Information and Inference: A Journal of the IMA (2020) **9**, 601–626
doi:10.1093/imaiai/iaz017
Advance Access publication on 16 October 2019

One-bit compressed sensing with partial Gaussian circulant matrices

Sjoerd Dirksen†
Mathematical Institute, Utrecht University, P.O. Box 80010, 3508 TA Utrecht, The Netherlands
†Corresponding author. Email: s.dirksen@uu.nl

Hans Christian Jung and Holger Rauhut
Lehrstuhl C für Mathematik (Analysis), RWTH Aachen University,
Pontdriesch 10, 52062 Aachen, Germany
jung@mathc.rwth-aachen.de jung@mathc.rwth-aachen.de

[Received on 16 March 2018; revised on 15 May 2019; accepted on 13 June 2019]

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, $m \simeq \delta^{-4} s \log(N/s\delta)$ 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 restricted isometry property property. Under a slightly worse dependence on δ, we establish stability with respect to approximate sparsity, as well as full vector recovery results, i.e., estimation of both vector norm and direction.

Keywords: compressed sensing; quantization; circulant matrices; restricted isometry properties.

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 the empirical fact 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 *nonlinear* measurements of the form $y = Q(Ax)$, where $Q : \mathbb{R}^m \to \mathscr{A}^m$ is a quantizer and $\mathscr{A}$ denotes a finite quantization alphabet.

In this paper, we study the measurement model

$$y = \operatorname{sign}(Ax + \tau), \tag{1.1}$$

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 [5]. 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. The use of random thresholds is referred to as *dithering* and has originally been introduced for the purpose of removing visual artefacts in quantization of images [26]; see also [12, Section V.E) for an overview. More recently, dithering has re-appeared in the compressed sensing literature; see, e.g., [2,4,17]. In particular, the recent works [2,17] have shown that dithering allows one to completely reconstruct the signal (instead of only its direction) from one-bit measurements of the form (1.1) under certain circumstances.

Until now, recovery results for the one-bit compressed sensing model (1.1) 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 synthetic aperture radar imaging, Fourier optical imaging and channel estimation (see, e.g., [27] 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 $0 < \delta \leq 1$, we assume that the sparsity s is small enough, i.e.,

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

In addition, we suppose that the (expected) number of measurements satisies

$$m \gtrsim \begin{cases} \delta^{-1} s \log(eN/(s\delta)) & \text{if } 0 < \delta \leq \left(\log^2(s)\log(N)\right)^{-1} \\ \delta^{-1/2} s \log(s) \log^{3/2}(N) & \text{if } \left(\log^2(s)\log(N)\right)^{-1} < \delta \leq 1. \end{cases} \tag{1.2}$$

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

$$x^{\#}_{\mathrm{HT}} = H_s(A^* \operatorname{sign}(Ax)) \tag{HT}$$

and via the program

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

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

Our first result shows that under (1.2) the following holds with high probability: for any s-sparse x with $\|x\|_2 = 1$ the hard thresholding reconstruction $x^{\#}_{\mathrm{HT}}$ satisfies $\|x - x^{\#}_{\mathrm{HT}}\|_2 \leq \delta^{1/4}$. Moreover, under slightly stronger conditions (see Theorem 4.2 with '$\delta = \delta^{1/4}$') for any vector satisfying $\|x\|_1 \leq \sqrt{s}$ and $\|x\|_2 = 1$, any solution $x^{\#}_{\mathrm{LP}}$ to (LP) satisfies $\|x - x^{\#}_{\mathrm{LP}}\|_2 \leq \delta^{1/8}$. 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, $\mathcal{N}(0, R^2)$-distributed random variables. If a condition similar to (1.2) 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 $\|x\|_2 \leq R$, any solution

$x^{\#}_{\mathrm{CP}}$ to the second-order cone program

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

satisfies $\|x - x^{\#}_{\mathrm{CP}}\|_2 \leqslant R\delta^{1/8}$.

Our analysis relies on an observation of Foucart [9], who showed that it is sufficient for the matrix A to satisfy an ℓ_1/ℓ_2 *restricted isometry 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 5.1, which is the main technical result of our work. The final section of the paper discusses two additional consequences of these RIP results. In Corollary 6.1 we follow the work [7] to derive a recovery guarantee for (unquantized) outlier robust compressed sensing with Gaussian circulant matrices. In Theorem 6.1 we use a recent result from [14] 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 (ℓ_2-)restricted isometry property is a classical way of analyzing the performance of compressive sensing [10]. The restricted isometry constant δ_s is defined as the smallest constant δ such that

$$(1-\delta)\|x\|_2^2 \leqslant \|Ax\|_2^2 \leqslant (1+\delta)\|x\|_2^2 \quad \text{for all } \ s\text{-sparse } \ x \in \mathbb{R}^N. \tag{2.1}$$

If $\delta_{2s} < 1/\sqrt{2}$ then all s-sparse signals can be reconstructed via ℓ_1-minimization exactly; see, e.g., [6,10]. Stability under noise and sparsity defects can be shown as well and similar guarantees also hold for other reconstruction algorithms [10]. It is well known that Gaussian random matrices satisfy $\delta_s \leqslant \delta$ with probability at least $1-\eta$ if $m \gtrsim \delta^{-2}(s\log(eN/s) + \log(1/\eta))$ [10, Chapter 9].

The situation that A is a subsampled random circulant matrix (see below for a formal definition) has been analysed in several contributions [16,18,20,24,25,27]. The best available result states [18] that a properly normalized (deterministically) subsampled random circulant matrix (generated by a (sub)gaussian random vector) satisfies $\delta_s \leqslant \delta$ with probability at least $1-\eta$ if

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

The original contribution [27] 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

[10] in [20] in a small sparsity regime $s \lesssim \sqrt{N/\log(N)}$ under the optimal condition

$$m \gtrsim s\log(eN/s). \tag{2.2}$$

Non-uniform recovery results have been shown in [16,24,25], 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, $\tau = \mathbf{0}$. 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 [15, 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 $\|x - x'\|_2 \leqslant \delta$. 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 $\|x^\# - x\|_2 \gtrsim s/(m + s^{3/2})$ [15, Theorem 1]. That is, the dependence of m on δ can in general not be improved. It was shown in [11, Theorem 7] that this optimal error dependence can be obtained using a polynomial time algorithm if the measurement matrix is modified. Specifically, the work [11] showed that if $m \gtrsim \delta^{-1}m'\log(m'/\delta)$ and $A = A_2A_1$, 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 $m' \simeq s\log(N/s)$ 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 [22] proposed the reconstruction program (LP). They showed that using $m \gtrsim \delta^{-1}s\log^2(N/s)$ Gaussian measurements one can recover every x with $\|x\|_1 \leqslant \sqrt{s}$ and $\|x\|_2 = 1$ via (LP) with reconstruction error $\delta^{1/5}$. In [23] 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 [17] 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 [17] 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 of the signal.

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 [23] 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 $Q_\alpha : \mathbb{R}^m \to (\alpha\mathbb{Z} + \alpha/2)^m$ defined by $Q_\alpha(z) = \left(\alpha\lfloor z_i/\alpha\rfloor + \alpha/2\right)_{i=1}^m$. 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 [18] that with high probability one can recover any s-sparse vector up to a reconstruction error α

from its quantized measurements $Q_\alpha(Ax)$. It follows from [20] that by using random subsampling and imposing a small sparsity assumption similar to ours, this number of measurements can be decreased to $m \gtrsim s\log(N/s)$. In these results, the recovery error does not improve beyond the resolution α of the quantizer, even if one takes more measurements. In Theorem 6.1 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 [8]. In particular, [8, Theorem 5] proves that based on $m \simeq s\log^2 s\log^2 N$ one-bit sigma-delta quantized measurements, one can use a convex program to find an approximand 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. We let A_i, $i = 1, \ldots, m$, denote the rows of a matrix $A \in \mathbb{R}^{m\times N}$. 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 the 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 $R_I \in \{0,1\}^{|I|\times N}$ denote the restriction matrix such that $R_I(x) = x_I$. Further, $\|x\|_p$ denotes the usual ℓ_p-norm of a vector $x \in \mathbb{R}^N$ and $B_{\ell_p^N}$ denotes the corresponding unit ball; $\|A\|_{\ell_2\to\ell_2}$ denotes the spectral norm of a matrix A and $\|A\|_F$ its Frobenius norm. For an event E, 1_E is the indicator function of E.

Throughout the text, we write $A \lesssim B$ (respectively $A \gtrsim B$) if there is an absolute constant $c > 0$ such that the inequality $A \leqslant cB$ (respectively $A \geqslant cB$) holds. Further, we write $A \simeq B$ if both $A \lesssim B$ and $A \gtrsim B$ hold simultaneously.

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 $\|x\|_1 \leqslant \sqrt{s}\|x\|_2$. 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 the diagonal matrix, respectively, generated by x. That is,

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

We study the following linear measurement matrix. We consider a vector $\theta = (\theta_i)_{i\in[N]}$ of i.i.d. random selectors, i.e., a sequence of independent Bernoulli random variables θ_i satisfying $1 - \mathbb{P}(\theta_i = 0) = \mathbb{P}(\theta_i = 1) = m/N$, 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 $m = \mathbb{E}|I|$, i.e., m corresponds to the expected number of measurements in this model. However, the Chernoff bound implies that $m/2 \leqslant |I| \leqslant 3m/2$ with probability at least $1-\mathrm{e}^{-cm}$, so the true number of measurements $|I|$ is with high probability comparable to m.

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 $0 < \delta, \eta \leqslant 1$ and $s \in [N]$ such that

$$s \lesssim \min\left\{\sqrt{\delta^2 N/\log(N)}, \delta^2 N/\log(1/\eta)\right\}. \tag{4.1}$$

Set $\delta_0 := \delta_0(s, N) := (\log^2(s)\log(N))^{-1/2}$. If $0 < \delta \leqslant \delta_0$ suppose that

$$m \gtrsim \delta^{-2} s \log\left(eN/(s\delta\eta)\right). \tag{4.2}$$

If $\delta_0 < \delta \leqslant 1$ suppose that

$$m \gtrsim \delta^{-1} s \max\left\{\frac{\log(N)}{\delta_0}, \delta_0 \log(1/\eta), \frac{\log(1/\eta)}{s\delta_0}\right\}. \tag{4.3}$$

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

Let us remark that for polynomially scaling probabilities $\eta = N^{-\alpha}$, the second and third term in the maximum in (4.3) can be bounded by a constant c_α times the first term and then (4.3) 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 (4.1) imposes an implicit restriction on δ. In particular, if $\delta \lesssim \sqrt{\log(N)/N}$, then (4.1) excludes all non-trivial sparsities $s \geqslant 1$. However, we expect that the requirement (4.1) of small sparsity is only an artefact of our proof and that recovery can also be expected for larger sparsities under conditions similar to (4.2) and (4.3) with possibly more logarithmic factors; see also [**20**] for an analogous phenomenon in standard compressed sensing. In fact, our proof relies on the $\mathrm{RIP}_{1,2}$ property (see (4.2) below) and [**20**] provides at least the lower $\mathrm{RIP}_{1,2}$ bound also for larger sparsities.

Under a slightly worse scaling in δ than in (4.2) and (4.3), we can recover any effectively sparse signal on the unit sphere via the linear program (LP).

THEOREM 4.2 Let $0 < \delta, \eta \leqslant 1$ and introduce $\delta_0 = (\log^2(s)\log(N))^{-1/2}$. If $0 < \delta \leqslant \delta_0^{1/2}$ assume that

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

and if $\delta_0^{1/2} < \delta \leqslant 1$ assume that

$$s \lesssim \min\left\{\delta^{4/3}\sqrt{N/\log^2(N)}, \delta^2 N/\log(1/\eta)\right\},$$
$$m \gtrsim \delta^{-4/3} s \max\left\{\frac{\log(N)}{\delta_0^{4/3}}, \delta_0^{2/3}\log(1/\eta), \frac{\log(1/\eta)}{\delta_0^{2/3}\delta^{4/3}s}\right\}.$$

Then the following holds with probability exceeding $1-\eta$: for every $x \in \mathbb{R}^N$ with $\|x\|_1 \leqslant s$ and $\|x\|_2 = 1$, any solution $x^\#_{\mathrm{LP}}$ to (LP) satisfies $\|x - x^\#_{\mathrm{LP}}\|_2 \lesssim \sqrt{\delta}$.

We will prove Theorem 4.1 by using a recent observation of Foucart [9]. 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 $A \in \mathbb{R}^{m\times N}$ *satisfies* $\mathrm{RIP}_{1,2}(s,\delta)$ if

$$(1-\delta)\|x\|_2 \leqslant \|Ax\|_1 \leqslant (1+\delta)\|x\|_2, \qquad \text{for all } x \in \Sigma_{s,N}, \tag{4.4}$$

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

$$(1-\delta)\|x\|_2 \leqslant \|Ax\|_1 \leqslant (1+\delta)\|x\|_2, \qquad \text{for all } x \in \Sigma^{\mathrm{eff}}_{s,N}. \tag{4.5}$$

LEMMA 4.1 [9, Theorem 8] Suppose that A satisfies $\mathrm{RIP}_{1,2}(2s,\delta)$. Then, for every $x \in \mathbb{R}^N$ with $\|x\|_0 \leqslant s$ and $\|x\|_2 = 1$, the hard thresholding reconstruction $x^\#_{\mathrm{HT}}$ satisfies $\|x - x^\#_{\mathrm{HT}}\|_2 \leqslant 2\sqrt{5\delta}$.

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

REMARK 4.1 It is in general not possible to extend Theorem 4.1 to subgaussian circulant matrices. Indeed, consider any random measurement matrix $A \in \mathbb{R}^{m\times N}$ with entries in $\{-1,1\}$, e.g., a randomly subsampled circulant matrix generated by a Rademacher random vector. Suppose that the threshold vector τ in (1.1) is zero and consider, for $0 < \lambda < 1$, the normalized 2-sparse vectors

$$x_{+\lambda} = (1+\lambda^2)^{-1/2}(1,\lambda,0,\ldots,0), \qquad x_{-\lambda} = (1+\lambda^2)^{-1/2}(1,-\lambda,0,\ldots,0). \tag{4.6}$$

Then $x_{+\lambda}$ and $x_{-\lambda}$ produce identical one-bit measurements, i.e., $\mathrm{sign}(Ax_\lambda) = \mathrm{sign}(Ax_{-\lambda})$.

As a consequence, Lemma 4.1 implies that A cannot satisfy $\mathrm{RIP}_{1,2}(4,\delta)$ for small values of δ. Indeed, if A satisfies this property then Lemma 4.1 implies

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

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

By excluding extremely sparse vectors via a suitable ℓ_∞-norm bound, it might be possible to circumvent this counterexample. In fact, in the case of unstructured subgaussian random matrices, positive recovery results for sparse vectors with such an additional constraint were shown in [**1**].

So far, our recovery results only allow one 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 4.3 Let $A=R_I\Gamma_g$ and let $\tau_1,\dots,\tau_m$ be independent $\mathcal{N}(0,R^2)$-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 $\|x\|_1\leqslant\sqrt{s}\|x\|_2$ and $\|x\|_2\leqslant R$, any solution $x^{\#}_{\mathrm{CP}}$ to the second-order cone program (CP) satisfies $\|x-x^{\#}_{\mathrm{CP}}\|_2\leqslant R\sqrt{\delta}$.

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

$$\min_{z\in\mathbb{R}^N}\|z\|_1 \qquad \text{s.t.} \qquad \operatorname{sign}(C[z,R])=\operatorname{sign}(C[x,R]),\ \|z\|_2\leqslant R. \tag{4.7}$$

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

$$B:=R_I[\Gamma_g\ h]=R_I\begin{bmatrix} g_N & g_1 & g_2 & \cdots & g_{N-2} & g_{N-1} & h_1\\ g_{N-1} & g_N & g_1 & \cdots & g_{N-3} & g_{N-2} & h_2\\ g_{N-2} & g_{N-1} & g_N & \cdots & g_{N-4} & g_{N-3} & h_3\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ g_1 & g_2 & g_3 & \cdots & g_{N-1} & g_N & h_N\end{bmatrix}, \tag{4.8}$$

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

LEMMA 4.2 Let $\delta<1/5$. Suppose that C satisfies $\mathrm{RIP}^{\mathrm{eff}}_{1,2}(36(\sqrt{s}+1)^2,\delta)$. Then, for any $x\in\mathbb{R}^N$ satisfying $\|x\|_1\leqslant\sqrt{s}\|x\|_2$ and $\|x\|_2\leqslant R$, any solution $x^{\#}$ to (4.7] satisfies

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

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

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

$$\|u-v\|_2 \leqslant 2\left\| \frac{[u,1]}{\|[u,1]\|_2} - \frac{[v,1]}{\|[v,1]\|_2} \right\|_2 \tag{4.9}$$

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 (4.7] and write

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

Since $x/R, x^\#/R \in B_{\ell_2^N}$, (4.9) implies that

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

By the parallellogram identity,

$$\left\| \frac{\bar{x}-\bar{x}^\#}{2} \right\|_2^2 = \frac{\|\bar{x}\|_2^2 + \|\bar{x}^\#\|_2^2}{2} - \left\| \frac{\bar{x}+\bar{x}^\#}{2} \right\|_2^2. \tag{4.10}$$

Let us observe that $[x,R]$ and $[x^\#,R]$ are $(\sqrt{s}+1)^2$-effectively sparse. Indeed, by optimality of $x^\#$ for (4.7) and s-effective sparsity of x,

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

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

$$z := \frac{\bar{x}+\bar{x}^\#}{2} \in \Sigma^{\mathrm{eff}}_{36(\sqrt{s}+1)^2, N+1}. \tag{4.11}$$

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

$$\left\| \frac{\bar{x}+\bar{x}^\#}{2} \right\|_2 \geqslant \frac{1}{1+\delta} \left\| C\Big(\frac{\bar{x}+\bar{x}^\#}{2}\Big) \right\|_1 = \frac{\|C\bar{x}\|_1 + \|C\bar{x}^\#\|_1}{2(1+\delta)} \geqslant \frac{(1-\delta)}{(1+\delta)}. \tag{4.12}$$

Hence, (4.10) implies

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

Let us now prove (4.11). Since $[x,R]$ and $[x^\#,R]$ are $(\sqrt{s}+1)^2$-effectively sparse,

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

It remains to bound $\|z\|_2$ from below. In (4.12) we already observed that

$$\|Cz\|_1 = \frac{1}{2}\left\|\frac{C[x,R]}{\|[x,R]\|_2}\right\|_1 + \frac{1}{2}\left\|\frac{C[x^\#,R]}{\|[x^\#,R]\|_2}\right\|_1 \geqslant (1-\delta). \tag{4.13}$$

Set $t = 8s+8$, then $t \geqslant 4(\sqrt{s}+1)^2 \geqslant t/2$. 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 $k \geqslant 1$,

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

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

$$\begin{aligned}
\|Cz\|_1 &\leqslant \sum_{k\geqslant 0}\|Cz_{T_k}\|_1 \leqslant (1+\delta)\left(\|z_{T_0}\|_2 + \sum_{k\geqslant 1}\|z_{T_k}\|_2\right) \\
&\leqslant (1+\delta)\|z\|_2 + \frac{(1+\delta)}{\sqrt{t}}\|z\|_1 \leqslant (1+\delta)\|z\|_2 + \frac{(1+\delta)}{\sqrt{t}}(\sqrt{s}+1). \\
&\leqslant (1+\delta)\|z\|_2 + \frac{1}{2}(1+\delta).
\end{aligned} \tag{4.14}$$

Since $\delta \leqslant 1/5$, (4.13] and (4.14] together yield

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

□

5. Proof of the $\mathrm{RIP}_{1,2}$ properties

We will now prove the RIP properties required in Lemmas 4.1 and 4.2. For our analysis we recall two standard concentration inequalities. The Hanson–Wright inequality [**13**] states that if g is a standard Gaussian vector and $B \in \mathbb{R}^{N\times N}$, then for all $t \geqslant 0$,

$$\mathbb{P}(|g^TBg - \mathbb{E}g^TBg| \geqslant t) \leqslant \exp\left(-c\min\left\{\frac{t^2}{\|B\|_F^2}, \frac{t}{\|B\|_{\ell_2\to\ell_2}}\right\}\right), \tag{5.1}$$

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 [**28**] 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 for all $t \geqslant 0$,

$$\mathbb{P}\left(\left|\sup_{x\in T}\langle x,g\rangle - \mathbb{E}\sup_{x\in T}\langle x,g\rangle\right| \geqslant t\right) \leqslant 2\mathrm{e}^{-t^2/2\sigma^2}, \tag{5.2}$$

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 $\|y\|_1 \leqslant \sqrt{s}$ and $\|y\|_2 = 1$. Let g be a standard Gaussian vector. For any $t > 0$,

$$\mathbb{P}\left(\left|\frac{1}{N}\|\Gamma_g y\|_2^2 - 1\right| \geqslant t\right) \leqslant 2\mathrm{e}^{-cN\min\{t^2,t\}/s} \tag{5.3}$$

and

$$\mathbb{P}\left(\left|\frac{1}{N}\sqrt{\frac{\pi}{2}}\|\Gamma_g y\|_1 - 1\right| \geqslant t\right) \leqslant 2\mathrm{e}^{-Nt^2/\pi s}. \tag{5.4}$$

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

$$\mathbb{P}\left(\left|\frac{1}{N}\sqrt{\frac{\pi}{2}}\|[\Gamma_g\ h]y\|_1 - 1\right| \geqslant t\right) \leqslant 2\mathrm{e}^{-Nt^2/\pi s}. \tag{5.5}$$

Proof. Note that $\Gamma_g y = \Gamma_y g$. By the Hanson–Wright inequality (5.1),

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

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

$$\|\Gamma_y^*\Gamma_y\|_{\ell_2\to\ell_2} = \|\Gamma_y\|_{\ell_2\to\ell_2}^2 \leqslant s \tag{5.6}$$

and

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

To prove (5.4) 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 (5.2) for suprema of Gaussian processes applied with $T = \Gamma_y^* B_{\ell_\infty^N}$,

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

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 (5.6),

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

The proof of (5.5) 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

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

Moreover,

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

The concentration inequality (5.2) for suprema of Gaussian processes now implies

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

where

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

For any $z \in B_{\ell_\infty^N}$, we obtain using (5.6) and the Cauchy–Schwarz inequality,

$$\begin{aligned}\|[\Gamma_{y_{[N]}} \; y_{N+1}\mathrm{Id}_N]^* z\|_2^2 &= \|\Gamma_{y_{[N]}}^* z\|_2^2 + y_{N+1}^2 \;\|z\|_2^2 \\ &\leqslant \left(\|\Gamma_{y_{[N]}}\|_{\ell_2 \to \ell_2}^2 + y_{N+1}^2\right)\|z\|_2^2 \leqslant \left(s\|y_{[N]}\|_2^2 + y_{N+1}^2\right)N \leqslant sN.\end{aligned}$$

This completes the proof. □

By Lemmas 4.1 and 4.2, our main recovery results in Theorems 4.1 and 4.3 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 5.1 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 $\mathrm{RIP}_{1,2}(s, \delta)$ 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 $\mathrm{RIP}_{1,2}^{\mathrm{eff}}(s, \delta)$ with probability at least $1 - \eta$.

Proof. Let $\kappa > 0$ be a number to be chosen later. Let $\mathscr{N}_{\delta/(1+\kappa)} \subset \Sigma_{s,N}$ be a minimal $\delta/(1+\kappa)$-net for $\Sigma_{s,N}$ with respect to the Euclidean norm. It is well known, see, e.g., [10, Proposition C.3], that

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

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

$$\begin{aligned} E_I &= \left\{\frac{m}{2} \leqslant |I| \leqslant \frac{3m}{2}\right\} \\ E_{\mathrm{RIP}} &= \left\{\forall\, z \in \Sigma_{2s,N}\ \ \frac{1}{\sqrt{m}}\|Az\|_2 \leqslant C(1+\kappa)\right\} \\ E_{\Gamma,\ell_1} &= \left\{\forall y \in \mathscr{N}_{\delta/(1+\kappa)}\ \ \left|\frac{1}{N}\sqrt{\frac{\pi}{2}}\|\Gamma_g y\|_1 - 1\right| \leqslant \delta\right\} \\ E &= \left\{\forall y \in \mathscr{N}_{\delta/(1+\kappa)}\ \ \left|\frac{1}{m}\sqrt{\frac{\pi}{2}}\|Ay\|_1 - \frac{1}{N}\sqrt{\frac{\pi}{2}}\|\Gamma_g y\|_1\right| \leqslant 2\delta\right\}, \end{aligned} \tag{5.7}$$

respectively. Under E_I and E_{RIP},

$$\begin{aligned} \left|\frac{1}{m}\sqrt{\frac{\pi}{2}}\|Ax\|_1 - \frac{1}{m}\sqrt{\frac{\pi}{2}}\|Ay\|_1\right| &\leqslant \frac{\delta}{1+\kappa}\frac{1}{m}\sqrt{\frac{\pi}{2}}\left\|A\left(\frac{x-y}{\|x-y\|_2}\right)\right\|_1 \\ &\leqslant \frac{\delta}{1+\kappa}\frac{|I|}{m}\sup_{z\in\Sigma_{2s,N}}\frac{1}{\sqrt{|I|}}\sqrt{\frac{\pi}{2}}\|Az\|_2 \leqslant C\sqrt{\frac{9\pi}{8}}\delta. \end{aligned}$$

Therefore, if the events in (5.7) hold simultaneously, then by the triangle inequality

$$\begin{aligned} \left|\frac{1}{m}\sqrt{\frac{\pi}{2}}\|Ax\|_1 - 1\right| &\leqslant \left|\frac{1}{m}\sqrt{\frac{\pi}{2}}\|Ax\|_1 - \frac{1}{m}\sqrt{\frac{\pi}{2}}\|Ay\|_1\right| + \left|\frac{1}{m}\sqrt{\frac{\pi}{2}}\|Ay\|_1 - \frac{1}{N}\sqrt{\frac{\pi}{2}}\|\Gamma_g y\|_1\right| \\ &\quad + \left|\frac{1}{N}\sqrt{\frac{\pi}{2}}\|\Gamma_g y\|_1 - 1\right| \leqslant \big(C\sqrt{9\pi/8}+3\big)\delta. \end{aligned}$$

Hence, it remains to show that the events in (5.7) hold with probability at least $1-\eta$. The Chernoff bound immediately yields $\mathbb{P}(E_I^c) \leqslant \mathrm{e}^{-cm}$. By Theorem B.1, under the event E_I, if

$$m \gtrsim \kappa^{-2} s\big(\log^2(s)\log^2(N) + \log(1/\eta)\big) \tag{5.8}$$

then by Lemma C.3(a),

$$\begin{aligned}\mathbb{P}_{\theta,g}(E^c_{\mathrm{RIP}}) &= \mathbb{E}_\theta \mathbb{P}_g\left(\exists z \in \Sigma_{2s,N} \ \ \frac{1}{\sqrt{|I|}}\|R_I \Gamma_g z\|_2 \geqslant C(1+\kappa)\right)\\ &\leqslant \mathbb{E}_\theta \mathbb{P}_g\left(\exists z \in \Sigma_{2s,N} \ \ \frac{1}{\sqrt{|I|}}\|R_I \Gamma_g z\|_2 1_{E_I} \geqslant C(1+\kappa)\right) + \mathbb{P}\big(E_I^c\big)\\ &\leqslant 2\eta.\end{aligned}$$

Moreover, by (5.4) and a union bound,

$$\begin{aligned}\mathbb{P}\left(\sup_{y\in\mathcal{N}_{\delta/(1+\kappa)}}\left|\frac{1}{N}\sqrt{\frac{\pi}{2}}\|\Gamma_g y\|_1 - 1\right| \geqslant \delta\right) &\leqslant |\mathcal{N}_{\delta/(1+\kappa)}| 2\mathrm{e}^{-\delta^2 N/\pi s}\\ &\leqslant 2\mathrm{e}^{s\log(3e(1+\kappa)N/(s\delta)) - \delta^2 N/\pi s},\end{aligned}$$

so $\mathbb{P}(E^c_{\Gamma,\ell_1}) \leqslant 2\eta$ if

$$N \geqslant \pi\delta^{-2}s^2 \log\big(3eN(1+\kappa)/(s\delta)\big) + \pi\delta^{-2}s\log(1/\eta). \tag{5.9}$$

Thus, it remains to show that $\mathbb{P}(E^c) \leqslant c\eta$ for an absolute constant $c > 0$. 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.2,

$$\mathbb{P}_\theta\left(\sup_{y\in\mathcal{N}_{\delta/(1+\kappa)}} |X_y| \geqslant 2\delta\right) \leqslant \mathbb{P}_\theta\left(\sup_{y\in\mathcal{N}_{\delta/(1+\kappa)}} |X_y - X'_y| \geqslant \delta\right) + \sup_{y\in\mathcal{N}_{\delta/(1+\kappa)}} \mathbb{P}_\theta\big(|X_y| \geqslant \delta\big)$$

and so, by the law of total probability,

$$\begin{aligned}\mathbb{P}_{\theta,g}\Big(\sup_{y\in\mathcal{N}_{\delta/(1+\kappa)}} |X_y| \geqslant 2\delta\Big) \leqslant \mathbb{P}_{\theta,g}\Big(&\sup_{y\in\mathcal{N}_{\delta/(1+\kappa)}} |X_y - X'_y| \geqslant \delta\Big)\\ &+ \mathbb{E}_g \sup_{y\in\mathcal{N}_{\delta/(1+\kappa)}} \mathbb{P}_\theta(|X_y| \geqslant \delta).\end{aligned} \tag{5.10}$$

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

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

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

$$\begin{aligned}
&\mathbb{P}_{\theta,\theta',g}\left(\sup_{y\in\mathcal{N}_{\delta/(1+\kappa)}} |X_y - X'_y| \geqslant \delta\right)\\
&\quad= \mathbb{P}_{\varepsilon,\theta,\theta',g}\left(\frac{1}{m}\sqrt{\frac{\pi}{2}}\sup_{y\in\mathcal{N}_{\delta/(1+\kappa)}}\left|\sum_{i=1}^{N}\varepsilon_i(\theta_i-\theta'_i)|\langle(\Gamma_g)_i, y\rangle|\right| \geqslant \delta\right)\\
&\quad\leqslant 2\mathbb{P}_{\varepsilon,\theta,g}\left(\frac{1}{m}\sqrt{\frac{\pi}{2}}\sup_{y\in\mathcal{N}_{\delta/(1+\kappa)}}\left|\sum_{i=1}^{N}\varepsilon_i\theta_i\left|\langle(\Gamma_g)_i, y\rangle\right|\right| \geqslant \delta/2\right)\\
&\quad\leqslant 2\mathbb{P}_{\varepsilon,\theta,g}\left(\frac{1}{m}\sqrt{\frac{\pi}{2}}\sup_{y\in\mathcal{N}_{\delta/(1+\kappa)}}\left|\sum_{i=1}^{N}\varepsilon_i\theta_i 1_{E_{\mathrm{RIP}}}|\langle(\Gamma_g)_i, y\rangle|\right| \geqslant \delta/2\right) + \mathbb{P}_{\theta,g}(E^c_{\mathrm{RIP}}),
\end{aligned}$$

where we used Lemma C.3(a) in the last step. By Hoeffding's inequality applied to $(\varepsilon_i)_{i\in[N]}$ and assuming E_I,

$$\begin{aligned}
\mathbb{P}_{\varepsilon}\left(\frac{1}{m}\sqrt{\frac{\pi}{2}}\left|\sum_{i=1}^{N}\varepsilon_i\theta_i 1_{E_{\mathrm{RIP}}}|\langle(\Gamma_g)_i, y\rangle|\right| \geqslant \delta/2\right) &\leqslant 2\exp\left(-\frac{m^2\delta^2}{2\pi\sum_{i=1}^{N}\theta_i 1_{E_{\mathrm{RIP}}}|\langle(\Gamma_g)_i, y\rangle|^2}\right)\\
&= 2\exp\left(-\frac{m\delta^2}{2\pi 1_{E_{\mathrm{RIP}}}\frac{1}{m}\|Ay\|_2^2}\right)\\
&\leqslant 2\mathrm{e}^{-\frac{m\delta^2}{2\pi C^2(1+\kappa)^2}}.
\end{aligned}$$

Hence, a union bound yields

$$\mathbb{P}_{\varepsilon,\theta,g}\left(\frac{1}{m}\sqrt{\frac{\pi}{2}}\sup_{y\in\mathcal{N}_{\delta/(1+\kappa)}}\left|\sum_{i=1}^{N}\varepsilon_i\theta_i 1_{E_{\mathrm{RIP}}}|\langle(\Gamma_g)_i, y\rangle|\right| \geqslant \delta/2\right) \leqslant 2|\mathcal{N}_{\delta/(1+\kappa)}|\,\mathrm{e}^{-\frac{m\delta^2}{2\pi C^2(1+\kappa)^2}} \leqslant \eta, \tag{5.11}$$

provided that

$$m \gtrsim \frac{(1+\kappa)^2}{\delta^2}\left(s\log\left(\frac{3e(1+\kappa)N}{\delta s}\right) + \log(1/\eta)\right). \tag{5.12}$$

To bound the second term on the right-hand side of (5.10), consider the event

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

By (5.3) and a union bound,

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

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 (up to a constant) the same as (5.9). This shows that

$$\mathbb{E}_g \sup_{y\in\mathcal{N}_{\delta/(1+\kappa)}} \mathbb{P}_\theta(|X_y| \geqslant \delta) \leqslant \mathbb{E}_g \sup_{y\in\mathcal{N}_{\delta/(1+\kappa)}} \mathbb{P}_\theta\left(|X_y 1_{E_{\Gamma,\ell_2}}| \geqslant \delta\right) + \eta. \tag{5.13}$$

Now recall the following facts. If X is a random variable, X' is an independent copy, and med(X) is a median of X, then for any $\delta > 0$ (see Lemma C.1],

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

and

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

Combining these, we find

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

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., [**19**, Lemma 6.3] or [**10**, Lemma 8.4],

$$\begin{aligned}
\left(\mathbb{E}_\theta((X - \mathbb{E}_\theta X)^2)^{1/2}\right) &= \left(\mathbb{E}_\theta \left|\frac{1}{m}\sqrt{\frac{\pi}{2}}\|Ay\|_1 - \mathbb{E}_\theta\left(\frac{1}{m}\sqrt{\frac{\pi}{2}}\|Ay\|_1\right)\right|^2\right)^{1/2} 1_{E_{\Gamma,\ell_2}} \\
&\leqslant \frac{2}{m}\sqrt{\frac{\pi}{2}}\left(\mathbb{E}_{\theta,\varepsilon}\left|\sum_{i=1}^N \varepsilon_i\theta_i|\langle(\Gamma_g)_i, y\rangle|\right|^2\right)^{1/2} 1_{E_{\Gamma,\ell_2}} \\
&= \frac{2}{m}\sqrt{\frac{\pi}{2}}\left(\mathbb{E}_\theta \sum_{i=1}^N \theta_i|\langle(\Gamma_g)_i, y\rangle|^2\right)^{1/2} 1_{E_{\Gamma,\ell_2}} = \frac{1}{\sqrt{m}}\sqrt{2\pi}\frac{1}{\sqrt{N}}\|\Gamma_g y\|_2 1_{E_{\Gamma,\ell_2}} \\
&\leqslant \frac{2\sqrt{2\pi}}{\sqrt{m}} \leqslant \delta/2
\end{aligned}$$

if $m \geqslant 32\pi\delta^{-2}$, the latter being a weaker condition than (5.12]. In summary, we find

$$\begin{aligned}
\mathbb{P}_\theta(|X_y 1_{E_{\Gamma,\ell_2}}| \geqslant \delta) &\leqslant 2\mathbb{P}(|X_y - X'_y|1_{E_{\Gamma,\ell_2}} \geqslant \delta/2) \\
&\leqslant 4\mathbb{P}_{\theta,\varepsilon}\left(\frac{1}{m}\sqrt{\frac{\pi}{2}}\left|\sum_{i=1}^N \varepsilon_i\theta_i|\langle(\Gamma_g)_i, y\rangle|1_{E_{\Gamma,\ell_2}}\right| \geqslant \frac{\delta}{4}\right).
\end{aligned}$$

Now apply Hoeffding's inequality with respect to $(\varepsilon_i)_{i\in[N]}$ to obtain

$$\mathbb{P}_\theta(|X_y 1_{E_{\Gamma,\ell_2}}| \geqslant \delta) \leqslant 8\mathbb{E}_\theta \exp\left(\frac{-m^2\frac{2}{\pi}\delta^2}{16\sum_{i=1}^N \theta_i |\langle(\Gamma_g)_i, y\rangle|^2}\right) = 8\mathbb{E}_\theta \exp\left(\frac{-m\delta^2}{8\pi\frac{1}{m}\|Ay\|_2^2}\right)$$
$$\leqslant 8\mathrm{e}^{-\frac{m\delta^2}{8\pi C^2(1+\kappa)^2}} + 8\mathbb{P}_\theta(E_{\mathrm{RIP}}^c),$$

where Lemma C.3(b) was used in the last step. If $m \geqslant 8\pi\frac{C^2(1+\kappa)^2}{\delta^2}\log(1/\eta)$ (which is again weaker than (5.12)), we find using (5.13),

$$\mathbb{E}_g \sup_{y\in N_\delta} \mathbb{P}_\theta(|X_y| \geqslant \delta) \leqslant 8\eta + 8\mathbb{P}_{\theta,g}(E_{\mathrm{RIP}}^c) \leqslant 16\eta. \tag{5.14}$$

Combining the estimates (5.11) and (5.14) it follows that $\mathbb{P}(E^c) \leqslant c\eta$. In order to show the result we still need to choose $\kappa > 0$ and distinguish two cases to this end.

Case 1: Assume that $0 < \delta \leqslant \delta_0 = \left(\log(s)\sqrt{\log(N)}\right)^{-1}$ and choose $\kappa = 1$. A non-trivial $s \geqslant 1$ is only allowed by (4.1) if $\delta \gtrsim 1/\sqrt{N}$. In this situation we have $\log(3e(1+\kappa)N/(s\delta)) \simeq \log(N)$. Then (4.1) implies (5.9). Moreover, (4.2) implies both (5.8) and (5.12) for our conditions on the parameters κ and δ.

Case 2: If $\delta_0 < \delta \leqslant 1$ we choose $\kappa = \sqrt{\delta\log(s)}\log^{1/4}(N) > 1$. Again, a non-trivial $s \geqslant 1$ implies $\delta \gtrsim 1/\sqrt{N}$ by (4.1) and also in this case we have $\log(3e(1+\kappa)N/(s\delta)) \simeq \log(N)$. Plugging our choice of κ into (5.8) and (5.12), we observe that both these conditions are implied by (4.3).

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

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

with probability at least $1-\eta$. Let us first recall that $C_{s,N} \subset 2\mathrm{conv}(\Sigma_{s,N})$ [**22**, Lemma 3.1]. Hence, under E_{RIP},

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

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

$$\log|\mathcal{N}_{\delta/(1+\kappa)}| \lesssim \frac{(1+\kappa)^2}{\delta^2}\left(\mathbb{E}\sup_{x\in C_{s,N}}\langle g,x\rangle\right)^2 \leqslant \frac{4(1+\kappa)^2}{\delta^2}\left(\mathbb{E}\sup_{x\in\Sigma_{s,N}}\langle g,x\rangle\right)^2 \tag{5.15}$$
$$\lesssim \frac{4(1+\kappa)^2}{\delta^2} s\log(eN/s), \tag{5.16}$$

where the final inequality is [23, Lemma 2.3]. By now chasing through the argument above we arrive at the three conditions (replacing (5.8), (5.9) and (5.12), respectively)

$$
\begin{aligned}
m &\gtrsim \kappa^{-2} s(\log^2(s)\log^2(N) + \log(1/\eta)) \\
N &\gtrsim \frac{(1+\kappa)^2}{\delta^4} s^2 \log(eN/s) + s\frac{1}{\delta^2}\log(1/\eta) \\
m &\gtrsim \frac{(1+\kappa)^4}{\delta^4} s\log(eN/s) + \frac{(1+\kappa)^2}{\delta^2}\log(1/\eta).
\end{aligned}
\tag{5.17}
$$

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

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

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

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

$$
\begin{aligned}
E_{\mathrm{RIP},B} &= \left\{\forall\, z \in \Sigma_{2s,N+1} \;\; \frac{1}{\sqrt{m}}\|Bz\|_2 \leqslant 2 + C(1+\kappa)\right\} \\
E_{\Gamma,h,\ell_2} &= \left\{\forall y \in N_\delta \;\; \frac{1}{\sqrt{N}}\|[\Gamma_g\, h]y\|_2 \leqslant 4\right\} \\
E_{\Gamma,h,\ell_1} &= \left\{\forall y \in N_\delta \;\; \left|\frac{1}{N}\sqrt{\frac{\pi}{2}}\|[\Gamma_g\, h]y\|_1 - 1\right| \leqslant \delta\right\}
\end{aligned}
\tag{5.18}
$$

hold with probability at least $1 - c\eta$. For E_{Γ,h,ℓ_1} this is immediate from (5.5) and a union bound. For $E_{\mathrm{RIP},B}$, observe that

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

We have already seen that the event $E_I = \{\frac{m}{2} \leqslant |I| \leqslant \frac{3m}{2}\}$ holds with probability $1 - \eta$. Under this event, the Hanson–Wright inequality (5.1) yields

$$
\mathbb{P}_h\left(\frac{1}{m}\|D_\theta h\|_2^2 \geqslant 2\right) \leqslant \mathbb{P}_h\left(\frac{1}{|I|}\|D_\theta h\|_2^2 \geqslant \frac{4}{3}\right) \leqslant \mathrm{e}^{-c|I|} \leqslant \mathrm{e}^{-cm/2} \leqslant \eta
$$

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

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

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

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

under the conditions (5.17). Very similarly, one shows that E_{Γ,h,ℓ_2} holds with probability at least $1-\eta$ under (5.17). As before, distinguishing two cases for δ one arrives at the statement of the theorem. Finally, we note that a rescaling argument leads to a failure probability of $1-\eta$ instead of $1-c\eta$. □

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 [**7**] and for compressed sensing involving uniformly scalar quantized measurements [**14**,**21**]. In this section, we briefly sketch the implications of Theorem 5.1 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

$$m \gtrsim s \max\left\{\log^{4/3}(s)\log^{5/3}(N), \frac{\log(1/\eta)}{\log^{2/3}(s)\log^{1/3}(N)}, \frac{\log^{2/3}(s)\log^{1/3}(N)\log(1/\eta)}{s}\right\}. \tag{6.1}$$

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

$$\min_{z\in\mathbb{C}^n}\|z\|_1 \qquad \text{s.t.} \qquad \|y-Az\|_1 \leqslant \varepsilon$$

satisfies

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

Proof. As is argued in the proof of [**7**, Theorem III.3], it suffices to show that with probability at least $1-\eta$,

$$\frac{1}{m}\|Ax\|_1 \geqslant c\|x\|_2, \qquad \text{for all } x\in\Sigma_{s,N}^{\mathrm{eff}}, \tag{6.2}$$

for a universal constant $c>0$. Hence, the result immediately follows from Theorem 5.1 by choosing $\delta = c$ constant. □

Note that after estimating $\log(s)\leqslant\log(N)$ and for inverse polynomial probability $\eta = N^{-2}$, say, condition (6.1) takes the simpler form

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

In addition, we can use Theorem 5.1 to derive the following reconstruction result involving a uniform scalar quantizer with dithering of the form $\tau+u$, where τ is Gaussian and u is uniformly distributed.

Let $Q_\alpha : \mathbb{R}^m \to (\alpha\mathbb{Z} + \alpha/2)^m$ be the uniform scalar quantizer with resolution α defined by $Q_\alpha(z) = \big(\alpha\lfloor z_i/\alpha\rfloor + \alpha/2\big)_{i=1}^m$.

THEOREM 6.1 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 $[0, \alpha]$ and are independent of τ. Assume that, for $0 < \epsilon, \eta \leqslant 1$,

$$\begin{aligned} s &\lesssim \min\left\{\sqrt{N/\log^2(N)}, N/\log(1/\eta)\right\} \\ m &\gtrsim s \max\left\{\log^{4/3}(s)\log^{5/3}(N), \frac{\log(1/\eta)}{\log^{2/3}(s)\log^{1/3}(N)}, \frac{\log^{2/3}(s)\log^{1/3}(N)\log(1/\eta)}{s},\right. \\ &\qquad\qquad \left. R^2\alpha^{-2}\epsilon^{-6}\log(eN/s), \frac{\log(1/\eta)}{\epsilon^2 s}\right\}. \end{aligned} \tag{6.3}$$

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

$$\min \|z\|_1 \quad \text{s.t.} \quad \|z\|_2 \leqslant R,\ Q_\alpha\left(\sqrt{\frac{\pi}{2}}Az + \tau + u\right) = Q_\alpha\left(\sqrt{\frac{\pi}{2}}Ax + \tau + u\right) \tag{6.4}$$

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

Note that the program (6.4) is convex. Indeed, the second condition in (6.4) is equivalent to

$$\frac{\alpha}{2} \leqslant \left(\sqrt{\frac{\pi}{2}}Az + \tau + u\right)_i - Q_\alpha\left(\sqrt{\frac{\pi}{2}}Ax + \tau + u\right)_i < \frac{\alpha}{2}, \qquad i = 1, \ldots, m.$$

Furthermore, the reconstruction error in Theorem 6.1 tends to zero if α tends to 0. To see this, apply Theorem 6.1 for $\epsilon = \alpha^{-1/3}\kappa$ and $\kappa > 0$. This roughly speaking shows that $m \gtrsim R^2\kappa^{-6}\log(eN/s)$ measurements suffice to yield an error $\|x^\# - x\|_2 \lesssim \alpha^{2/3}\kappa$.

To place Theorem 6.1 into context, let us compare this result to [**21**, Proposition 1], which concerns unstructured Gaussian measurement matrices and dithering with uniformly distributed thresholds. The latter result implies that if $m \gtrsim \alpha^{-2}\epsilon^{-4}s\log(en/s)$ then with probability at least $1 - 2e^{-c\epsilon m}$ one can reconstruct any effectively s-sparse vector in $B_{\ell_2^N}$ up to error $\epsilon(1+\alpha)$ via a program analogous to (6.4). The slightly better scaling in ϵ suggests that the dependence of m in ϵ in (6.3) is likely not optimal (for partial circulant matrices).

Proof. Let $x^\#$ be any solution to (6.4). Since x is feasible for (6.4),

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

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

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

Moreover, by the last condition in (6.4),

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

where $B = R_I[\Gamma_g\ h]$ and 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 (6.3), Theorem 5.1 implies that

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

with probability at least $1 - \eta$. By [14, Proposition 1] we obtain that under this event, $\sqrt{\frac{\pi}{2}}B$ satisfies with probability at least $1 - \eta$ a quantized version of $\mathrm{RIP}_{1,2}$: for some universal constant c and any $z, z' \in \Sigma^{\mathrm{eff}}_{(\sqrt{s}+1)^2, N+1} \cap \sqrt{2}R\, B_{\ell_2^{N+1}}$,

$$\frac{1}{2}\|z - z'\|_2 - c\alpha\epsilon \leqslant \left\| Q_\alpha\left(\sqrt{\frac{\pi}{2}}Bz + u\right) - Q_\alpha\left(\sqrt{\frac{\pi}{2}}Bz' + u\right)\right\|_1 \leqslant \frac{1}{2}\|z - z'\|_2 + c\alpha\epsilon \tag{6.5}$$

provided that

$$m \gtrsim \epsilon^{-2}\mathscr{N}(\Sigma^{\mathrm{eff}}_{(\sqrt{s}+1)^2, N+1} \cap \sqrt{2}R\, B_{\ell_2^{N+1}}, \|\cdot\|_2, \alpha\epsilon^2) + \epsilon^{-2}\log(1/\eta), \tag{6.6}$$

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

$$\begin{aligned} \Sigma^{\mathrm{eff}}_{(\sqrt{s}+1)^2, N+1} \cap \sqrt{2}R\, B_{\ell_2^{N+1}} &\subset \sqrt{2}R\{x \in \mathbb{R}^{N+1}\ \ \|x\|_1 \leq \sqrt{2s},\ \|x\|_2 \leqslant 1\} \\ &\subset 2\sqrt{2}R\, \mathrm{conv}(\Sigma_{2s, N+1}), \end{aligned}$$

where the final inclusion holds by [22, Lemma 3.1]. Therefore, similarly to (5.15), Sudakov's inequality (Theorem C.1) implies that

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

and it follows that (6.6) is satisfied under our assumptions (6.3). We can now apply (6.5) with $z = [x, R]$ and $z' = [x^{\#}, R]$ and use $\|z - z'\|_2 = \|x - x^{\#}\|_2$ to obtain the asserted error bound. □

Acknowledgements

S.D. would like to thank Laurent Jacques, Christian Kümmerle and Rayan Saab for stimulating discussions.

Funding

All authors acknowledge funding from the Deutsche Forschungsgemeinschaft (DFG) through the project Quantized Compressive Spectrum Sensing (QuaCoSS) which is part of the Priority Program SPP 1798 Compressive Sensing in Information Processing (COSIP). H.J. and H.R. acknowledge funding by the German Israel Foundation (GIF) through the project Analysis of Structured Random Matrices in Recovery Problems (G-1266).

References

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

2. Baraniuk, R. G., Foucart, S., Needell, D., Plan, Y. & Wootters, M. (2017) Exponential decay of reconstruction error from binary measurements of sparse signals. *IEEE Trans. Inform. Theory*, **63**, 3368–3385.

3. Boucheron, S., Lugosi, G. & Massart, P. (2013) *Concentration Inequalities*. Oxford: Oxford University Press.

4. Boufounos, P. T. (2012) Universal rate-efficient scalar quantization. *IEEE Trans. Inform. Theory*, **58**, 1861–1872.

5. Boufounos, P. T. & Baraniuk, R. G. (2008) 1-bit compressive sensing. *2008 42nd Annual Conference on Information Sciences and Systems*. Princeton, NJ: IEEE, pp. 16–21.

6. Cai, T. & Zhang, A. (2014) Sparse representation of a polytope and recovery of sparse signals and low-rank matrices. *IEEE Trans. Inform. Theory*, **60**, 122–132.

7. Dirksen, S., Lecué, G. & Rauhut, H. (2018) On the gap between restricted isometry properties and sparse recovery conditions. *IEEE Trans. Inform. Theory*, **64**, 5478–5487.

8. Feng, J.-M., Krahmer, F. & Saab, R. (2017) Quantized compressed sensing for partial random circulant matrices. ArXiv:1702.04711.

9. Foucart, S. (2017) Flavors of compressive sensing. *Approximation Theory XV: San Antonio 2016* (G. E. Fasshauer & L. L. Schumaker, eds). Cham: Springer International Publishing, pp. 61–104.

10. Foucart, S. & Rauhut, H. (2013) *A Mathematical Introduction to Compressive Sensing*. Applied and Numerical Harmonic Analysis. New York, NY: Birkhäuser.

11. Gopi, S., Netrapalli, P., Jain, P. & Nori, A. (2013) One-bit compressed sensing: provable support and vector recovery. *Proceedings of the 30th International Conference on Machine Learning* (S. Dasgupta & D. McAllester eds). *Proceedings of Machine Learning Research*, vol. 28. Atlanta, Georgia, USA: PMLR, pp. 154–162.

12. Gray, R. M. & Neuhoff, D. L. (1998) Quantization. *IEEE Trans. Inform. Theory*, **44**, 2325–2383.

13. Hanson, D. & Wright, F. (1971) A bound on tail probabilities for quadratic forms in independent random variables. *Ann. Math. Statist.*, **42**, 1079–1083.

14. Jacques, L. & Cambareri, V. (2016) Time for dithering: fast and quantized random embeddings via the restricted isometry property. ArXiv:1607.00816.

15. Jacques, L., Laska, J. N., Boufounos, P. T. & Baraniuk, R. G. (2013) Robust 1-bit compressive sensing via binary stable embeddings of sparse vectors. *IEEE Trans. Inform. Theory*, **59**, 2082–2102.

16. James, D. & Rauhut, H. (2015) Nonuniform sparse recovery with random convolutions. *2015 International Conference on Sampling Theory and Applications*. Washington, DC: IEEE, pp. 34–38.

17. Knudson, K., Saab, R. & Ward, R. (2016) One-bit compressive sensing with norm estimation. *IEEE Trans. Inform. Theory*, **62**, 2748–2758.
18. Krahmer, F., Mendelson, S. & Rauhut, H. (2014) Suprema of chaos processes and the restricted isometry property. *Comm. Pure Appl. Math.*, **67**, 1877–1904.
19. Ledoux, M. & Talagrand, M. (1991) *Probability in Banach Spaces*. Berlin: Springer.
20. Mendelson, S., Rauhut, H. & Ward, R. (2016) Improved bounds for sparse recovery from subsampled random convolutions. ArXiv:1610.04983.
21. Moshtaghpour, A., Jacques, L., Cambareri, V., Degraux, K. & Vleeschouwer, C. D. (2016) Consistent basis pursuit for signal and matrix estimates in quantized compressed sensing. *IEEE Signal Process. Lett.*, **23**, 25–29.
22. Plan, Y. & Vershynin, R. (2013a) One-bit compressed sensing by linear programming. *Comm. Pure Appl. Math.*, **66**, 1275–1297.
23. Plan, Y. & Vershynin, R. (2013b) Robust 1-bit compressed sensing and sparse logistic regression: a convex programming approach. *IEEE Trans. Inform. Theory*, **59**, 482–494.
24. Rauhut, H. (2009) Circulant and Toeplitz matrices in compressed sensing. *Proceedings of SPARS'09*.
25. Rauhut, H. (2010) Compressive sensing and structured random matrices. *Theoretical Foundations and Numerical Methods for Sparse Recovery* (M. Fornasier ed.). Radon Series on Computational and Applied Mathematics, vol. 9. Berlin/New York: de Gruyter, 1–92.
26. Roberts, L. (1962) Picture coding using pseudo-random noise. *IRE Trans. Inform. Theory*, **8**(2) 145–154.
27. Romberg, J. (2009) Compressive sensing by random convolution. *SIAM J. Imaging Sci.*, **2**, 1098–1128.
28. Rudelson, M. & Vershynin, R. (2013) Hanson–Wright inequality and sub-Gaussian concentration. *Electron. Commun. Probab.*, **18**, 1–9.

A. Uniform scalar quantization and recovery via ℓ_∞-constrained ℓ_1-minimization

Let $Q_\alpha : \mathbb{R}^m \to (\alpha\mathbb{Z} + \alpha/2)^m$ be the uniform scalar quantizer with resolution α defined in Section 6. Suppose that we observe (undithered) quantized measurements $y = Q_\alpha(Ax)$. To reconstruct the signal we consider an ℓ_1-minimization problem with a quantization consistency constraint, i.e.,

$$\min \|z\|_1 \qquad \text{s.t.} \qquad y = Q_\alpha(Az). \tag{A.1}$$

The following result concerning reconstruction via (A.1) based on the standard RIP apparently has not been observed before (but see [7] for a discussion of non-optimal results in the literature 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 (2.1) for $\delta_{2s} < 4/\sqrt{41} \approx 0.62$. Then for any $x \in \mathbb{R}^N$ and $y = Q_\alpha(Ax)$ any solution $x^\sharp$ to (A.1) satisfies

$$\|x - x^\sharp\|_2 \lesssim \alpha + s^{-1/2} \inf_{w\in\mathbb{R}^N, \|w\|_0 \leqslant s} \|x - w\|_1. \tag{A.2}$$

Proof. The optimization problem (A.1) is closely related to the ℓ_∞-constrained ℓ_1-minimization problem

$$\min \|z\|_1 \qquad \text{s.t.} \qquad \|Az - y\|_\infty \leqslant \alpha/2. \tag{A.3}$$

In fact, either a minimizer of (A.1) exists, in which case it is also a minimizer of (A.3) or no minimizer of (A.1) exists, in which case the theorem is void. (A minimizer of (A.3) always exists so that it may be preferred in practice. The error bound (A.2) still holds for (A.3), but every minimizer x^* of (A.3) is quantization inconsistent (i.e., $Q_\alpha(Ax^*) \neq Q_\alpha(Ax)$) in the case that no minimizer of (A.1) exists.)

This close relation of (A.1) and (A.3) suggests to study versions of the null space property, see, e.g., [10, Chapter 4]. By [10, Theorem 6.13], the bound on the restricted isometry constants of $\frac{1}{\sqrt{m}}A$ implies the ℓ_2-robust null space property in the form

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

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

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

which is the ℓ_∞-robust null space property of order s. By [10, Theorem 4.25] this implies that any minimizer x^* of (A.3) satisfies

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

This concludes the proof. □

By [18], see also Section 2, a partial random circulant matrix $\frac{1}{\sqrt{|I|}}R_I\Gamma_g$ with subsampling on a fixed (deterministic) set $I \subset [N]$ generated by a standard Gaussian random vector satisfies $\delta_{2s} \leqslant 0.6$ 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 $y = Q_\alpha(R_I\Gamma_g)$ via (A.1) under this condition on the number of measurements.

B. Upper RIP bound for partial random circulant matrices

The proof of our main result only requires the upper (ℓ_2)-RIP bound in (2.1). For (deterministic) subsampling of random circulant matrices we can deduce the following bound on m from [18], valid also for values of $\kappa \geqslant 1$.

Theorem B.1 Let I be a fixed (deterministic) subset of $[N]$ and let $A = R_I\Gamma_g$ be the associated subsampled Gaussian circulant matrix. Fix $\kappa > 0$. If

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

then $\sup_{x\in\Sigma_{s,N}} \frac{1}{\sqrt{|I|}}\|Ax\|_2 \leqslant C(1+\kappa)$ with probability at least $1 - \eta$.

Proof. As argued in [18], we can write $\frac{1}{\sqrt{|I|}}Ax = V_x g$ with $V_x = \frac{1}{\sqrt{|I|}}R_I\Gamma_x$. Denote, $\mathscr{A}_{s,N} = \{V_x : x \in \Sigma_{s,N}\}$. It follows then from [18, Theorem 3.5(a)] that for every $p \geqslant 1$,

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

where $\gamma_2(\mathscr{A}_{s,N}, \|\cdot\|_{\ell_2\to\ell_2})$ denotes the γ_2-functional of the set $\mathscr{A}_{s,N}$ with respect to the spectral norm, $d_{\ell_2\to\ell_2}$ and d_F denote the diameters in the spectral and Frobenius norms, respectively; see [18] for details. These parameters have been estimated in [18, Section 4],

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

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

Moreover, the moment bound (B.1) implies (see, e.g., [18, Prop. 2.6]) that for all $t \geqslant 1$,

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

where $c, C > 0$ are absolute constants. Requiring that the right hand is bounded by η gives the statement of the theorem. □

C. Some tools from probability

In our proof we use Sudakov's inequality; see, e.g., [19, Theorem 3.18] for a proof.

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

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

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

LEMMA C.1 Let X be a random variable and let med(X) be a median of X. Then,

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

Moreover, if X' is an independent copy of X, then for any $\delta > 0$,

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

The estimates in Lemma C.1 are well known. We provide a proof for convenience.

Proof. We may assume that $\sigma := (\mathbb{E}(X-\mathbb{E}X)^2)^{1/2} < \infty$. Fix any $\varepsilon > 0$. By Cantelli's inequality,

$$\mathbb{P}(X - \mathbb{E}X \geqslant (\sigma+\varepsilon)) \leqslant \frac{\sigma^2}{\sigma^2 + (\sigma+\varepsilon)^2} < \frac{1}{2}$$

and

$$\mathbb{P}(X - \mathbb{E}X \leqslant -(\sigma+\varepsilon)) \leqslant \frac{\sigma^2}{\sigma^2 + (\sigma+\varepsilon)^2} < \frac{1}{2}.$$

Hence,

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

Taking $\varepsilon \to 0$ yields the first statement. To prove the second statement, observe that since X and X' are independent,

$$\mathbb{P}(X' \leqslant \mathrm{med}(X))\mathbb{P}(X \geqslant \mathrm{med}(X) + \delta) \leqslant \mathbb{P}(X \geqslant X' + \delta).$$

Since X and X' are identically distributed, $\mathbb{P}(X' \leqslant \mathrm{med}(X)) \geqslant \frac{1}{2}$ and it follows that

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

By replacing X, X' by $-X, -X'$ and using that $-\operatorname{med}(X)$ is a median of $-X$ we find

$$\mathbb{P}(X - \operatorname{med}(X) \leqslant -\delta) \leqslant 2\mathbb{P}(X - X' \leqslant -\delta).$$

Combining these estimates, we find

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

□

The following probability bound related to symmetrization follows in the same way as in [19, equation (6.3)].

LEMMA C.2 Let (X_t) be a family of random variables indexed by a countable set T and let (X'_t) be an independent copy of (X_t). Then, for $x, y > 0$,

$$\mathbb{P}\left(\sup_{t\in T} |X_t| \geqslant x + y\right) \leqslant \mathbb{P}\left(\sup_{t\in T} |X_t - X'_t| \geqslant x\right) + \sup_{t\in T} \mathbb{P}(|X_t| \geqslant y).$$

In our proofs we repeatedly use the following simple facts.

LEMMA C.3 Let X be a real-valued random variable and $\mathcal{F}$ be an event.

(a) For any $s > 0$,

$$\mathbb{P}(X > s) \leqslant \mathbb{P}(1_{\mathscr{F}} X > s) + \mathbb{P}(\mathscr{F}^c).$$

(b) If $X \leqslant L$ almost surely for some $L \in \mathbb{R}$ then

$$\mathbb{E}X \leqslant \mathbb{E}\left[1_{\mathscr{F}} X\right] + L\mathbb{P}(\mathscr{F}^c).$$

Proof. By the union bound

$$\begin{aligned}\mathbb{P}(X > s) &= \mathbb{P}\left(\{1_{\mathscr{F}} X > s\} \cup \{1_{\mathscr{F}^c} X > s\}\right) \leqslant \mathbb{P}(1_{\mathscr{F}} X > s) + \mathbb{P}(1_{\mathcal{F}^c} X > s)\\ &\leqslant \mathbb{P}(1_{\mathscr{F}} X > s) + \mathbb{P}(\mathcal{F}^c),\end{aligned}$$

where the last step used that $s > 0$. If $X \leqslant L$ almost surely, then

$$\mathbb{E}X = \mathbb{E}[1_{\mathscr{F}} X] + \mathbb{E}[1_{\mathscr{F}^c} X] \leqslant \mathbb{E}[1_{\mathscr{F}} X] + L\mathbb{P}(\mathscr{F}^c).$$

□