An RIP-based approach to ΣΔ quantization for compressed sensing

Joe-Mei Feng and Felix Krahmer

October 1, 2013

Abstract

In this paper, we provide new approach to estimating the error of reconstruction from ΣΔ quantized measurements for compressed sensing. Our method is based on the restricted isometry property (RIP) of a certain projection of the measurement matrix. The main application of our result is the error analysis for partial random circulant matrices. This is the first time such bounds are provided for a structured measurement matrix with the fast multiplication property. Our results also recover the best-known reconstruction error bounds for Gaussian and subgaussian measurement matrix.

1 Introduction

1.1 Compressed sensing

Compressed sensing has drawn significant attention since the seminal works by Candés, Romberg, Tao [6], and Donoho [10]. The theory of compressed sensing is based on the observation that various cases of natural signals are approximately sparse with respect to certain bases or frames. The basic idea is to recover such signals from a small number of linear measurements. Hence the problem turns into an undetermined linear system. Various criteria have been proposed to determine whether such a system has a unique sparse solution. In this paper we will work with the restricted isometry property (RIP), which has been shown by Candés et al. [8] to guarantee uniqueness.

Definition 1. A matrix $A \in \mathbb{R}^{m\times N}$ has the restricted isometry property (RIP) of order s if there exists $0 < \delta < 1$ such that for all s-sparse vectors $x \in \mathbb{C}^N$, i.e., vectors that have at most s non-zero components,

$$(1-\delta)\|x\|_2^2 \leq \|Ax\|_2^2 \leq (1+\delta)\|x\|_2^2.$$

The smallest such δ is called the RIP-constant of order s denoted by δ_s.

Finding the RIP-constant of a measurement matrix is in general an NP hard problem [23]. That is why most papers work with random matrices.

Examples of random matrices known to have the RIP include subgaussian, partial random circulant, and partial Fourier matrices. A subgaussian matrix has independent random entries whose tail is dominated by a Gaussian (cf. Definition 3). Such matrices have been shown to have the RIP provided $m = \Omega(s\log(eN/s))$ [18]. This order of the embedding dimension m is known to be optimal. Examples of subgaussian matrices include Gaussian and Bernoulli matrices. A partial random circulant matrix is the matrix representation of a deterministically subsampled random convolution. Such a matrix has been shown to have the RIP provided $m = \Omega(s(\log s)^2(\log N)^2)$ [11]. A partial random Fourier matrix consists of m rows of a discrete Fourier matrix drawn at random. The RIP for such a matrix has been shown provided $m = \Omega(s\log^3 s\log N))$ [21]. These embedding dimensions are slightly worse than for subgaussian matrices, but such structured matrices are preferred in many applications, e.g., because of their reduced randomness and their fast multiplication properties.

1.2 Quantized compressed sensing

While the measurements in compressed sensing are usually not representable by finite bits, we need to quantize these measurements such that the system can be operated on computers. To quantize the measurements we seek to represent our measurements by finite many symbols from a finite alphabet consisting of real numbers. The extreme case of considering the set of only two elements $\{-1, 1\}$ is also called 1-bit quantization. The most intuitive method to quantize the measurements is to map each of the measurements to the closest element from the alphabet set. Since this method processes the quantization independently for each measurement, it is also called memoryless scalar quantization (MSQ).

Most of the literature on MSQ quantization compressed sensing up to date consider 1-bit quantization [5, 17, 19, 1], which bias down to considering only the measurement signs. Jacques et al. [17] show for Gaussian measurements or measurements drawn uniformly from the unit sphere, the worst case error is bounded by $O(\frac{s}{m}\log\frac{mN}{s})$. Later, for Gaussian measurements, Gupta et al. [14] demonstrate that one may tractably recover the support of the signal from $O(s\log n)$ measurements. Plan et al. [19] show that one can again for Gaussian measurements, reconstruct the direction of an s-sparse signal via convex optimization, with accuracy $O(\frac{s}{m})^{\frac{1}{5}}$ up to logarithmic factors with high probability. Later Ai et al. [1] derive similar results for subgaussian measurements under additional assumptions on the size of the signal entries.

However in [17] it is shown that the reconstruction ℓ_2-error can never be better than $\Omega(\frac{s}{m})$. To break this bottleneck of MSQ, $\Sigma\Delta$ quantization for compressed sensing has drawn attention recently. $\Sigma\Delta$ quantizes a vector as a whole rather than the components individually, i.e., the quantized values depend on previous quantization steps. Güntürk et al. [13] show that for rth order $\Sigma\Delta$ quantization in Gaussian compressed sensing case, the ℓ_2-error is bounded by $O((\frac{s}{m})^{\alpha(r-\frac{1}{2})})$ for any $0 < \alpha < 1$ with high probability. More recently, in [18], this result has been generalized to subgaussian measurements. Indeed for r large enough this breaks the MSQ bottleneck.

1.3 Contributions

The primary contribution of this paper is that the restricted isometry property (RIP) is applied to estimate the error bound for $\Sigma\Delta$ quantized compressed sensing. That is, once we know the RIP-constant of a modification of the measurement matrix, we can estimate the reconstruction error.

In the following results, we assume that the $\Sigma\Delta$ quantized measurements with quantization alphabet $\mathscr{Z} = \Delta\mathbb{Z}$ are given to us. We refer the readers to Section 2.2 for details on the quantization scheme employed. A special role is played by the rth power finite difference matrix D; denoting the singular value decomposition of D^{-r} by $D^{-r} = U_{D^{-r}} S_{D^{-r}} V^*_{D^{-r}}$, we obtain our main theorem below.

Theorem 1. *Given an s-sparse signal $x \in \mathbb{R}^N$, denoted by $\Phi \in \mathbb{R}^{m\times N}$ a measurement matrix, and q the rth order $\Sigma\Delta$-quantized measurements of Φx with step size Δ. Suppose Φ has the RIP such that the support set T can be determined. Choose L as the Sobolev dual matrix of Φ_T and reconstruct the signal by $\hat{x} = Lq$ (see Section 2.2.2 for details), if $\sqrt{\frac{1}{\ell}} P_\ell V^*_{D^{-r}} \Phi$, $\ell \leq m$, has RIP-constant $\delta_s \leq \delta$, where P_ℓ maps a vector to its first ℓ components. then the reconstruction error is bounded above by*

$$\|x - \hat{x}\|_2 \leq \frac{\Delta}{2c_2(r)\sqrt{(1-\delta)}} (\frac{m}{\ell})^{-r+\frac{1}{2}},$$

where $c_2(r) > 0$ is a constant depending only on r.

Note from Theorem 1, the smaller ℓ is the better the bound. However, ℓ has to be large enough such that $\frac{1}{\sqrt{m}}(P_\ell V^*_{D^{-r}} \Phi)$ has the RIP-constant $\delta_s \leq \delta$.

This result can be applied to obtain recovery guarantees for various compressed sensing setting such as Gaussian, subgaussian and partial random circulant measurements. We will show in later section that our result covers which in [13] and [18]. Since the result of reconstruction bound on random circulant matrix has not been proposed so far, we state it as the following.

Theorem 2. *Given an s-sparse signal $x \in \mathbb{R}^N$, let $\Phi x = P_\Omega(\xi * x) \in \mathbb{R}^{m\times N}$ be a partial random circulant matrix, where the ξ_i's are independent mean-zero, ρ-subgaussian random variables (see Definition 3) of variance one and q the rth order $\Sigma\Delta$-quantization of Φx with step size Δ. If*

$$m \geq C_1 s \log^{\frac{2}{1-2\alpha}} s \log^{\frac{2}{1-2\alpha}} N,$$

where $0 \leq \alpha \leq \frac{1}{2}$, C_1 depends only on ρ, and m large enough such that the support set T is determined, choosing L as the Sobolev dual matrix of Φ_T, reconstructing the signal by $\hat{x} = Lq$, then, with probability at least 0.99, the reconstructed error is bounded by

$$\|x - \hat{x}\|_2 \lesssim_\rho \Delta(\frac{m}{s})^{-\alpha(r-\frac{1}{2})},$$

the notation $\lesssim_\rho$ means less than or equal to up to a constant depending on ρ.

1.4 Organization

The following paper is organized as the following. We first introduce the background and previous results on $\Sigma\Delta$ quantization, suprema of chaos process, and the restricted isometry property for partial random circulant matrices. Further we present our main result in Section 3, where we show how the RIP is used to estimate reconstruction error for quantized compressed sensing. In Section 4, we show that our result can cover the best-known bounds for Gausssian and subgaussian measurement matrices. In the last section, we apply our result on partial random circulant measurement matrices. This is the first time such bounds are provided for a structured measurement matrix with the fast multiplication property.

2 Background and previous results

2.1 Notations

Through out this paper we use the following notations. The set $D_{s,N} = \{x \in \mathbb{C} | \|x\|_2 \leq 1, \|x\|_0 \leq s\}$ is the set of unit norm s-sparse vectors. ℓ_0-norm $\|\cdot\|_0$ counts the number of non-zero components of a vector. Denote the Frobenius norm by $\|A\|_F = \sqrt{trA^*A}$, and ℓ_2-operation norm $\|A\|_{2\to 2} = \sup_{\|x\|_2=1} \|Ax\|_2$. Based on last two norms $d_F(\mathcal{B}) = \sup_{A\in\mathcal{B}} \|A\|_F$, and $d_{2\to 2}(\mathcal{B}) = \sup_{A\in\mathcal{B}} \|A\|_{2\to 2}$. We also assign orders to singular value of a matrix as $\sigma_i(\cdot)$ and $\sigma_{\min}(\cdot)$ indicating the ith largest and the smallest singular value of a matrix respectively. Also denoted by $\gtrsim$, $\geq$ up to a positive constant, while $\lesssim$ means $\leq$ up to a positive constant. Denoted by † the Moore-Penrose inversion operator.

2.2 $\Sigma\Delta$ Quantization

$\Sigma\Delta$ quantization is originally introduced as an efficient quantizer for redundant representations of oversampled band-limited functions [16]. Later on the mathematical analysis is provided by [9, 12] and many follow-up papers, and further the scheme has been extended to frame expansions [3].

Given an alphabet $\mathcal{Z}$ such that $\mathcal{Z} = \Delta\mathbb{Z}$, the idea of rth order $\Sigma\Delta$ quantization is to quantize each component of a vector taking the previous r quantization steps into account. More explicitly, a greedy rth order $\Sigma\Delta$ quantization maps a sequence of inputs (y_j) to elements $q_i \in \mathcal{Z}$ via an internal state variable u_i, explicitly

$$(\Delta^r u)_i = \sum_{j=0}^{r} \left(\begin{array}{c} r \\ j \end{array} \right) (-1)^j u_{i-j} = y_i - q_i, \tag{1}$$

where q_i is chosen such that $|u_i|$ is minimized.

Setting initial conditions $(u_i)_{i=0}^{-\infty} = 0$, then equation (1) can be expressed as

$$D^r u = y - q,$$

where the finite difference matrix is given by

$$D_{ij} \equiv \begin{cases} 1 & , \text{ if } i = j, \\ -1 & , \text{ if } i = j+1, \\ 0 & , \text{ otherwise.} \end{cases} \tag{2}$$

2.2.1 Coarse recovery

Given an s-sparse signal x, and an $m \times N$ measurement matrix Φ, where $m \ll N$, we acquire measurements $y = \Phi x$. Applying rth order $\Sigma\Delta$ quantization to y, we have q. If treat q as perturbed measurements, i.e., $q = y + e = \Phi x + e$, then by [13], we can determine the support set. This is proved by a modified version of Proposition 4.1 in [13] and Theorem 1 in [7], which says

Proposition 1. *Given $x \in \mathbb{R}^N$ an s-sparse signal, denote e a noise vector with $\|e\|_2 \leq \epsilon$, and let $\Phi \in \mathbb{R}^{N\times m}$ be a measurement matrix. Reconstruct x from $q = \Phi x + e$ via ℓ_1 minimization we obtain x', i.e.,*

$$\hat{x} = arg\min \|z\|_1 \text{ subject to } \|\Phi z - q\|_2 \leq \epsilon.$$

If $\frac{1}{\sqrt{m}}\Phi$ has RIP-constants such that $\delta_{3s} + 3\delta_{4s} < 2$, then $\|x - \hat{x}\|_2 \leq K\frac{1}{\sqrt{m}}\eta$, and denote T the support set of x and $s = |T|$, and if $\min_{j\in T}|x_j| \geq K2^{r-\frac{1}{2}}\Delta$, $j \in T$, for some positive constant K, then the largest s components of x' is T.

Note the factor $\frac{1}{\sqrt{m}}$ which makes $\frac{1}{\sqrt{m}}\Phi$ a unit-norm-columned matrix, while in the compressed sensing literature, this is common to normalize the measurement matrix such that it has unit-norm columns. However for quantization compressed sensing, if the measurement size varies depending on the scale $\frac{1}{\sqrt{m}}$, then to analysis the reconstruction error it is not justified to use a fix step size Δ. To fairly compare the error while we let m grow, we should leave our measurement entries independent of m, therefore, in this paper the measurement matrices are not normalized, and instead, for convenience we set each entry of the measurement matrices to have variance one.

2.2.2 $\Sigma\Delta$ error estimate and Sobolev dual

According to the previous section, denote the support set of x by T, we solve x by multiplying a left inverse of Φ_T, say L. Then the reconstruction ℓ_2-error is given by

$$\|x - \hat{x}\|_2 = \|Ly - Lq\| = \|L(y - q)\|_2 = \|L(D^r u)\|_2 \leq \|LD^r\|_{2\to 2}\|u\|_2.$$

The Sobolev dual matrix $L_{sob,r}$, first introduced in [4], is a left inverse of Φ_T defined to minimize $\|LD^r\|_{2\to 2}$, i.e.,

$$\min_L \|LD^r\|_{2\to 2} \quad \text{subject to } L\Phi_T = I.$$

The geometric intuition is that the frame is smoothly varying.

Using the explicit formula $L_{sob,r}D^r = (D^{-r}\Phi_T)^\dagger$, we obtain the error bound

$$\|x - \hat{x}\|_2 \leq \|(D^{-r}\Phi_T)^\dagger\|_{2\to 2}\|u\|_2 = \frac{1}{\sigma_{min}(D^{-r}\Phi_T)}\|u\|_2. \tag{3}$$

Recall that $\|u\|_2 \leq 2^{-1}\Delta\sqrt{m}$, once we find a bound for $\sigma_{min}(D^{-r}\Phi_T)$ from below we can bound $\|x - \hat{x}\|_2$ from above.

A key ingredient to bound this singular value is the following result from the study of Toeplitz matrices , which depends highly on Weyl's inequality [15] (see also for example in [13]).

Proposition 2. *Let r be any positive integer and D be as in (2). There are positive constants $c_{s1}(r)$ and $c_{s2}(r)$, independent of m, such that*

$$c_{s1}(r)(\frac{m}{j})^r \leq \sigma_j(D^{-r}) \leq c_{s2}(r)(\frac{m}{j})^r, \; j = 1, \ldots, m. \tag{4}$$

2.3 Suprema of chaos process

Based on [11], in this section we will see one of the most important tools for our main result, which is Theorem 3.

Definition 2. For a metric space (T, d), an admissible sequence of T is a collection of subsets of T, $\{T_r : r \geq 1\}$, $|T_r| \leq 2^{2^r}$ and $|T_0| = 1$, define γ_2 function introduced by Talagrand [22] as

$$\gamma_2(T, d) = \inf_{T_r} \sup_{t \in T} \sum_{r=0}^{\infty} 2^{r/2} d(t, T_r),$$

γ_2 function can be bounded by a Dudley integral [22]. More specificaly, the γ_2 function used in the following theorem is bounded by

$$\gamma_2(\mathcal{B}, \|\cdot\|_{2\to2}) \lesssim \int_0^{d_{2\to2}(\mathcal{B})} \log^{1/2} N(\mathcal{B}, \|\cdot\|_{2\to2}, u) du,$$

where $N(\mathcal{B}, \|\cdot\|_{2\to2}, u)$ is the covering number of $\mathcal{B}$ with respect to the norm $\|\cdot\|_{2\to2}$ and the radius u.

A key ingredient for out main result will be the following theorem.

Theorem 3. *[11] Let $\mathcal{B}$ be a set of matrices, and let ξ be a random vector whose entries are independent, mean zero, variance one, and ρ-subgaussian random variables (see Definition 3). Then*

$$\mathbb{P}(C_{\mathcal{B}} \geq c_1 E + t) \leq 2\exp(-c_2 \min\{\frac{t^2}{V^2}, \frac{t}{U}\}), \tag{5}$$

where $c_1, c_2 > 0$ are constants depending on ρ,

$$C_{\mathcal{B}} := \sup_{A \in \mathcal{B}} |\|A\xi\|_2^2 - \mathbb{E}\|A\xi\|_2^2|,$$

$$E = \gamma_2(\mathcal{B}, \|\cdot\|_{2\to2})(\gamma_2(\mathcal{B}, \|\cdot\|_{2\to2}) + d_F(\mathcal{B})) + d_F(\mathcal{B}) d_{2\to2}(\mathcal{B}),$$

where

$$V = d_{2\to2}(\mathcal{B})(\gamma_2(\mathcal{B}, \|\cdot\|_{2\to2}) + d_F(\mathcal{B})),$$

and

$$U = d_{2\to2}^2(\mathcal{B}).$$

2.4 The restricted isometry property for partial random circulant matrices

In this section we will show the RIP for partial random matrices, which is selected from [11].

The main ingredient to define a partial random circulant matrix is an ρ-subgaussian random variable as below.

Definition 3. A random variable X is called ρ-subgaussian if $\mathbb{P}(|X| \geq t) \leq 2\exp(-t^2/2\rho^2)$

Based on this definition we can define a partial random circulant matrix corresponding to the matrix representation of a subsampled random convolution. Then the definition of a partial random circulant matrix is

Definition 4. Given a random vector $\xi = (\xi_i)_{i=1}^N$, where the ξ_i's are independent mean-zero, ρ-subgaussian random variables of variance one. A partial random circulant matrix $\Phi_{m,N}$, generated by ξ is defined via

$$\Phi_{m,N}x = P_\Omega(\xi * x),$$

where P_Ω is a deterministic projection to m components of a vector.

[11] states the RIP for partial random circulant matrix $\frac{1}{\sqrt{m}}\Phi$ as below.

Theorem 4. *Let $\xi = (\xi_j)_{j=1}^N$ be a random vector with independent mean-zero, variance one, ρ-subgaussian entries. If for $s \le N$ and $\eta, \delta \in (0,1)$,*

$$m \ge c\delta^{-2} s \max\{(\log s)^2(\log N)^2, \log(\eta^{-1})\}, \tag{6}$$

then with probability at least $1-\eta$, the RIP-constant of the partial random circulant matrix $\Phi \in \mathbb{R}^{m\times N}$ generated by ξ satisfies $\delta_s \le \delta$. The constant c depends only on ρ.

3 RIP-based error analysis

In this section we will give the quantized compressed sensing problem a mathematical model, and explain how we approach the reconstruction error via the restricted isometry property. In the next two sections we show its applications. From Section 2.2.2, the main issue to estimate the reconstruction error is to estimate $\sigma_{min}(D^{-r}\Phi_T)$. Comparing to $\sigma_{min}(D^{-r}\Phi)$, since the domain is restricted, $\sigma_{min}(D^{-r}\Phi_T) \ge \sigma_{min}(D^{-r}\Phi)$. The intuition is to find a suprema of the effective smallest singular value of $\sigma_{min}(D^{-r}\Phi)$ while the domain is restricted on $D_{s,N}$. This restriction motivates the concept of RIP. In the following proof we show how RIP can be applied to find this effective smallest singular value.

Proof of Theorem 1. Recall that $D^{-r} = U_{D^{-r}}S_{D^{-r}}V_{D^{-r}}^*$. Then, as s is a diagonal matrix,

$$\begin{aligned}
\sigma_{\min}(D^{-r}\Phi_T) &= \sigma_{\min}(S_{D^{-r}}V_{D^{-r}}^*\Phi_T)\\
&\ge \sigma_{\min}(P_\ell S_{D^{-r}}V_{D^{-r}}^*\Phi_T)\\
&= \sigma_{\min}((P_\ell S_{D^{-r}}P_\ell^*)(P_\ell V_{D^{-r}}^*\Phi_T)_{\ell\times s})\\
&\ge s_\ell \sigma_{\min}(P_\ell V_{D^{-r}}^*\Phi_T)_{\ell\times s},
\end{aligned}$$

Next we need to bound $\sigma_{\min}(P_\ell V_{D^{-r}}^*\Phi_T)$ uniformly over all support set T.

If $\frac{1}{\sqrt{\ell}}P_\ell V_{D^{-r}}^*\Phi$ RIP-constant $\delta_s \le \delta$ then $\sigma_{\min}(P_\ell V_{D^{-r}}^*\Phi_T)_{\ell\times s}$ is uniformly bounded from below by

$$\sqrt{\ell}\sqrt{1-\delta}. \tag{7}$$

The theorem follows by applying (3), (4), (7) as

$$\frac{1}{\sigma_{min}(D^{-r}\Phi_T)}\|u\|_2 \le \frac{\Delta}{2c_2(r)\sqrt{(1-\delta)}}(\frac{m}{\ell})^{-r+\frac{1}{2}} \tag{8}$$

□

4 Gaussian and subgaussian matrices

Given Φ a standard Gaussian random matrix. Since $(P_\ell V_{D-r}^*\Phi)$ is also a standard Gaussian random matrix due to rotation invariance, with $\ell = \Omega(s\log N)$, $\frac{1}{\sqrt{\ell}}(P_\ell V_{D-r}^*\Phi)$ has the RIP-constant $\delta_s < \delta$ with high probability [20]. Since $s \le \ell \le m$,

$$\frac{m}{\ell} \le (\frac{m}{s})^\alpha,\ \ \alpha \in (0,1).$$

Provided that $\ell = \Omega(s\log N)$,

$$\begin{aligned}
&\frac{m}{\ell} \lesssim (\frac{m}{s})^\alpha \\
\Rightarrow &\frac{m}{s\log N} \lesssim (\frac{m}{s})^\alpha \\
\Rightarrow &m \gtrsim s(\log N)^{\frac{1}{\alpha-1}}.
\end{aligned}$$

Apply Theorem 1 directly, we obtain

$$\|x-\hat{x}\|_2 \lesssim \Delta(\frac{m}{s})^{-\alpha(r-\frac{1}{2})},$$

with high probability. This therefore recovers the result in [13].

By similar steps, this can also be generalized to subgaussian measurement matrices, which recovering the result in [2]. As this argument involves some additional technical steps, we refrain from presenting the details here.

5 Partial random circulant matrices

In this section we apply our main theorem, Theorem 1, on partial random circulant matrices to obtain Theorem 2. To prove Theorem 2 we need the following lemma.

Lemma 1. *Let $\xi \in \mathbb{R}^N$ be a random vector with independent mean-zero, variance one, ρ-subgaussian entries, and denote by $\Phi \in \mathcal{R}^{m\times n}$ a partial random circulant matrix generated by ξ. If $\ell \le m$ satisfies*

$$\ell \ge C\sqrt{sm}(\log s)(\log N),$$

where C is a constant depending only on ρ, then, with probability at least 0.99, $\sqrt{\frac{1}{\ell}}P_\ell V_{D-1}^*\Phi$ has the restricted isometry property with constant $\delta_s \le 0.1$.

Proof. The proof of this lemma is based on Theorem 3

Define $\mathcal{B} = \{\sqrt{\frac{1}{\ell}}P_\ell V_{D-r}^* V_x : x \in D_{s,N}\}$, then

$$\begin{aligned}
\delta_s(\sqrt{\frac{1}{\ell}}P_\ell V_{D-1}^*\Phi) &= \sup_{x\in D_{s,N}} \left|\|\sqrt{\frac{1}{\ell}}P_\ell V_{D-r}^* V_x\xi\|_2^2 - \mathbb{E}\|\sqrt{\frac{1}{\ell}}P_\ell V_{D-r}^* V_x\xi\|_2^2\right| \\
&= \sup_{x\in D_{s,N}} \left|\|\sqrt{\frac{1}{\ell}}P_\ell V_{D-r}^* V_x\xi\|_2^2 - \mathbb{E}\|x\|_2^2\right| \\
&= \sup_{A\in\mathcal{B}} \left|\|A\xi\|_2^2 - \mathbb{E}\|A\xi\|_2^2\right| \\
&= C_{\mathcal{B}},
\end{aligned}$$

since for $A \in \mathcal{B}$

$$\mathbb{E}\|A\xi\|_2^2 = \mathbb{E}\xi^* A^* A\xi = \|A\|_F^2 = \|\sqrt{\frac{1}{\ell}} P_\ell V_{D-r}^* V_x\|_F^2 = \|\sqrt{\frac{1}{\ell}} P_\ell V_x\|_F^2 = \|x\|_2^2.$$

Estimating the terms in Theorem 3 using the definition of $\mathcal{B}$ and the corresponding bounds for $\mathcal{A} = \{\frac{1}{\sqrt{m}} V_x : x \in D_{s,N}\}$ derived in [11], we obtain

$$d_F(\mathcal{B}) = \sqrt{\frac{1}{\ell}} \sqrt{m} d_F(\mathcal{A}) \leq \sqrt{\frac{m}{\ell}},$$

and

$$d_{2\to 2}(\mathcal{B}) = \sqrt{\frac{m}{\ell}} d_{2\to 2}(\mathcal{A}) \leq \sqrt{\frac{s}{\ell}}.$$

To estimate $\gamma_2(\mathcal{B}, \|\cdot\|_{2\to 2})$, we need to estimate the covering numbers of the set $\mathcal{B}$ with respect to the norm $\ell^{-1/2}\|\cdot\|_\infty$, while the norm used in the estimation of $\gamma_2(\mathcal{A}, \|\cdot\|_{2\to 2})$ in [11] is $m^{-1/2}\|\cdot\|_\infty$. Paralleling to the arguments to

$$\gamma_2(\mathcal{A}, \|\cdot\|_{2\to 2}) \lesssim \sqrt{\frac{s}{m}} (\log s)(\log N),$$

we obtain

$$\gamma_2(\mathcal{B}, \|\cdot\|_{2\to 2}) \lesssim \sqrt{\frac{s}{\ell}} (\log s)(\log N).$$

Suppose $\ell \geq 3c_1\sqrt{sm}(\log s)(\log N)$, where c_1 is the same as in Theorem 3, and assume $t = \frac{\delta}{2}$ then by definition of E

$$c_1 E + \frac{\delta}{2} \leq c_1\{\sqrt{\frac{s}{\ell}}(\log s)(\log N)[\sqrt{\frac{s}{\ell}}(\log s)(\log N) + \sqrt{\frac{m}{\ell}}] + \sqrt{\frac{m}{\ell}}\sqrt{\frac{s}{\ell}}\} + \frac{\delta}{2} \leq \delta, \tag{9}$$

Further by definition of V and U, one has for C large enough

$$2\exp(-c_2 \frac{t^2}{V^2}) \leq 2\exp(-c_2 (\frac{\frac{\delta}{2}}{\sqrt{\frac{s}{\ell}}[\sqrt{\frac{s}{\ell}}(\log s)(\log N) + \sqrt{\frac{m}{\ell}}]})^2) \leq 0.01 \tag{10}$$

and

$$2\exp(-c_2 \frac{t}{U}) \leq 2\exp(-c_2 (\frac{\delta/2}{s/\ell})) \leq 0.01. \tag{11}$$

Combining equation (10) and (11), we obtain

$$2\exp(-c_2 \min\{\frac{t^2}{V^2}, \frac{t}{U}\}) \leq 0,01. \tag{12}$$

Setting $\delta = 0.1$, we obtain from equation (5) that

$$P(\delta_s \geq 0.1) \leq 0.01.$$

This proves the lemma.

□

Proof of Theorem 2. To determine the support set of s-sparse signal x, we need from Theorem 4 that

$$m \gtrsim s \log^2 s \log^2 N. \tag{13}$$

Now choose $\delta = 0.1$, and $\ell \gtrsim C\sqrt{sm}(\log s)(\log N)$, where C is as in Lemma 1, combining Lemma 1 and Theorem 1. We obtain that the reconstruction error is bounded by

$$\begin{aligned}
\|x - \hat{x}\|_2 &\lesssim \Delta(\frac{m}{\ell})^{-r+\frac{1}{2}} \\
&\approx \Delta(\frac{m}{\sqrt{sm}(\log s)(\log N)})^{-r+\frac{1}{2}} \\
&\lesssim \Delta((\frac{m}{s}))^{-\alpha(r-\frac{1}{2})},
\end{aligned}$$

This follows from the fact that $(m/s)^{\frac{1-2\alpha}{2}} \gtrsim \log s \log N$. This inequality only holds for $\frac{1-2\alpha}{2} \geq 0$, and thus $\alpha \leq \frac{1}{2}$. Together with equation (13), we obtain that $0 \leq \alpha \leq \frac{1}{2}$ This concludes the proof.

□

References

[1] Albert Ai, Alex Lapanowski, Yaniv Plan, and Roman Vershynin. One-bit compressed sensing with non-gaussian measurements. *Linear Algebra and its Applications*, 2013.

[2] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin. A simple proof of the restricted isometry property for random matrices. *Constr. Approx.*, 28(3):253–263, 2008.

[3] John J Benedetto, Alexander M Powell, and Ozgur Yilmaz. Sigma-delta ($\Sigma\Delta$) quantization and finite frames. *Information Theory, IEEE Transactions on*, 52(5):1990–2005, 2006.

[4] James Blum, Mark Lammers, Alexander M Powell, and Özgür Yılmaz. Sobolev duals in frame theory and sigma-delta quantization. *Journal of Fourier Analysis and Applications*, 16(3):365–381, 2010.

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

[6] Emmanuel J Candés, Justin K Romberg, and Terence Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications on pure and applied mathematics*, 59(8):1207–1223, 2006.

[7] Emmanuel J Candés, Justin K Romberg, and Terence Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications on pure and applied mathematics*, 59(8):1207–1223, 2006.

[8] Emmanuel J Candes and Terence Tao. Decoding by linear programming. *Information Theory, IEEE Transactions on*, 51(12):4203–4215, 2005.

[9] Ingrid Daubechies and Ron DeVore. Approximating a bandlimited function using very coarsely quantized data: A family of stable sigma-delta modulators of arbitrary order. *The Annals of Mathematics*, 158(2):679–710, 2003.

[10] D. L. Donoho. Compressed sensing. *IEEE Trans. Inform. Theory*, 52(4):1289–1306, 2006.

[11] Felix Krahmer, Shahar Mendelson, and Holger Rauhut. Suprema of chaos processes and the restricted isometry property. *Comm. Pure Appl. Math.*, to appear.

[12] C Sinan Güntürk. One-bit sigma-delta quantization with exponential accuracy. *Communications on Pure and Applied Mathematics*, 56(11):1608–1630, 2003.

[13] C Sinan Güntürk, Mark Lammers, Alexander M Powell, Rayan Saab, and Ö Yılmaz. Sobolev duals for random frames and $\sigma\delta$ quantization of compressed sensing measurements. *Foundations of Computational Mathematics*, 13(1):1–36, 2013.

[14] Ankit Gupta, Robert Nowak, and Benjamin Recht. Sample complexity for 1-bit compressed sensing and sparse classification. In *Information Theory Proceedings (ISIT), 2010 IEEE International Symposium on*, pages 1553–1557. IEEE, 2010.

[15] Roger A Horn and Charles R Johnson. *Matrix analysis*. Cambridge university press, 2012.

[16] Hi Inose, Y Yasuda, and J Murakami. A telemetering system by code modulation-Δ-Σ modulation. *Space Electronics and Telemetry, IRE Transactions on*, (3):204–209, 1962.

[17] Laurent Jacques, Jason N Laska, Petros T Boufounos, and Richard G Baraniuk. Robust 1-bit compressive sensing via binary stable embeddings of sparse vectors. 2011.

[18] Felix Krahmer, Rayan Saab, and Rachel Ward. Root-exponential accuracy for coarse quantization of finite frame expansions. *Information Theory, IEEE Transactions on*, 58(2):1069–1079, 2012.

[19] Yaniv Plan and Roman Vershynin. One-bit compressed sensing by linear programming. *Communications on Pure and Applied Mathematics*, 66(8):1163–1333, 2013.

[20] Holger Rauhut. Compressive sensing and structured random matrices. *Theoretical foundations and numerical methods for sparse recovery*, 9:1–92, 2010.

[21] Mark Rudelson and Roman Vershynin. On sparse reconstruction from fourier and gaussian measurements. *Communications on Pure and Applied Mathematics*, 61(8):1025–1045, 2008.

[22] Michel Talagrand. *The generic chaining: upper and lower bounds for stochastic processes*. Springer, 2005.

[23] Andreas M Tillmann and Marc E Pfetsch. The computational complexity of the restricted isometry property, the nullspace property, and related concepts in compressed sensing. *arXiv preprint arXiv:1205.2081*, 2012.