Sigma-Delta quantization of sub-Gaussian compressed sensing measurements

Felix Krahmer
Institute for Numerical and Applied Mathematics
University of Göttingen
Lotzestraße 16-18
37085 Göttingen, Germany
Email: f.krahmer@math.uni-goettingen.de

Rayan Saab
Department of Mathematics
Duke University
Durham, NC 27708, USA
Email: rayans@math.duke.edu

Özgür Yılmaz
Department of Mathematics
University of British Columbia
1984 Mathematics Road
Vancouver, BC V6T 1Z2, Canada
Email: oyilmaz@math.ubc.ca

***Abstract*—Recently, it has been shown that for the setup of compressed sensing with Gaussian measurements that $\Sigma\Delta$ quantization can be effectively incorporated into the sensing mechanism [1]. In contrast to independently quantized measurements, the resulting schemes yield better reconstruction accuracy with a higher number of measurements even at a constant number of bits per signal. The original analysis of this method, however, crucially depends on the rotation invariance of the Gaussian measurements and hence does not directly generalize to other classes of measurements. In this note, we present a refined analysis that allows for a generalization to arbitrary sub-Gaussian measurements.**

I. Introduction

Compressed Sensing [2], [3] is a recent paradigm in signal processing based on the observation that many natural signals are approximately sparse in suitable representation systems, that is, they have only few significant coefficients. The underlying idea is that such signals are intrinsically low-dimensional, so the number of linear measurements necessary to allow for recovery of the signal should be considerably smaller than the signal dimension. Here taking m linear measurements of a signal $x \in \mathbb{R}^N$ is to be understood as considering the measurement vector $y = Ax$, where $A \in \mathbb{R}^{m\times N}$ is a fixed measurement matrix. As it turns out, a number of measurements proportional to $s\log(N/s)$ can allow for stable and robust recovery of signals with s non-vanishing entries in dimension N, provided the measurement matrix is suitably chosen. As no deterministic constructions for such matrices are known, this choice typically involves a random matrix construction.

Note that the resulting linear system to be solved to recover the signal is underdetermined, so the regularizing assumption of sparsity is crucial. However, once it has been determined which s coefficients are significant, the system becomes redundant by at least a logarithmic factor.

In order for the signal to be processed digitally, the measurements must, in a second step, be *quantized*. That is, the measurement vector, whose entries can a priori take arbitrary real values, must be represented by a sequence of values from a given finite alphabet. At this stage, the redundancy mentioned above can be exploited by applying a Sigma-Delta quantization scheme. Such coarse quantization schemes, originally designed for quantizing oversampled bandlimited signals [4], translate redundancy in a signal representation to more accurate quantized representations even though the alphabet size representing each sample is fixed. The idea is that the quantized representations are chosen dynamically using a feedback loop such that the quantization error made in a given sample partly compensates for the error made in previous samples.

This idea has been transferred to the setup of quantizing frame representations in $\mathbb{R}^N$ in [5]. As it turned out in subsequent works, higher accuracy can be achieved if instead of the Moore-Penrose pseudoinverse of the frame matrix, the so called canonical dual frame, an alternative dual frame, the so-called *Sobolev dual* is used for reconstruction [6]. In compressed sensing, once the support is identified, the measurement vector is nothing but a frame representation of the signal, so similar ideas apply.

For measurement matrices with independent standard normal entries, this scenario has been analyzed in [1]. For recovery, the authors proceed via a two-stage approach. In a first stage, standard compressed sensing techniques are used to estimate the support of the signal from the quantized measurements. In a second step, once the support has been identified, a Sobolev dual is used to obtain a more precise estimate of the signal coefficients.

The analysis in [1] is specific to Gaussian measurements. In this note, we generalize their results to arbitrary sub-Gaussian measurements. This more general setup includes important examples like Bernoulli matrices.

II. Sigma-Delta quantization

A. Greedy quantization schemes

Denote by $\mathcal{A}$ the $2L$ level *mid-rise* alphabet

$$\mathcal{A} = \Big\{ \pm (2j+1)\delta/2, \quad j \in \{0, ..., L-1\} \Big\}$$

and let $Q : \mathbb{R} \mapsto \mathcal{A}$ denote the scalar quantizer, which is defined via its action

$$Q(x) = \arg\min_{q\in\mathcal{A}} |x - q|.$$

The rth order *greedy* $\Sigma\Delta$ quantization scheme, defined via

$$q_i = Q\Big(\sum_{j=1}^{r}(-1)^{j-1}\binom{r}{j}u_{i-j} + y_i\Big)$$
$$u_i = \sum_{j=1}^{r}(-1)^{j-1}\binom{r}{j}u_{i-j} + y_i - q_i, \qquad (1)$$

maps a sequence of inputs $(y_i)_{i=1}^m$ to a sequence $(q_i)_{i=1}^m$ whose elements take on values from $\mathcal{A}$. Note that condition (1) can be rewritten as

$$(\Delta^r u)_i = y_i - q_i,$$

where Δ is the finite difference operator.

It is easily seen by induction that for bounded input sequences $\|y\|_\infty < (L - 2^{r-1} - 3/2)$, such schemes satisfy

$$\|u\|_\infty \leq \delta/2.$$

In other words, the scheme is *stable*, that is, its state sequence remains bounded. Note that to satisfy this stability condition, the number of levels L must increase with r.

B. Sigma-Delta error analysis

If $y = Ex \in \mathbb{R}^m$ is a vector of frame coefficients that is $\Sigma\Delta$ quantized to yield the vector $q \in \mathcal{A}^m$, then linear reconstruction of x from q using some dual frame F of E (i.e., $FE = I$) produces the estimate $\hat{x} := Fq$. We would like to control the reconstruction error $\eta := x - \hat{x}$. Writing the state variable equations (II-A) in vector form, we have

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

where D is the $m \times m$ difference matrix with entries given by

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

Thus,

$$\eta = x - Fq = F(y - q) = FD^r u.$$

Working with with stable $\Sigma\Delta$ schemes, one can control $\|u\|_2$ via $\|u\|_\infty$. Thus, it remains to bound the operator norm $\|FD^r\| := \|FD^r\|_{\ell_2^m \mapsto \ell_2^k}$ and a natural choice for F is

$$F := \arg\min_{G:GE=I} \|GD^r\| = (D^{-r}E)^\dagger D^{-r}. \qquad (2)$$

This so-called Sobolev dual frame was first proposed in [6]. Here $A^\dagger := (A^*A)^{-1}A^*$ is the $k \times m$ Moore-Penrose (left) inverse of the $m \times k$ matrix A. Since (2) implies that $FD^r = (D^{-r}E)^\dagger$, the singular values of $D^{-r}E$ will play a key role in this paper.

An important property of the matrix D is given in the following proposition .

Proposition 1 ([1], Proposition 3.1): There are constants c_1, c_2 depending only on r such that the singular values of the matrix D^{-r} satisfy

$$c_1(r)\left(\frac{m}{j}\right)^r \leq \sigma_j(D^{-r}) \leq c_2(r)\left(\frac{m}{j}\right)^r.$$

III. Preliminaries

Here and throughout, $x \sim \mathcal{D}$ denotes that the random variable x is drawn according to a distribution $\mathcal{D}$. Furthermore, $\mathcal{N}(0, \sigma^2)$ denotes the zero-mean Gaussian distribution with variance σ^2. The following definition provides a means to compare the tail decay of two distributions.

Definition 2: If two random variables $\eta \sim \mathcal{D}_1$ and $\xi \sim \mathcal{D}_2$ satisfy $P(|\eta| > t) \leq KP(|\xi| > t)$ for some constant K and all $t \geq 0$, then we say that η is K-dominated by ξ (or, alternatively, by $\mathcal{D}_2$).

Definition 3: A random variable is sub-Gaussian with parameter $c > 0$ if it is e-dominated by $\mathcal{N}(0, c^2)$.

Remark 4: One can also define sub-Gaussian random variables via their moments or, in case of zero mean, their moment generating functions. See [7] for a proof that all these definitions are equivalent.

Remark 5: Exmaples of sub-Gaussian random variables include Gaussian random variables, all bounded random variables (such as Bernoulli), and their linear combinations.

Definition 6: We say that a matrix E is sub-Gaussian with parameter c if its entries are independent sub-Gaussian random variables with mean zero, variance one, and parameter c.

IV. Main results

In this section, we present our main results, generalizing the theorems of [1] on the singular values of $D^{-r}E$ to sub-Gaussian matrix entries and leveraging these results to establish recovery guarantees from $\Sigma\Delta$ quantized compressed sensing measurements.

Proposition 7: Let E be an $m \times k$ sub-Gaussian matrix with parameter c, let $S = \mathrm{diag}(s)$ be a diagonal matrix, and let V be an orthonormal matrix, both of size $m \times m$. Further, let $r \in \mathbb{Z}^+$ and suppose that $s_j \geq C_1^r \left(\frac{m}{j}\right)^r$, where C_1 is a positive constant that may depend on r. Then there exist constants $C_2, C_3 > 0$ (depending on c and C_1) such that for $0 < \alpha < 1$ and $\lambda := \frac{m}{k} \geq C_2^{\frac{1}{1-\alpha}}$

$$\mathbb{P}\left(\sigma_{\min}(\tfrac{1}{\sqrt{m}}SV^*E) \leq \lambda^{\alpha(r-1/2)}\right) \leq 2\exp(-C_3 m^{1-\alpha}k^\alpha).$$

In particular, C_3 depends only on c, while C_2 can be expressed as $f(c)C_1^{-\frac{2r}{2r-1}}$ provided $C_1 \leq 1/2$.

Proof: The matrix SV^*E has dimensions m and k, so by the Courant min-max principle applied to the transpose one has

$$\sigma_{\min}(SV^*E) = \min_{\substack{W \subset \mathbb{R}^m \\ \dim W = m-k+1}} \sup_{z \in W: \|z\|_2 = 1} \|E^*VSz\|_2$$

Noting that, for $m \geq \widetilde{k} := C_4 m^{1-\alpha}k^\alpha > k$, where the constant C_4 will be determined later, each $m - k + 1$-dimensional subspace intersects the span $V_{\widetilde{k}}$ of the first $\widetilde{k}$ standard basis vectors in at least a $\widetilde{k} - k + 1$-dimensional space, this expression is bounded from below by

$$\begin{aligned}
&\min_{\substack{W\subset V_{\widetilde{k}}\\ \dim W=\widetilde{k}-k+1}} \sup_{z\in W:\|z\|_2=1} \|E^*VSz\|_2\\
&\quad\geq \min_{\substack{W\subset V_{\widetilde{k}}\\ \dim W=\widetilde{k}-k+1}} \sup_{z\in W:\|z\|_2=s_{\widetilde{k}}} \|E^*Vz\|_2\\
&\quad= \min_{\substack{W\subset \mathbb{R}^{\widetilde{k}}\\ \dim W=\widetilde{k}-k+1}} \sup_{z\in W:\|z\|_2=1} s_{\widetilde{k}}\|E^*VP_{\widetilde{k}}^*z\|_2.
\end{aligned} \tag{3}$$

The inequality follows from the observation that $V_{\widetilde{k}}$ is invariant under S and the smallest singular value of $S|_{V_{\widetilde{k}}}$ is $s_{\widetilde{k}}$. In the last step, $P_{\widetilde{k}}$ denotes the projection of an m-dimensional vector onto its first $\widetilde{k}$ components. We note that (3), again by the Courant min-max principle, is equal to

$$s_{\widetilde{k}}\sigma_k(E^*VP_{\widetilde{k}}^*) = s_{\widetilde{k}}\sigma_{min}(P_{\widetilde{k}}V^*E) = s_{\widetilde{k}} \inf_{y\in S^{k-1}} \|P_{\widetilde{k}}V^*Ey\|_2$$

Now, as $\mathbb{E}\|P_{\widetilde{k}}V^*Ey\|_2^2=\widetilde{k}$,

$$\begin{aligned}
&\inf_{y\in S^{k-1}} \|P_{\widetilde{k}}V^*Ey\|_2^2\\
&\qquad\geq \Big(\widetilde{k}-\sup_{y\in S^{k-1}} \big|\|P_{\widetilde{k}}V^*Ey\|_2^2-\mathbb{E}\|P_{\widetilde{k}}V^*Ey\|_2^2\big|\Big).
\end{aligned}$$

Thus, noting that

$$\begin{aligned}
\frac{\lambda^{\alpha(r-1/2)}}{s_{\widetilde{k}}} &< m^{\alpha(r-1/2)}k^{-\alpha(r-1/2)}C_1^{-r}m^{-r}\widetilde{k}^r\\
&= C_1^{-r}C_4^{r-\frac{1}{2}}\frac{\sqrt{\widetilde{k}}}{\sqrt{m}}
\end{aligned}$$

and that by choosing $C_4=\min(\frac{1}{2}C_1^{\frac{2r}{2r-1}},\frac{1}{2})$ we ensure that $1-C_1^{-2r}C_4^{2r-1}\geq\frac{1}{2}$,

$$\begin{aligned}
&\mathbb{P}(\sigma_{\min}(\tfrac{1}{\sqrt{m}}SV^*E)\leq\lambda^{\alpha(r-1/2)})\\
&\leq \mathbb{P}(\sup_{y\in S^{k-1}} \big|\|\tfrac{1}{\sqrt{m}}P_{\widetilde{k}}V^*Ey\|_2^2-\mathbb{E}\|\tfrac{1}{\sqrt{m}}P_{\widetilde{k}}V^*Ey\|_2^2\big| \geq \frac{\widetilde{k}}{2m}).
\end{aligned} \tag{4}$$

Note that this choice of C_4 also ensures $\widetilde{k}\leq m$, which is required above. We will estimate (4) using the chaos bounds of [9], similarly to the proof of [9, Thm. A.1]. Indeed, we can write

$$\tfrac{1}{\sqrt{m}}P_{\widetilde{k}}V^*Ey = W_y\xi,$$

where ξ is a vector of length km with independent subgaussian entries of mean zero and variance 1, and

$$W_y=\frac{1}{\sqrt{m}}P_{\widetilde{k}}V^*\begin{pmatrix} y^T & 0 & \cdots & 0\\ 0 & y^T & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots\\ 0 & \cdots & 0 & y^T\end{pmatrix}.$$

In order to apply [9, Thm. 3.1], we need to estimate, for $\mathcal{A}=\{W_y : y\in S^{k-1}\}$, $d_F(\mathcal{A}):=\sup_{A\in\mathcal{A}}\|A\|_F$, $d_{2\to2}(\mathcal{A}):=\sup_{A\in\mathcal{A}}\|A\|_{2\to2}$, and the Talagrand functional $\gamma_2(\mathcal{A},\|\cdot\|_{2\to2})$ (see [8] for its definition). We obtain for $A=W_y\in\mathcal{A}$:

$$\|A\|_F^2=\frac{1}{m}\sum_{j=1}^{k}\sum_{\ell_1,\ell_2=1}^{\widetilde{k},m} y_j^2V_{\ell_1,\ell_2}^2=\frac{\widetilde{k}}{m},\quad\text{so}\quad d_F(\mathcal{A})=\sqrt{\frac{\widetilde{k}}{m}}.$$

Furthermore, we have, for $z\in\mathbb{R}^k$,

$$\begin{aligned}
\|W_z\|_{2\to2} &= \left\|\frac{1}{\sqrt{m}}P_{\widetilde{k}}V^*\begin{pmatrix} z^T & 0 & \cdots & 0\\ 0 & z^T & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots\\ 0 & \cdots & 0 & z^T\end{pmatrix}\right\|_{2\to2}\\
&\leq \left\|\frac{1}{\sqrt{m}}\begin{pmatrix} z^T & 0 & \cdots & 0\\ 0 & z^T & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots\\ 0 & \cdots & 0 & z^T\end{pmatrix}\right\|_{2\to2},
\end{aligned}$$

so the quantities $d_{2\to2}(\mathcal{A})$ and $\gamma_2(\mathcal{A},\|\cdot\|_{2\to2})$ can be estimated in exact analogy to [9, Thm. A.1]. This yields $d_{2\to2}(\mathcal{A})=\frac{1}{\sqrt{m}}$ and $\gamma_2(\mathcal{A},\|\cdot\|_{2\to2})\leq C_5\sqrt{\frac{k}{m}}$ for some constant C_5 depending only on c. With these estimates, we obtain for the quantities E, U, V in [9, Thm. 3.1]

$$\begin{aligned}
E &\leq (2C_5+2)\frac{\sqrt{k\widetilde{k}}}{m}\\
U &\leq \frac{1}{m}\\
V &\leq (C_5+1)\frac{\sqrt{\widetilde{k}}}{m},
\end{aligned}$$

so the resulting tail bound reads

$$\begin{aligned}
&\mathbb{P}\Big(\sup_{y\in S^{k-1}} \big|\|\tfrac{1}{\sqrt{m}}P_{\widetilde{k}}V^*Ey\|_2^2-\mathbb{E}\|\tfrac{1}{\sqrt{m}}P_{\widetilde{k}}V^*Ey\|_2^2\big| \geq\\
&\qquad\qquad c_1(2C_5+2)\frac{\sqrt{k\widetilde{k}}}{m}+t\Big)\\
&\quad\leq e^{-c_2\min\left(\frac{t^2m^2}{(C_5+1)\widetilde{k}},mt\right)}.
\end{aligned}$$

where c_1 and c_2 are the constants depending only on c as they appear in [9, Thm. 3.1]. Note that $k=\widetilde{k}\frac{\lambda^{-(1-\alpha)}}{C_4}$, so for oversampling rates $\lambda>\left((4c_1(2C_5+2))^2/C_4\right)^{\frac{1}{1-\alpha}}=:C_2^{\frac{1}{1-\alpha}}$, we obtain $c_1E\leq\frac{\widetilde{k}}{4m}$ and hence, choosing $t=\frac{\widetilde{k}}{4m}$, we obtain the result

$$\begin{aligned}
&\mathbb{P}(\sup_{y\in S^{k-1}} \big|\|\tfrac{1}{\sqrt{m}}P_{\widetilde{k}}V^*Ey\|_2^2-\mathbb{E}\|\tfrac{1}{\sqrt{m}}P_{\widetilde{k}}V^*Ey\|_2^2\big| \geq \frac{\widetilde{k}}{2m})\\
&\quad\leq e^{-C_3\widetilde{k}}
\end{aligned}$$

where, as desired, the constant $C_3:=\frac{c_2}{16(C_5+1)}$ depends only on the subgaussian parameter c. ■

Analogously to [1], we can use the above bounds to establish guarantees for recovery from $\Sigma\Delta$ quantized compressed sensing measurements.

Theorem 8: Let $\widetilde{\Phi}$ be an $m\times N$ sub-Gaussian matrix with parameter c and set $\Phi:=\frac{1}{\sqrt{m}}\widetilde{\Phi}$, let $r\in\mathbb{Z}^+$, and let $0<\alpha<$

1. Then exist constants C_6, C_7, C_8, C_9 depending only on r and c such that the following holds. Suppose that

$$\lambda := \frac{m}{k} \geq \Big(C_6 \log(eN/k)\Big)^{\frac{1}{1-\alpha}}.$$

Consider the $2L$-level rth order greedy $\Sigma\Delta$ schemes with step-size δ, denote by q the quantization output resulting from Φz where $z \in \mathbb{R}^N$, and denote by Δ a standard compressed sensing decoder. Then with probability exceeding $1 - 2e^{-C_7 m^{1-\alpha} k^{\alpha}}$ for all $z \in \Sigma_k^N$ having $\min_{j \in \mathrm{supp(z)}} |z_j| > C_8 \delta$:

1) the support of z, T, coincides with the support of the best k-term approximation of $\Delta(q)$.
2) denoting by E and F the sub-matrix of Φ corresponding to the support of z and its rth order Sobolev dual respectively, and by $x \in \mathbb{R}^k$ the restriction of z to its support, we have

$$\|x - Fq\|_2 \leq C_9 \lambda^{-\alpha(r-1/2)} \delta.$$

The proof traces the same steps as in [1]. Namely, 1) is a direct consequence of standard RIP-based recovery guarantees and 2) follows from a union bound over all submatrices consisting of k columns of Φ. This union bound determines the condition on λ and the probability. As the all the proof ingredients established above are identical to the corresponding results in [1], we omit the details.

V. Conclusion

Theorem 8 is a complete generalization of the main result of [1] to the scenario of sub-Gaussian matrices. Up to constants, the resulting embedding dimensions are the same as in the Gaussian case.

References

[1] C. Güntürk, M. Lammers, A. Powell, R. Saab, and Ö. Yılmaz, Sobolev duals for random frames and quantization of compressed sensing measurements," *Foundations of Computational Mathematics*, vol. 13, pp. 1–36, 2013.

[2] E. J. Candès, J. Romberg, and T. Tao, “Stable signal recovery from incomplete and inaccurate measurements,” *Comm. Pure Appl. Math.*, vol. 59, pp. 1207–1223, 2006.

[3] D. Donoho, “Compressed sensing.” *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[4] H. Inose and Y. Yasuda, “A unity bit coding method by negative feedback,” *Proceedings of the IEEE*, vol. 51, no. 11, pp. 1524–1535, 1963.

[5] J. Benedetto, A. Powell, and O. Yılmaz, “Sigma-delta ($\Sigma\Delta$) quantization and finite frames,” *IEEE Trans. Inform. Theory*, vol. 52, no. 5, pp. 1990–2005, 2006.

[6] J. Blum, M. Lammers, A. Powell, and O. Yılmaz, “Sobolev duals in frame theory and sigma-delta quantization,” *J. Fourier Anal. and Appl.*, vol. 16, no. 3, pp. 365–381, 2010.

[7] R. Vershynin, “Introduction to the non-asymptotic analysis of random matrices,” in *Compressed Sensing: Theory and Applications*, Y. Eldar and G. Kutyniok, Eds. Cambridge: Cambridge Univ Press, 2012, pp. xii+544.

[8] M. Ledoux and M. Talagrand, *Probability in Banach Spaces: Isoperimetry and Processes*. New York: Springer, 1991.

[9] F. Krahmer, S. Mendelson, and H. Rauhut, Suprema of chaos processes and the restricted isometry property, *Communications on Pure and Applied Mathematics*, to appear.