3385, 2017.

[3] S. Boucheron, G. Lugosi, and P. Massart. *Concentration inequalities*. Oxford University Press, Oxford, 2013. A nonasymptotic theory of independence, With a foreword by Michel Ledoux.

[4] P. T. Boufounos and R. G. Baraniuk. 1-bit compressive sensing. In *2008 42nd Annual Conference on Information Sciences and Systems*, pages 16–21. IEEE, 2008.

[5] T. Cai and A. Zhang. Sparse Representation of a Polytope and Recovery of Sparse Signals and Low-Rank Matrices,. *IEEE Trans. Inform. Theory*, 60(1):122–132, Jan, 2014.

[6] S. Dirksen, G. Lecué, and H. Rauhut. On the gap between restricted isometry properties and sparse recovery conditions. *IEEE Trans. Inform. Theory*, to appear. ArXiv:1504.05073.

[7] J.-M. Feng, F. Krahmer, and R. Saab. Quantized compressed sensing for partial random circulant matrices. *ArXiv:1702.04711*, 2017.

[8] S. Foucart. *Flavors of Compressive Sensing*, pages 61–104. Springer International Publishing, Cham, 2017.

[9] S. Foucart and H. Rauhut. *A Mathematical Introduction to Compressive Sensing*. Applied and Numerical Harmonic Analysis. Birkhäuser, 2013.

[10] S. Gopi, P. Netrapalli, P. Jain, and A. Nori. One-bit compressed sensing: Provable support and vector recovery. In S. Dasgupta and D. McAllester, editors, *Proceedings of the 30th International Conference on Machine Learning*, volume 28 of *Proceedings of Machine Learning Research*, pages 154–162, Atlanta, Georgia, USA, 17–19 Jun 2013. PMLR.

[11] D. Hanson and F. Wright. A bound on tail probabilities for quadratic forms in independent random variables. *Ann. Math. Statist.*, 42:1079–1083, 1971.

[12] L. Jacques and V. Cambareri. Time for dithering: fast and quantized random embeddings via the restricted isometry property. *ArXiv:1607.00816*, 2016.

[13] 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*, 59(4):2082–2102, 2013.

[14] D. James and H. Rauhut. Nonuniform sparse recovery with random convolutions. In *Sampling Theory and Applications (SampTA), 2015 International Conference on*, pages 34–38. IEEE, 2015.

[15] K. Knudson, R. Saab, and R. Ward. One-bit compressive sensing with norm estimation. *IEEE Trans. Inform. Theory*, 62(5):2748–2758, 2016.

[16] F. Krahmer, S. Mendelson, and H. Rauhut. Suprema of chaos processes and the restricted isometry property. *Comm. Pure Appl. Math.*, 67(11):1877–1904, 2014.

[17] M. Ledoux and M. Talagrand. *Probability in Banach spaces*. Springer-Verlag, Berlin, 1991. Isoperimetry and processes.

[18] S. Mendelson, H. Rauhut, and R. Ward. Improved bounds for sparse recovery from subsampled random convolutions. *ArXiv:1610.04983*, 2016.

[19] A. Moshtaghpour, L. Jacques, V. Cambareri, K. Degraux, and C. D. Vleeschouwer. Consistent basis pursuit for signal and matrix estimates in quantized compressed sensing. *IEEE Signal Processing Letters*, 23(1):25–29, Jan 2016.

[20] Y. Plan and R. Vershynin. One-bit compressed sensing by linear programming. *Comm. Pure Appl. Math.*, 66(8):1275–1297, 2013.

[21] Y. Plan and R. Vershynin. Robust 1-bit compressed sensing and sparse logistic regression: a convex programming approach. *IEEE Trans. Inform. Theory*, 59(1):482–494, 2013.

[22] H. Rauhut. Circulant and Toeplitz matrices in compressed sensing. In *Proc. SPARS'09*, 2009.

[23] H. Rauhut. Compressive sensing and structured random matrices. In M. Fornasier, editor, *Theoretical foundations and numerical methods for sparse recovery*, volume 9 of *Radon Series Comp. Appl. Math.*, pages 1–92. deGruyter, 2010.

[24] J. Romberg. Compressive sensing by random convolution. *SIAM J. Imaging Sci.*, 2(4):1098–1128, 2009.

[25] M. Rudelson and R. Vershynin. Hanson-Wright inequality and sub-gaussian concentration. *Electron. Commun. Probab.*, 18:no. 82, 1–9, 2013.