Error Decay of (almost) Consistent Signal Estimations from Quantized Random Gaussian Projections

Laurent Jacques

June 3, 2014

Abstract

This paper provides new error bounds on consistent reconstruction methods for signals observed from quantized random sensing. Those signal estimation techniques guarantee a perfect matching between the available quantized data and a reobservation of the estimated signal under the same sensing model. Focusing on dithered uniform scalar quantization of resolution \( \delta > 0 \), we prove first that, given a random Gaussian frame of \( \mathbb{R}^N \) with \( M \) vectors, the worst case \( \ell_2 \)-error of consistent signal reconstruction decays with high probability as \( O\left(\frac{N}{M} \log \frac{M}{\sqrt{N}}\right) \) uniformly for all signals of the unit ball \( \mathbb{B}^N \subset \mathbb{R}^N \). Up to a log factor, this matches a known lower bound in \( \Omega\left(\frac{N}{M}\right) \). Equivalently, with a minimal number of frame coefficients behaving like \( M = O\left(\frac{N}{\epsilon_0} \log \frac{\sqrt{N}}{\epsilon_0}\right) \), any vectors in \( \mathbb{B}^N \) with \( M \) identical quantized projections are at most \( \epsilon_0 \) apart with high probability. Second, in the context of Quantized Compressed Sensing with \( M \) random Gaussian measurements and under the same scalar quantization scheme, consistent reconstructions of \( K \)-sparse signals of \( \mathbb{R}^N \) have a worst case error that decreases with high probability as \( O\left(\frac{K}{M} \log \frac{MK}{\sqrt{K}}\right) \) uniformly for all such signals. Finally, we show that the strict consistency condition can be slightly relaxed, e.g., allowing for a bounded level of error in the quantization process, while still guaranteeing a proximity between the original and the estimated signal. In particular, if this quantization error is of order \( O(1) \) with respect to \( M \), similar worst case error decays are reached for reconstruction methods adjusted to such an approximate consistency.

1 Introduction

Since the advent of the digital signal processing era and of analog to digital converters an intense field of research has been concerned by the following non-linear sensing model

\[ q = Q[Ax] \in \mathcal{J}, \tag{1} \]

where \( A \in \mathbb{R}^{M \times N} \) is a matrix representing a linear transformation (or sensing) of a signal \( x \) taken in some bounded subset \( \mathcal{K} \) of \( \mathbb{R}^N \) and \( Q \) stands for a quantization of \( Ax \) mapping \( \mathcal{K} \subset \mathbb{R}^M \) to a finite set of vectors \( \mathcal{J} \subset \mathbb{R}^M \), e.g., using a given number of bits \([7, 16]\).

Since the bounded space \( \mathcal{K} \) contains a priori an infinite number of signals, the model (1) is of course lossy and \( x \) cannot be recovered exactly from \( q \). Quantifying this loss of information in function of both the signal reconstruction method and of the key elements \( A, N, M, \mathcal{K} \) and \( Q \) has therefore been the source of numerous research at the frontier of information theory, high dimensional geometry, signal processing and statistics.

*LJ is with the ICTEAM institute, ELEN Department, Université catholique de Louvain (UCL), Belgium.
The general model (1) is for instance the one adopted in Quantized Compressed Sensing (QCS) [11, 19] when the signal $x$ is assumed sparse or compressible in an orthonormal basis $\Psi$ of $\mathbb{R}^N$ and when the sensing matrix is generated randomly, e.g., from random Gaussian ensembles [10]. When $M \geq N$, Eq. (1) is also a model for frame coefficient quantization (FCQ) of signals in $\mathbb{R}^N$, i.e., when the coefficients $Ax = (a_1^T x, \ldots, a_M^T x)^T$ of $x$ in an overcomplete frame of $\mathbb{R}^N$ are quantized in $q = Q(Ax) \in J$, $A = (a_1, \ldots, a_M)^T$ representing then the matrix whose row set $\{a_j \in \mathbb{R}^N : 1 \leq j \leq M\}$ collects the frame vectors $\Phi$ [14, 13].

In this work, we restrict the analysis of (1) to a scalar, regular and uniform quantizer, i.e., when $Q$ is a scalar operation applied componentwise on $Ax \in \mathbb{R}^M$, and when the 1-D quantization cells defined by the set function $Q^{-1}[c] = \{\lambda : Q[\lambda] = c\} \subset \mathbb{R}$ are convex and have all the same size. We refer the reader for instance to [17] for a review of scalar and $\Sigma\Delta$-quantization [17] in the QCS literature, to [14] for a theoretical analysis of non-regular scalar quantizers, or to [27] for an example of vector quantization by frame permutation.

In particular, we adopt a uniform midrise quantizer
\[
Q_\delta(\lambda) := \delta(\lfloor \frac{1}{2} \rfloor + \frac{1}{2}) \in \delta(\mathbb{Z} + \frac{1}{2}) =: \mathbb{Z}_\delta
\]
of resolution $\delta > 0$ applied componentwise on vectors.

Moreover, we focus on the interplay of $Q_\delta$ with both a random Gaussian matrix $A = \Phi \sim \mathcal{N}(0,1)^{M \times N}$ and a dithering $\xi \sim \mathcal{U}([0,\delta])$, where $\mathcal{D}^{M \times N}(\eta)$ and $\mathcal{D}^M(\eta)$ denote a $M \times N$ random matrix or a $M$-length random vector, respectively, whose entries are identically and independently distributed as the probability distribution $\mathcal{D}(\eta)$ of parameters $\eta = (\eta_1, \ldots, \eta_P)$, e.g., the standard normal distribution $\mathcal{N}(0,1)$ or the uniform distribution $\mathcal{U}([0,\delta])$. Such a dithering is often used for improving the statistical properties of the quantizer by randomizing the unquantized input location inside the quantization cell [10]. As will become clear later (see Sec. 9), this uniform dithering allows us also to bridge our analysis with a geometrical probability context inspired by Buffon’s needle problem [8, 21].

Consequently, given a signal $x$ in a bounded set $K \subset \mathbb{R}^N$, the quantized sensing scenario studied in this paper reads
\[
q = Q_\delta(\Phi x + \xi) \in \mathbb{Z}_\delta^M.
\]
This is either a sensing model for QCS with a random Gaussian sensing $\Phi$, or a quantization scheme for FCQ when the overcomplete frame is made of $M$ vectors in $\mathbb{R}^N$ (with $M \geq N$) that are randomly and independently drawn from $\mathcal{N}(0, \mathbb{I}_{N \times N})$. This guarantees that they are also linearly independent with probability 1, i.e., we obtain a random Gaussian frame (RGF) of $\mathbb{R}^N$.

Our main objective in order to quantify the information loss in (3) while trying to estimate $x$ is to characterize the worst case error
\[
\mathcal{E}_\delta(\Phi, \xi, K) := \max_{x \in K} \|x - x^*\|
\]
of any consistent reconstruction method whose output $x^*$ is determined by the following formal program:
\[
\text{find any } x^* \in \mathbb{R}^N \text{ such that } Q_\delta(\Phi x^* + \xi) = Q_\delta(\Phi x + \xi) \text{ and } x^* \in K.
\]
In the case of FCQ of signals in a RGF (i.e., $M \geq N$), we set $K = \mathbb{B}^N$, while in the context of QCS we take $K = \Sigma_K(\Psi) \cap \mathbb{B}^N$ where $\Sigma_K(\Psi) = \{v \in \Psi \alpha \in \mathbb{R}^N : \|\alpha\|_0 \leq K\}$, with $|\alpha|_0 := \#\{j \in [N] : \alpha_j \neq 0\}$, is the space of $K$-sparse signals in the orthonormal basis $\Psi \in \mathbb{R}^{N \times N}$. For the sake of simplicity, we work with the canonical basis $\Psi = \mathbb{I}_{N \times N}$. However, all our results can be applied to $\Psi \neq \mathbb{I}_{N \times N}$ from the rotational invariance of the random Gaussian matrix $\Phi \sim \mathcal{N}^{M \times N}(0,1)$ in $\mathbb{R}^N$ [11].
We acknowledge the fact that for $\mathcal{K} = \Sigma_K \cap \mathbb{B}^N$ the program \((4)\) is possibly NP hard, \textit{e.g.}, if $x^*$ is found by minimizing the $\ell_0$ “norm” under the consistency constraint \((26)\). However, similarly to the procedure developed in \cite{24}, we are anyway interested in studying its reconstruction error, remembering that similar ideal reconstructions in CS and in QCS have often driven the determination of feasible programs \cite{11,13,19,20,23,28,32}.

As a corollary, the asymptotic decay of $E_\delta$ in function of $M$, $N$, $\delta$ and of the probability $\eta$ can be established.

**Corollary 1.** Given $M \geq 0$, $0 < \eta < 1$, $\delta > 0$ with $\delta = O(1)$, there exists a constant $C > 0$ such that

$$P[E_\delta(\Phi, \eta, \mathbb{B}^N) \leq C \left( \frac{N}{M} \log \frac{M}{\sqrt{N}} + \log \frac{2}{\eta} \right)] \geq 1 - \eta,$$

where the probability is computed with respected to both the RGF $\Phi \sim \mathcal{N}^{M \times N}(0,1)$ and the random dithering $\xi \sim \mathcal{U}^{M}([0,\delta])$.

Loosely speaking, this corollary says that if $\delta = O(1)$, with a probability higher than $1 - \eta$, $E_\delta(\Phi, \xi, \mathbb{B}^N) = O(\frac{N}{M} \log \frac{M}{\sqrt{N}} + \log \frac{2}{\eta}). \quad (5)$

As shown in Sec. \cite{4} it is then straightforward to generalize Theorem \cite{1} to the set of sparse signals.

**Theorem 2.** Let us fix $\epsilon_0 > 0$, $0 < \eta < 1$, $\delta > 0$ and $M$ such that

$$M \geq \left( \frac{4\delta + 2\epsilon_0}{\epsilon_0} \right) \left( 2K \log \left( \frac{56N}{\sqrt{K \epsilon_0}} \right) + \log \frac{1}{2\eta} \right).$$

Let us randomly draw a Gaussian sensing matrix $\Phi \sim \mathcal{N}^{M \times N}(0,1)$ and a dithering $\xi \sim \mathcal{U}^{M}([0,\delta])$. Then, with a probability higher than $1 - \eta$, for all $x \in \mathcal{K} = \Sigma_K \cap \mathbb{B}^N$ sensed by \cite{3}, any solution $x^*$ to \cite{4} is such that $\|x - x^*\| \leq \epsilon_0$, or, equivalently, $E_\delta(\Phi, \xi, \Sigma_K \cap \mathbb{B}^N) \leq \epsilon_0$. 

Our first contribution is an upper bound on this worst case error for consistent signal reconstruction in the context of Random Gaussian Frame Coefficient Quantization (RG-FCQ).

**Theorem 1.** Let us fix $\epsilon_0 > 0$, $0 < \eta < 1$, $\delta > 0$ and $M \geq N$ such that

$$M \geq \frac{4\delta + 2\epsilon_0}{\epsilon_0} \left( N \log \left( \frac{29 \sqrt{N}}{\epsilon_0} \right) + \log \frac{1}{2\eta} \right).$$

Let us randomly draw a RGF $\Phi \sim \mathcal{N}^{M \times N}(0,1)$ and a dithering $\xi \sim \mathcal{U}^{M}([0,\delta])$. Then, with a probability higher than $1 - \eta$, for all $x \in \mathcal{K} = \mathbb{B}^N$ sensed by \cite{3}, any solution $x^*$ to \cite{4} is such that $\|x - x^*\| \leq \epsilon_0$, or, equivalently, $E_\delta(\Phi, \xi, \mathbb{B}^N) \leq \epsilon_0$.
In such a case, as for Corollary 1 in the context of RGF, Sec. 4 also shows that, for any $M \geq 0$,

$$\mathcal{E}_\delta(\Phi, \xi, \Sigma_K \cap \mathbb{B}^N) = O\left( \frac{K}{M} \log\left( \frac{MN}{\sqrt{\delta}} \right) + \log \frac{2}{\eta} \right),$$

with a probability higher than $1 - \eta$.

As explained in Sec. 2, this matches existing error bounds for 1-bit compressed sensing in the case of Gaussian random projections [23]. It also improves previous known bounds decaying as $O(1/\sqrt{M})$ for linear reconstruction methods in FCQ [14] and as $O(\sqrt{K/M})$ for QCS, while a known lower bound in $\Omega(K/M)$ exists [7]. Our result behaves also similarly to the bound on the mean worst case error (mean established with respect to $\Phi$ and $\xi$ in the context of our notations) of consistent reconstruction methods obtained in [30] in the case of random frames over $\mathbb{S}^{N-1}$.

As a last contribution, we show that slight deviations to strict consistency is also possible while keeping similar proximity relations between almost consistent vectors. This occurs for instance if the sensing model [3] is moderately corrupted, i.e., if there exists a noise $\eta \in \mathbb{R}^M$ such that, for any $x \in \mathcal{K}$,

$$\|Q_\delta((\Phi x + \eta) + \xi) - Q_\delta(\Phi x + \xi)\|_1 \leq r \delta.$$

In such a case, we can relax the formal reconstruction (4) and ask to reconstruct $x^*$ from $\mathcal{K}$, find any $x^* \in \mathbb{R}^N$ such that $\|Q_\delta(\Phi x^* + \xi) - Q_\delta(\Phi x + \xi)\|_1 \leq r \delta$ and $x^* \in \mathcal{K}$. \hspace{1cm} (6)

This leads to study a new worst case reconstruction error associated to the definition of relaxed consistency cells $C^r_x := \{x^* \in \mathcal{K} : \|Q_\delta(\Phi x^* + \xi) - Q_\delta(\Phi x + \xi)\|_1 \leq r \delta\}$, i.e.,

$$\mathcal{E}^r_x(\Phi, \xi, \mathcal{K}) := \max \max_{x \in \mathcal{K}} \|x - x^*\|.$$

Sec. 3 proves that, with a probability higher than $1 - \eta$ with respect to the random draw of $\Phi \sim \mathcal{N}^{M \times N}(0, 1)$ and $\xi \sim \mathcal{U}^M([0, \delta])$, the worst case error on almost consistent vectors behaves like

$$\mathcal{E}^r_x(\Phi, \xi, \mathbb{B}^N) = O\left( \frac{N+r}{M} \log\left( \frac{\max(M, N)}{N} \right) + \log\left( \frac{1}{2\eta} \right) \right)$$

for RG-FCQ if $M \geq N$, while for QCS and $K$-sparse signals,

$$\mathcal{E}^r_x(\Phi, \xi, \Sigma_K \cap \mathbb{B}^N) = O\left( \frac{K+r}{M} \log\left( \frac{\max(N, M)}{K} \right) + \log\left( \frac{1}{2\eta} \right) \right).$$

These two results show that, in an almost consistent regime where $r = O(1)$ with respect to $M$, e.g., if the dynamics of the noise components $n_j$ decreases with $j$, the impact of the noise $n$ on the reconstruction of $x$ is controlled and does not change the asymptotic decay of the worst case reconstruction error. However, in the case of stronger noise where $r = O(M)$, the worst case reconstruction error can suffer from a systematic bias in $O(r/M)$. This bias is due to our ignorance of the locations of the corrupted quantized projections in the almost consistent sensing above.

The rest of the paper is structured as follows. In Sec. 2, our main results are discussed under the light of related prior works, i.e., with respect to known upper and lower bounds on reconstruction errors for RG-FCQ and for the QCS setting. Sec. 3 provides the proof of our main results above in the context of RG-FCQ, while Sec. 4 focuses on QCS by extending the analysis to the set of $K$-sparse signals of $\mathbb{R}^N$. Sec. 5 studies the proximity of almost consistent vectors. Our last section, Sec. 6, provides finally the proof of a key Lemma sustaining all our developments. This one is connected to a geometric equivalence between quantization of random projections in $\mathbb{R}^N$ and a variant of Buffon’s needle problem [18] [21].
Conventions: We find useful to present here the rest of the notations used throughout this paper. Domain dimensions are denoted by capital roman letters, e.g., $M, N, \ldots$. Vectors and matrices are associated to bold symbols while lowercase light letters are associated to scalar values, e.g., $\Phi \in \mathbb{R}^{M \times N}$ or $u \in \mathbb{R}^M$. The identity matrix in $\mathbb{R}^D$ reads $I_{D \times D}$. The $i$th component of a vector (or of a vector function) $u$ reads either $u_i$ or $(u)_i$, while the vector $u_i$ may refer to the $i$th element of a set of vectors. The set of indices in $\mathbb{R}^D$ is $[D] = \{1, \ldots, D\}$ and for any $S \subset [D]$ of cardinality $S = \#S$, $u_S \in \mathbb{R}^{\#S}$ denotes the restriction of $u$ to $S$. For materializing this last operation, we also introduce the linear restriction operator $\mathcal{R}_S$ such that $\mathcal{R}_S u = u_S$, i.e., $\mathcal{R}_S = ((1_{M \times M_S})^T$, where $B_S$ denotes the matrix obtained by restricting the columns of $B \in \mathbb{R}^{D \times D}$ to those indexed in $S$. For any $p \geq 1$, the $\ell_p$-norm of $u$ is $\|u\|_p^p = \sum_i |u_i|^p$ with $\|\cdot\| = \|\cdot\|_2$. The $(N-1)$-sphere in $\mathbb{R}^N$ is $S^{N-1} = \{x \in \mathbb{R}^N : \|x\| = 1\}$ while the unit ball is denoted $B^N = \{x \in \mathbb{R}^N : \|x\| \leq 1\}$. More generally, we note $\mathbb{E}_S^N(q) = \{x \in \mathbb{R}^N : \|x - q\| \leq s\}$. For asymptotic relations, we use the common Landau family of notations, i.e., the symbols $O$, $\Omega$ and $\Theta$ [24]. The positive thresholding function is defined by $(\lambda)_+ := \frac{1}{2}(\lambda + |\lambda|)$ for any $\lambda \in \mathbb{R}$.

2 Background and Discussion

In the context of uniform scalar quantization of resolution $\delta > 0$, many works have addressed the model [1] by observing that the distortion induced by quantization compared to a linear model $Ax$ is the one of an additive measurement noise $n = Q(Ax) - Ax$ with $n_i \in [-\delta/2, \delta/2]$, i.e.,

$$q = Ax + n.$$  

(7)

When the resolution is small compared to the dynamics of $Ax$, i.e., under the high resolution assumption [17] [16] [19], or if a random dithering is added to the quantization [16], each component of the noise can be assumed uniformly distributed within $[-\delta/2, \delta/2]$. This allows one to assume $n$ independent of $Ax$ and to bound the power of this noise, i.e., $\mathbb{E}\|n\|^2 = M \delta^2/12$ and $\|n\|^2 \leq \frac{1}{12} \delta^2(M + \sqrt{M})$ with high probability for $\zeta = O(1)$ (see, e.g., [19]).

When a noise $n$ of bounded power $\|n\| \leq \varepsilon$ corrupts the compressive observation of a sparse signal $x$ as in [7], a worst case reconstruction error in

$${\|x - x^*\|} = O(\varepsilon/\sqrt{M}),$$

can be reached by various reconstruction methods (e.g., Basis Pursuit DeNoise [10] [12] or Iterative Hard Thresholding [2]) as soon as the rescaled sensing matrix $\frac{1}{\sqrt{M}} A$ respects the restricted isometry property (RIP) [10].

When the compressive observations of a $K$-sparse signal undergo uniform scalar quantization, it is then expected that, with high probability, $\|x - x^*\| = O(\delta)$ by setting $\varepsilon^2 = \frac{1}{12} \delta^2(M + \zeta \sqrt{M})$. The constancy of this error with respect to $M$ is also known as the classical error limit of the pulse code modulation scheme (PCM) in CS [17].

However, most of these reconstruction techniques enforce a $\ell_2$-norm fidelity with $q$, e.g., by imposing $\|Ax^* - q\| \leq \varepsilon$, and the reconstructed signal is not guaranteed to be consistent with the observations, i.e., $Q(Ax^*) \neq q$. The knowledge of the sensing model is thus not fully exploited for reconstructing $x$ from $q$.

In the context of frame coefficient quantization (FCQ), it is also known that linear reconstruction methods resynthesizing the signal from its frame coefficients corrupted by an additive noise of variance $\sigma^2$ have a root mean square error (RMSE) lower bounded by [15] [30]

$$(\mathbb{E}\|x - x^*\|^2)^{\frac{1}{2}} \geq N \sigma/\sqrt{M},$$

5
where the frame is assumed made of unit norm elements and the expectation is taken with respect to noise.

A better approach for improving the reconstruction error decay in QCS or in FCQ consists in explicitly enforcing consistency between the estimated signal and the quantized data, as formally described by the formal program (4). Such a procedure was initially introduced in [14] in the context of quantized overcomplete expansion of signals. It was also shown there that, given a random model on the generation of the sensed signal, the RMSE of any reconstruction method is lower bounded by $\Omega(N/M)$. Interestingly, the same lower bound can also be obtained on the worst case error reconstruction without requiring a random generation model on the source [7]. While conjectured for general $(M/N)$-redundant frames, the combination of a tight frame formed by an oversampled Discrete Fourier Transform (DFT) with a consistent signal reconstruction reaches this lower bound, i.e., in this case the RMSE is upper bounded by $O(N/M)$ [14].

Consistent reconstruction methods have been applied to QCS in the high resolution regime (i.e., for $\delta \ll 1$ when $K \subset \mathbb{B}_N$) [7, 13, 20, 20], for uniform (or bounded) noise for FCQ [30], in the extreme 1-bit CS setting where quantization reduces to the application of a sign operator [5, 23, 28, 29], or even for non-regular quantization scheme in CS [4, 6].

Recently, Powell and Whitehouse in [30] have analyzed a model equivalent to (3) by adopting a geometric standpoint. In particular, adapting their work to our notations, they have studied the sensing model

$$q = Ax + n,$$

where $A = (a_1^T, \ldots, a_M^T)^T \in \mathbb{R}^{M \times N}$ is a frame whose elements $a_j$ are drawn from a suitable distribution on $S^{N-1}$ and the uniform noise $n \sim U^M([-\delta, \delta])$ stands for, e.g., a dithered uniform scalar quantization of $Ax$. They observe that the consistent reconstruction polytope

$$Q_M := \{ u \in \mathbb{R}^N : \| Au - q \|_\infty \leq \delta \}$$

can be seen a translation of an error polytope $P_N$, i.e., for any consistent reconstruction $x^* \in Q_N$

$$(x^* - x) \in P_M := \{ u \in \mathbb{R}^N : \| Au - n \|_\infty \leq \delta \}.$$ 

Therefore, for a given $A$, analyzing the worst case error of any consistent reconstruction amounts to estimating the width of $P_N$, i.e.,

$$W_M = \max\{\|u\| : u \in P_N\} = \max\{\|u - x\| : u \in Q_N\}.$$ 

Authors in [30] estimate the expected worst case square error $\mathbb{E}|W_M|^2$ with respect to the distribution of the random vectors $\{a_j : 1 \leq j \leq M\}$ on $S^{N-1}$. Relating this estimation to coverage processes on the unit sphere [9], they show that, under general assumption on the distribution of these unit frame vectors,

$$(\mathbb{E}|W_M|^2)^{1/2} \leq \frac{C}{M},$$

with $C > 0$ depending on this distribution. In particular, for $M$ frame vectors uniformly drawn at random over $S^{N-1}$, $C = O(N^{3/2}/M)$ so that

$$(\mathbb{E}|W_M|^2)^{1/2} = O(\frac{N^{3/2}/M}{M}). \quad (8)$$

Despite a slightly different context where the results above focus on an expected worst case analysis, the behavior of these bounds is highly similar to the one we get in Corollary [1] for
consistent reconstruction of signals in the case of RG-FCQ: we observe that, for one random draw of this \((M/N)\)-redundant RGF and of the quantization dithering, \(\mathcal{E}_\delta = O(\frac{N}{M}(\log \frac{M}{N^2} + \log \frac{1}{\delta}))\) with probability higher than \(1 - \eta\).

At first sight the dependence in \(N^{3/2}\) of (8) may seem less optimal than the dependence in \(N\) of (3). However, the first bound is adjusted to random frame vectors drawn uniformly at random over \(\mathbb{S}^{N-1}\), i.e., they have all a unit norm while the RGF vectors have an expected length equal to \(\sqrt{N}\). Keeping in mind the difficulty to compare a bound on the expectation of a random event with a probabilistic bound on this event itself, we can notice, however, that rescaling the result of (8) to uniform random frames over the dilated sphere \(\sqrt{N} \mathbb{S}^{N-1}\), or conversely rescaling \(\delta\) into \(\delta/\sqrt{N}\) in (8), provides an error decay in \((\mathbb{E}[W_M]^2)^{1/2} = O(\frac{N^{\delta}}{M})\).

We can notice that the decay in \((\log M)/M\) of our bound (5) with respect to \(M\) suffers from an extra log factor compared to the decay of \((\mathbb{E}[W_M]^2)^{1/2}\). As will become clear later, this factor actually comes from a union bound argument for upper bounding the probability of failure of our error bounds over all elements of a covering set of \(\mathcal{K}\) (see Sec. 3).

To conclude this section, as pointed out by the known lower bounds described above, let us mention that regular scalar quantization provides a rather limited decay of the reconstruction error, both for FCQ and QCS contexts. Recent developments in vector quantization for FCQ [27], in the use of feedback quantization and of \(\Sigma\Delta\) scheme for FCQ [26] and QCS [7, 17], and finally non-regular quantization schemes where \(Q\) is periodic over its range [4, 6], provide all faster reconstruction error bounds decaying polynomially or even exponentially in \(M\). The implicit objective of this paper is therefore to improve our understanding of one of the simplest quantization scheme, basically a dithered round off operation when combined with random Gaussian projections.

### 3 Quantization of Random Gaussian Frames

This section is dedicated to the proofs of Theorem 1 and of its Corollary 1. Following an argument developed in [4] for non-regular scalar quantization, proving that

\[
\mathcal{E}_\delta(\Phi, \xi, \mathbb{B}^N) = \max_{x \in \mathbb{B}^N} \max_{y \in \mathcal{L}_x} \|x - y\| < \epsilon_0
\]

holds with a probability higher than \(1 - \eta\) on the random draw of a RGF \(\Phi = (\varphi_1, \cdots, \varphi_M)^T \sim \mathcal{N}^{M \times N}(0, 1)\) and of a dithering \(\xi \sim \mathcal{U}^\mathcal{M}([0, \delta])\), amounts to showing that

\[
P[\forall x, y \in \mathbb{B}^N, Q_\delta[\Phi x + \xi] = Q_\delta[\Phi y + \xi] \Rightarrow \|x - y\| \geq \epsilon_0] \geq 1 - \eta,
\]

where \(P\) is computed with respect to the random quantities \(\Phi, \xi\).

Taking the contraposition, we can alternatively demonstrate that,

\[
P_{\text{fail}} := P[\exists x, y \in \mathbb{B}^N, \|x - y\| \geq \epsilon_0 \text{ s.t. } Q_\delta[\Phi x + \xi] = Q_\delta[\Phi y + \xi] < \eta.
\]

For upper bounding \(P_{\text{fail}}\), we take a \(\varsigma\)-covering of the unit ball \(\mathbb{B}^N\), i.e., a finite point set \(\mathcal{L}_s\) such that for any \(v \in \mathbb{B}^N\), there exists a point \(\tilde{v} \in \mathcal{L}_s\) at most \(s\) far apart from \(v\), i.e., \(\|v - \tilde{v}\| \leq s\). The cardinality \(L_s = \#\mathcal{L}_s\) of this covering set is known to be bounded as \(L_s \leq (3/s)^N\) [4].

Therefore, if \(x, y \in \mathbb{B}^N\) are such that \(\|x - y\| \geq \epsilon_0\), taking their respective closest points \(\bar{x}, \bar{y} \in \mathcal{L}_s\), we have \(\|\bar{x} - \bar{y}\| \geq \epsilon_0 - 2s\). Consequently, it is easy to show that

\[
P_{\text{fail}} \leq P[\exists \bar{p}, \bar{q} \in \mathcal{L}_s : \|\bar{p} - \bar{q}\| \geq \epsilon_0 - 2s, \exists \bar{u} \in B_s(p), \exists v \in B_s(p) : Q_\delta[\Phi u + \xi] = Q_\delta[\Phi v + \xi].
\]
Indeed, if the event whose probability is measured by $P_{\text{fail}}$ is verified for $x$ and $y$, taking $p = \bar{x}$, $q = \bar{y}$, $u = x$ and $v = y$ shows that the event associated to the probability of the RHS above occurs.

If one can find an upper bound $P_0$ on
\[
\mathbb{P}[\exists u \in B_s(p), \exists v \in B_s(p), \mathcal{Q}^s_\delta(\Phi u + \xi) = \mathcal{Q}^s_\delta(\Phi v + \xi) | \|\bar{p} - \bar{q}\| > \epsilon_0 - 2s] \leq P_0
\]
that is independent of $p$ and $q$, since the number of pair of points in $L_s$ is bounded by $(L_s)^2 < \frac{1}{2}L_s^2$ independently of any conditions on them, an union bound provides
\[
P_{\text{fail}} \leq \frac{1}{2}L_s^2 P_0.
\]
The following key Lemma allows one to estimate $P_0$.

**Lemma 1.** Let $p, q$ be two points in $\mathbb{R}^N$. There exists a radius $s' \geq \frac{1}{8\sqrt{N}}\|\bar{p} - \bar{q}\|$ such that, for $\Phi \sim \mathcal{N}^{N \times M}(0, 1)$ and $\xi \sim \mathcal{U}^{M \times 1}([0, \delta])$, the probability
\[
P_s'(\alpha, M) := \mathbb{P}[\exists u \in B_s(p), \exists v \in B_{s'}(q), \mathcal{Q}^s_\delta(\Phi u + \xi) = \mathcal{Q}^s_\delta(\Phi v + \xi)]
\]
satisfies
\[
P_s'(\alpha, M) \leq \left(1 - \frac{2\alpha}{8 + 4\alpha}\right)^M,
\]
with $\alpha = \|\bar{p} - \bar{q}\|/\delta$.

As explained in its proof (see Sec. [5]), this Lemma is determined by an equivalence with Buffon’s Needle problem in $N$ dimensions [18], where the needle is actually replaced by a “dumbbell” shape whose two balls are associated to the two neighborhoods of $p$ and $q$.

The quantity $P_s(\alpha, M)$ defined in Lemma 1 increases with $\lambda > 0$. Therefore, for finding an estimation of $P_0$ which is associated to the covering radius $s$, we must guarantee that $s \leq s'$, knowing that $\|\bar{p} - \bar{q}\| \geq \epsilon_0 - 2s$ and $2s' \geq \frac{1}{4\sqrt{N}}\|\bar{p} - \bar{q}\|$. This is achieved by imposing
\[
\frac{1}{4\sqrt{N}}(\epsilon_0 - 2s) = 2s, \text{ i.e.,}
\]
\[
2s = \frac{\epsilon_0}{4\sqrt{N} + 1}.
\]
This provides also $\epsilon_0 - 2s = \frac{4\sqrt{N}}{4\sqrt{N} + 1} \epsilon_0 > \frac{4}{5} \epsilon_0$ if $N \geq 2$.

Consequently, using (10) and observing that $1 - 3\alpha/(4 + 8\alpha)$ decays with $\alpha$,
\[
\mathbb{P}[\exists u \in B_s(p), \exists v \in B_s(p), \mathcal{Q}^s_\delta(\Phi u + \xi) = \mathcal{Q}^s_\delta(\Phi v + \xi) | \|\bar{p} - \bar{q}\| \geq \epsilon_0 - 2s]
\]
\[
= P_s(\alpha, M) \leq P_s'(\alpha, M) \leq \left(1 - \frac{3\alpha}{8 + 4\alpha}\right)^M \leq \left(1 - \frac{2\epsilon_0}{8 + 4\epsilon_0}\right)^M \leq \exp(-\frac{M\epsilon_0}{4\epsilon_0 + 2\epsilon_0}),
\]
where the conditioning of $P_s'$ simply records the extra information without altering its definition.

We can then set $P_0 = \exp(-\frac{M\epsilon_0}{4\epsilon_0 + 2\epsilon_0})$ so that finally
\[
P_{\text{fail}} = \mathbb{P}[\mathcal{Q}^s_\delta(\Phi x + \xi) = \mathcal{Q}^s_\delta(\Phi y + \xi) | \|x - y\| \geq \epsilon_0]
\]
\[
\leq \frac{1}{2}2^N \exp(-\frac{M\epsilon_0}{4\epsilon_0 + 2\epsilon_0})
\]
\[
= \frac{1}{2} \exp(N \log(\frac{2\sqrt{N} + 6}{\epsilon_0} - \frac{M\epsilon_0}{4\epsilon_0 + 2\epsilon_0}))
\]
\[
\leq \frac{1}{2} \exp(N \log(\frac{2\sqrt{N}}{\epsilon_0} - \frac{M\epsilon_0}{4\epsilon_0 + 2\epsilon_0})).
\]
Therefore, if we want $P_{\text{fail}} \leq \eta$ for some $0 < \eta < 1$, it suffices to impose
\[ M \geq \frac{4\delta + 2\epsilon_0}{\epsilon_0} \left(N \log\left(\frac{29\sqrt{N}}{\epsilon_0}\right) + \log \frac{1}{2\eta}\right), \]
which determines the condition invoked in Theorem 1

Knowing that we have necessarily $\epsilon_0 \leq 2$ since $x, y \in \mathbb{B}^N$, a stronger condition for (9) to occur with the same lower bound on its probability reads
\[ M \geq \frac{4(\delta + 1)}{\epsilon_0} \left(N \log\left(\frac{29\sqrt{N}}{\epsilon_0}\right) + \log \frac{1}{2\eta}\right). \quad (11) \]

Alternatively, saturating this condition, we have
\[ \epsilon_0 = \frac{4(\delta + 1)}{M} \left(N \log\left(\frac{29\sqrt{N}}{\epsilon_0}\right) + \log \frac{1}{2\eta}\right) \leq \frac{4(\delta + 1)}{M} \left(N \log\left(\frac{5M}{2\sqrt{N}}\right) + \log \frac{1}{2\eta}\right), \]
where we used the fact that, from (11),
\[ \frac{M}{\sqrt{N}} \geq \frac{4}{\epsilon_0} \sqrt{N} \log\left(\frac{29\sqrt{N}}{\epsilon_0}\right) \geq \frac{4}{\epsilon_0} \log\left(\frac{29}{\sqrt{2}}\right) \sqrt{N} \geq \frac{2}{5} \frac{29}{\epsilon_0} \sqrt{N}, \]
which implies $\epsilon_0 \leq 2$ and assuming $N \geq 2$.

In other words, assuming $\delta = O(1)$, there exists a constant $C > 0$ such that,
\[ \mathbb{P}\left[\mathcal{E}_\delta(\Phi, \xi, \mathbb{B}^N) \leq C\left(\frac{N}{M} \log\left(\frac{M}{\sqrt{N}}\right) + \frac{1}{M} \log \frac{1}{2\eta}\right)\right] \geq 1 - \eta, \]
which proves Corollary 1.

### 4 Extension to $K$-Sparse Vectors of $\mathbb{R}^N$

We study now how the minimal number of measurements evolves in the statement of Theorem 1 when both the original signal and the consistent reconstruction are additionally assumed to be $K$-sparse in $\mathbb{B}^N \subset \mathbb{R}^N$, i.e., they belong to $K = \Sigma_K \cap \mathbb{B}^N$ with $\Sigma_K := \{w \in \mathbb{R}^N : \|w\|_0 \leq K\}$.

Notice first that, given a fixed support $T_0 \subset [N]$ with $\#T_0 = 2K$, thanks to the developments of Sec. 3,
\[ \mathbb{P}\left[\exists x, x^* \in \mathbb{B}^N : \|x - x^*\| \geq \epsilon_0, \, \text{supp} x \cup \text{supp} x^* \subset T_0 \right]: Q_\delta[\Phi x + \xi] = Q_\delta[\Phi x^* + \xi] \]
\[ \leq \frac{1}{2} \exp\left(2K \log\left(\frac{29\sqrt{N}}{\epsilon_0}\right) - \frac{M\epsilon_0}{4\delta + 2\epsilon_0}\right), \]

since the subspace of vectors supported in $T_0$ is equivalent to $\mathbb{R}^{2K}$.

Since there are no more than $\binom{N}{2K} \leq (\frac{eN}{2\kappa})^{2K}$ choices of $2K$-length supports in $[N]$, another union bound provides
\[ \mathbb{P}\left[\exists x, x^* \in \mathbb{B}^N \cap \Sigma_K : \|x - x^*\| \geq \epsilon_0 \right]: Q_\delta[\Phi x + \xi] = Q_\delta[\Phi x^* + \xi] \]
\[ \leq \mathbb{P}\left[\exists T \subset [N] : \#T = 2K, \{\exists x, x^* \in \mathbb{B}^N : \|x - x^*\| \geq \epsilon_0, \, \text{supp} x \cup \text{supp} x^* \subset T \right]: \]
\[ Q_\delta[\Phi x + \xi] = Q_\delta[\Phi x^* + \xi] \]
\[ \leq \frac{1}{2} \left(\frac{N}{2K}\right) \exp\left(2K \log\left(\frac{29\sqrt{N}}{\epsilon_0}\right) - \frac{M\epsilon_0}{4\delta + 2\epsilon_0}\right) \leq \frac{1}{2} \exp\left(2K \log\left(\frac{29\sqrt{N}}{\epsilon_0}\right) - \frac{M\epsilon_0}{4\delta + 2\epsilon_0}\right). \]

Again, willing to have this last probability smaller than $\eta \in (0, 1)$ leads to imposing
\[ M \geq \frac{4\delta + 2\epsilon_0}{\epsilon_0} \left(2K \log\left(\frac{29\sqrt{N}}{\epsilon_0}\right) + \log\left(\frac{1}{2\eta}\right)\right), \]
which, by noting that $29e / \sqrt{2} < 56$, provides the key condition of Theorem 2.

Since $\epsilon_0 \leq 2$, a stronger condition reads
\[ M \geq \frac{4(\delta + 1)}{\epsilon_0} \left( 2K \log \left( \frac{56N}{\sqrt{K\ epsilon_0}} \right) + \log \left( \frac{1}{2\eta} \right) \right), \]
which gives the crude estimation
\[ \frac{MN}{\sqrt{K^4}} \geq \frac{8N}{\epsilon_0 \sqrt{K}} \log \left( \frac{56N}{\sqrt{K\ epsilon_0}} \right) > \frac{8N}{\epsilon_0 \sqrt{K}} \log \left( \frac{56\sqrt{N}}{\epsilon_0} \right) > \frac{1}{2} \frac{56N}{\sqrt{K\ epsilon_0}}, \]
using $K \leq N$ and $N \geq 2$. Therefore, saturating the condition on $M$ above,
\[ \epsilon_0 = \frac{4(\delta + 1)}{M} \left( 2K \log \left( \frac{56N}{\sqrt{K\ epsilon_0}} \right) + \log \left( \frac{1}{2\eta} \right) \right) \leq \frac{4(\delta + 1)}{M} \left( 2K \log \left( \frac{2MN}{\sqrt{K^4}} \right) + \log \left( \frac{1}{2\eta} \right) \right), \]
which shows that, if $\delta = O(1)$, there exists a constant $C > 0$ for which
\[ \mathbb{P}[\mathcal{E}_T(\Phi, \xi, \Sigma_K \cap \mathbb{B}^N) \leq C \left( \frac{M}{K} \log \left( \frac{MN}{\sqrt{K^4}} \right) + \frac{1}{M} \log \left( \frac{1}{2\eta} \right) \right)] \geq 1 - \eta. \]
This determines the bound stated at the end of the Introduction.

5 Proximity of Almost Consistent Signals

The strict consistency between the quantized projections of two vectors of $\mathcal{K} \subset \mathbb{R}^N$ can be slightly relaxed while still keeping their maximal distance bounded. To show this, we follow a similar procedure developed in [22] in the case of 1-bit quantized random projections.

We may first observe that if $||Q_\delta(\Phi x + \xi) - Q_\delta(\Phi y + \xi)||_1 \leq r \delta$ for some $r \in \mathbb{N}$, at most $r$ measurements differ between $Q_\delta(\Phi x + \xi)$ and $Q_\delta(\Phi y + \xi)$. There exists thus a subset $T$ of $[M]$ with size at least $M - r$ such that $\mathcal{R}_T Q_\delta(\Phi x + \xi) = \mathcal{R}_T Q_\delta(\Phi y + \xi)$, with the corresponding restriction operator $\mathcal{R}_T$ defined in the Introduction.

Therefore, for $\mathcal{K} \subset \mathbb{R}^N$ and writing $[M]_r$ the set of all subsets of $[M]$ of size $M - r$, a union bound provides
\[
P_r := \mathbb{P}[\exists x, y \in \mathcal{K} : ||x - y|| \geq \epsilon_0 \text{ s.t. } ||Q_\delta(\Phi x + \xi) - Q_\delta(\Phi y + \xi)||_1 \leq r \delta] \\
\leq \mathbb{P}[\exists T \subset [M]_r, \exists x, y \in \mathcal{K} : ||x - y|| \geq \epsilon_0 \text{ s.t. } \mathcal{R}_T Q_\delta(\Phi x + \xi) = \mathcal{R}_T Q_\delta(\Phi y + \xi)] \\
\leq \bigcup_{T \subset [M]_r} P[\exists x, y \in \mathcal{K} : ||x - y|| \geq \epsilon_0 \text{ s.t. } Q_\delta(\mathcal{R}_T \Phi x + \xi_T) = Q_\delta(\mathcal{R}_T \Phi y + \xi_T)].
\]
We have now to split the analysis between the RG-FCQ and the QCS cases.

**RG-FCQ case:** If $\mathcal{K} = \mathbb{B}^N$ with $M \geq N$, using Sec. 2 and and $\binom{M}{M-r} = \binom{M}{r} \leq (eM/r)^r$,
\[
P_r \leq \frac{1}{2} \binom{M}{M-r} \exp(N \log \left( \frac{29\sqrt{N}}{\epsilon_0} \right) - \frac{1}{2\delta + 2\epsilon_0}) \\
\leq \frac{1}{2} \exp(r \log \left( \frac{eM}{r} \right) + N \log \left( \frac{29\sqrt{N}}{\epsilon_0} \right) - \frac{(M-r)\epsilon_0}{2\delta + 2\epsilon_0}).
\]
Willing to have this last probability smaller than $\eta \in (0, 1)$, we find that, as soon as
\[ M \geq r + \frac{4\delta + 2\epsilon_0}{\epsilon_0} \left( r \log \left( \frac{eM}{r} \right) + N \log \left( \frac{29\sqrt{N}}{\epsilon_0} \right) + \log \left( \frac{1}{2\eta} \right) \right), \]
and given $\Phi \sim N^{M \times N}(0, 1)$ and $\xi \sim U([0, \delta])$, the event
\[ \forall x, y \in \mathcal{K}, \ ||Q_\delta(\Phi x + \xi) - Q_\delta(\Phi y + \xi)||_1 \leq r \delta \Rightarrow ||x - y|| \leq \epsilon_0, \]

Therefore, from (12), if
\[ \delta \eta \]
Again, willing to have this last probability smaller than
\[ \frac{\delta}{2} \log(e M) \]
and given a random draw of
\[ \Phi \]
find that, as soon as
\[ r \]
assuming
\[ r \in \mathbb{N} \]
and forgetting the prime symbol, we find that, as soon as
\[ M \geq \frac{6(\delta + 1)}{\epsilon_0} N \log\left( \frac{29\max(N, M/5)}{\epsilon_0} \right) + \frac{4(\delta + 1)}{\epsilon_0} \log\left( \frac{1}{2^\eta} \right), \]
and given a random draw of
\[ \Phi \sim \mathcal{N}^{M \times N}(0, 1) \]
and
\[ \xi \sim \mathcal{U}^M([0, \delta]) \]
the event
\[ \forall x, y \in \mathcal{K}, \quad \| \mathcal{Q}_\delta(\Phi x + \xi) - \mathcal{Q}_\delta(\Phi y + \xi) \|_1 \leq r \delta \quad \Rightarrow \quad \| x - y \| \leq \frac{N + r}{N} \epsilon_0, \quad (12) \]
holds with probability higher than \( 1 - \eta \).
Saturating the condition on \( M \) above and since \( N \geq 2 \geq \epsilon_0 \), we find
\[ M \geq \frac{6(\delta + 1)}{\epsilon_0} N \log\left( \frac{29\max(N, M/5)}{\epsilon_0} \right) \geq \frac{6}{\epsilon_0} N \log(29) > \frac{2}{3} \frac{29N}{\epsilon_0}, \]
and
\[ \epsilon_0 = \frac{6(\delta + 1)}{M} N \log\left( \frac{29\max(N, M/5)}{\epsilon_0} \right) + \frac{4(\delta + 1)}{M} \log\left( \frac{1}{2^\eta} \right) \]
\[ < \frac{6(\delta + 1)}{M} N \log\left( \frac{3\max(N, M/5)}{2N\epsilon_0} \right) + \frac{4(\delta + 1)}{M} \log\left( \frac{1}{2^\eta} \right). \]
Therefore, from (12), if \( \delta = O(1) \), there is a \( C > 0 \) such that
\[ \mathbb{P}\left[ \mathcal{E}^r_{\delta}(\Phi, \xi, \mathbb{B}^N) \leq \frac{C N + r}{M} \left( \log\left( \frac{M\max(N, M)}{N} \right) + \log\left( \frac{1}{2^\eta} \right) \right) \right] \geq 1 - \eta. \]

**QCS case:** If \( \mathcal{K} = \Sigma_{\mathcal{K}} \cap \mathbb{B}^N \), using the proof of Sec. [4] we have similarly
\[ P_r \leq \frac{1}{2} \left( \frac{M}{M - r} \right) \exp(2K \log(\frac{29eN}{\sqrt{2K\epsilon_0}}) - \frac{(M - r)\epsilon_0}{3e + 2\epsilon_0}) \]
\[ \leq \frac{1}{2} \exp(r \log(\frac{eM}{r}) + 2K \log(\frac{29eN}{\sqrt{2K\epsilon_0}}) - \frac{(M - r)\epsilon_0}{3e + 2\epsilon_0}). \]
Again, willing to have this last probability smaller than \( \eta \in (0, 1) \), we find that, as soon as
\[ M \geq r + \frac{4\delta + 2\epsilon_0}{\epsilon_0} (r \log((M/r)) + 2K \log(\frac{56N}{\sqrt{14K\epsilon_0}}) + \log(\frac{1}{2^\eta})), \]
and given $\Phi \sim \mathcal{N}^{M \times N}(0, 1)$ and $\xi \sim \mathcal{U}^M([0, \delta])$, the event

$$\forall x, y \in \mathcal{K}, \quad \|Q_\delta(\Phi x + \xi) - Q_\delta(\Phi y + \xi)\|_1 \leq r \delta \quad \Rightarrow \quad \|x - y\| \leq \epsilon_0,$$  

holds with probability higher than $1 - \eta$.

Targeting a slightly stronger condition on $M$ by a series of crude upper bounds, we observe that

$$\frac{4\delta + 2\epsilon_0}{\epsilon_0} \left( r(1 + \frac{\epsilon_0}{4\delta + 2\epsilon_0}) \log(\frac{2M}{r}) + 2K \log(\frac{56N}{\sqrt{K} \epsilon_0}) + \log(\frac{1}{2\eta}) \right),$$

$$\leq \frac{4\delta + 2\epsilon_0}{\epsilon_0} \left( 2r \log(eM) + 2K \log(\frac{56N}{\sqrt{K} \epsilon_0}) + \log(\frac{1}{2\eta}) \right);$$

$$\leq \frac{4(\delta + 1)}{\epsilon_0} \left( 2r \log(\frac{2eM}{\epsilon_0}) + 2K \log(\frac{56N}{\sqrt{K} \epsilon_0}) + \log(\frac{1}{2\eta}) \right),$$

$$\leq \frac{4K(\delta + 1)}{(K + r) \epsilon_0} \left( 2r \log(\frac{2eMK}{(K + r) \epsilon_0}) + 2K \log(\frac{56N\sqrt{K}}{(K + r) \epsilon_0}) + \log(\frac{1}{2\eta}) \right);$$

$$\leq \frac{4K(\delta + 1)}{(K + r) \epsilon_0} \left( 2(r + K) \log(\frac{56 \max(N,M/10)}{\epsilon_0}) + \frac{4(\delta + 1)}{\epsilon_0} \log(\frac{1}{2\eta}) \right);$$

$$= \frac{8(\delta + 1)}{\epsilon_0} K \log(\frac{56 \max(N,M/10)}{\epsilon_0}) + \frac{4(\delta + 1)}{\epsilon_0} \log(\frac{1}{2\eta}),$$

using the variable change $\epsilon_0 = \frac{K + r}{K} \epsilon_0$, $\epsilon_0 \leq 2$ and $2\epsilon/56 < 1/10$, and with the same remark on the vanishing value of $r \log(eM/r)$ when $r = 0$.

Therefore, reexpressing everything in function of $\epsilon'_0$ and forgetting the prime symbol, we find that, as soon as

$$M \geq \frac{8(\delta + 1)}{\epsilon_0} K \log(\frac{56 \max(N,M/10)}{\epsilon_0}) + \frac{4(\delta + 1)}{M} \log(\frac{1}{2\eta}),$$

and given a random draw of $\Phi \sim \mathcal{N}^{M \times N}(0, 1)$ and $\xi \sim \mathcal{U}^M([0, \delta])$, the event

$$\forall x, y \in \mathcal{K}, \quad \|Q_\delta(\Phi x + \xi) - Q_\delta(\Phi y + \xi)\|_1 \leq r \delta \quad \Rightarrow \quad \|x - y\| \leq \frac{K + r}{K} \epsilon_0, \quad (13)$$

holds with probability higher than $1 - \eta$.

Saturating the condition on $M$ above, since for this $M$

$$M \geq \frac{8(\delta + 1)}{\epsilon_0} K \log(\frac{56 \max(N,M/10)}{\epsilon_0}) \geq \frac{8}{\epsilon_0} K \log(56) > \frac{1}{2} \frac{56}{\epsilon_0} K,$$

we find

$$\epsilon_0 = \frac{8(\delta + 1)}{M} K \log(\frac{56 \max(N,M/10)}{\epsilon_0}) + \frac{4(\delta + 1)}{M} \log(\frac{1}{2\eta})$$

$$< \frac{8(\delta + 1)}{M} K \log(\frac{2M \max(N,M/10)}{K}) + \frac{4(\delta + 1)}{M} \log(\frac{1}{2\eta}).$$

Consequently, using [13], if $\delta = O(1)$, there exists a $C > 0$ such that

$$\mathbb{P}(C_0 \leq C \mathcal{K} \cap \mathbb{B}^N \leq \mathcal{K} \mathcal{M} \log(\frac{M \max(N,M)}{K}) + \log(\frac{1}{2\eta})] \geq 1 - \eta.$$

### 6 Proof of Lemma [1]

Let us recall the context. We want to show that, given two points $\tilde{p}, \tilde{q} \in \mathbb{R}^N$, there exists a radius $s' \geq \frac{1}{8\sqrt{N}} |\tilde{p} - \tilde{q}|$ such that, for $\Phi \sim \mathcal{N}^{N \times M}(0, 1)$ and $\xi \sim \mathcal{U}^{M \times 1}([0, \delta])$, the probability

$$P_{s'}(\alpha, M) := \mathbb{P}(\exists u \in \mathbb{B}_{s'}(\tilde{p}), \exists v \in \mathbb{B}_{s'}(\tilde{q}), Q_\delta(\Phi u + \xi) = Q_\delta(\Phi v + \xi))$$


satisfies
\[ P_{\delta'}(\alpha, M) \leq \left( 1 - \frac{3\alpha}{8 + 4\alpha} \right)^M, \]
with \( \alpha = \|p - \tilde{q}\|/\delta \).

Notice first that we can focus on upper bounding the probability associated to a single projection by the random vector \( \varphi \sim N^{N+1}(0,1) \) quantized with \( Q_\delta \) with a scalar dithering \( \xi \sim U([0,\delta]) \), the result for \( M \) dithered quantized projections simply following by raising the single measurement bound to the power \( M \), i.e., \( P_{\delta'}(\alpha, M) \leq (P_{\delta'}(\alpha, 1))^M \).

We write \( \varphi = \phi \hat{\varphi} \), where \( \varphi \in \mathbb{S}^{N-1} \) is uniformly distributed at random over \( \mathbb{S}^{N-1} \) and the length \( \phi = \|\varphi\| \sim \chi(N) \) follows a \( \chi \) distribution with \( N \) degrees of freedom. We are going first to estimate the following conditional probability:

\[
P_{\delta'}(\alpha, 1|\phi) := \mathbb{P}\left[ \exists u \in \mathbb{B}_{\delta'}(\tilde{p}), \exists v \in \mathbb{B}_{\delta'}(\tilde{q}), \ Q_\delta[\varphi^T u + \xi] = Q_\delta[\varphi^T v + \xi] \ | \ |\varphi\| = \phi \right]
\]

\[
= \mathbb{P}\left[ \exists u \in \mathbb{B}_{\delta'}(p), \exists v \in \mathbb{B}_{\delta'}(q), \ Q_\delta[\varphi^T u + \xi] = Q_\delta[\varphi^T v + \xi] \ | \ |\varphi\| = \phi \right]
\]

\[
\text{with the variable change } r = \phi s', \ p = \phi \tilde{p} \text{ and } q = \phi \tilde{q} \text{ and } \phi \hat{\varphi} = \varphi. \text{ Notice that } 2r/||p - q|| = 2s'/||p - \tilde{q}||. \text{ Let us focus on this last probability keeping in mind the relationships between these parameters for estimating later a result which is not conditioned to the knowledge of } \phi.
\]

We follow the procedure described in [21]. In this work, from a generalization of the Buffon’s needle problem [8] [18] in \( N \) dimensions, it is shown that when \( r = 0 \), i.e., \( u = p \) and \( v = q \), computing \( P_{\delta'}(\alpha, 1|\phi) \) above is equivalent to estimating the probability that a segment (or needle) of length \( L = |p - q| \) uniformly “thrown” at random in \( \mathbb{R}^N \), both spatially and in orientation, does not intersect a fixed set of parallel \((N - 1)\)-dimensional hyperplanes spaced by a distance \( \delta \).

More precisely, given \( \varphi \in \mathbb{S}^{N-1} \) and \( \xi \in [0,\delta] \), the function \( f(v) := Q_\delta(\varphi^T v + \xi) \) is piecewise constant in \( \mathbb{R}^N \) and the frontiers where its value changes correspond to a set of parallel \((N - 1)\)-dimensional hyperplanes in \( \mathbb{R}^N \). These hyperplanes are equi-spaced with a separating distance \( \delta \) and they are all normal to the direction \( \hat{\varphi} \). Consequently, the quantity \( X := \frac{1}{\delta}(Q_\delta(\varphi^T p + \xi) - Q_\delta(\varphi^T q + \xi)) \in \mathbb{Z} \) counts the number of such hyperplanes intersecting the segment \( \overline{pq} \). In this scenario, this segment is thus fixed and the hyperplanes are randomly oriented and shifted by \( \hat{\varphi} \) and \( \xi \), respectively.

However, we can reverse the point of view and rather consider those hyperplanes as fixed and normal, e.g., to the first canonical axis \( e_1 \) of \( \mathbb{R}^N \). This is allowed by considering the affine mapping \( A_{\varphi,\xi} : \mathbb{R}^N \to \mathbb{R}^N \) implicitly defined by any combination of a rotation and of a translation in \( \mathbb{R}^N \) such that \( e_1^T A_{\varphi,\xi}(v) = \varphi^T v + \xi \) for all \( v \in \mathbb{R}^N \). In words, thanks to \( A_{\varphi,\xi} \), projecting a point \( v \in \mathbb{R}^N \) onto the random orientation \( \hat{\varphi} \) and shifting the result by \( \xi \) is equivalent to projecting the random point \( A_{\varphi,\xi}(v) \) onto \( e_1 \).

Therefore, denoting \( p' = A_{\varphi,\xi}(p) \) and \( q' = A_{\varphi,\xi}(q) \), it is easy to see that the \( L \)-length segment \( \overline{p'q'} \), i.e., our needle, is then oriented uniformly at random over \( \mathbb{S}^{N-1} \) while the distance of its centrum \( \frac{1}{2}(p' + q') \) to the closest hyperplane follows a uniform random variable over the interval \([0,\delta/2] \). Moreover, we have

\[
X = \frac{1}{\delta}(Q_\delta(e_1^T p') - Q_\delta(e_1^T q')), 
\]

so that \( X \) actually measures the number of intersections the segment \( \overline{p'q'} \) makes with the set of hyperplanes \( G_\delta = \bigcup_{k \in \mathbb{Z}} \{ x : e_1^T x = k \} \). In [21], the distribution of the discrete bounded random variable \( X \) is actually fully determined and denoted Buffon\((L/\delta, N)\).
For \( r > 0 \), Eq. (14) shows that we must now consider the two neighboring \( \ell_2 \)-balls of \( p \) and \( q \) in \( P'_s(\alpha, 1|\phi) \) and estimate the probability that at least two points of these balls share the same dithered quantized projection onto \( \varphi \). Following the same argument as above, this new problem is now equivalent to a new Buffon experiment if the previous needle is ended with two balls. In other words, we create a *dumbbell* shape formed by a segment of length \( L \) on the extremities of which two balls of radius \( r \) are centered (see Fig. 1).

It is then easy to see that \( P'_s(\alpha, 1|\phi) \) is equivalent to the probability that there is no hyperplane of \( G_\delta \) intersecting only the part of the segment outside of the two balls when the dumbbell is thrown randomly in \( \mathbb{R}^N \) as for previous Buffon’s needle. Otherwise, having such an intersection would mean that no pair of points (taken in distinct balls) lie in the same subvolume delimited by two consecutive hyperplanes, *i.e.*, they do not have the same quantized projection, and conversely.

Let us parametrize this dumbbell by its distance \( w \sim U([0, \delta/2]) \) (estimated from the middle of the segment) to the closest hyperplane \( G_\delta \) and by its orientation drawn uniformly at random in \( S^{N-1} \). By symmetry, only the angle \( \theta \in [0, \pi] \) made by the dumbbell with the normal vector \( e_1 \) to \( G_\delta \) is important in this parametrization [21]. Moreover, from Fig. 1, the absence of intersection amounts to imposing \( w \geq \frac{1}{2} L | \cos \theta | - r \). The probability \( P'_s(\alpha, 1|\phi) \) is thus obtained by

\[
P'_s(\alpha, 1|\phi) = \frac{4 \kappa_N}{\delta} \int_{\frac{\pi}{2}}^{\pi} (\sin \theta)^{N-2} d\theta \int_0^{\delta/2} \mathbb{I}(w \geq \frac{1}{2} L | \cos \theta | - r) \frac{2}{\delta} dw,
\]

where \( \kappa_N(\sin \theta)^{N-2} d\theta \) is the area (normalized to the one of \( S^{N-1} \)) of the thin spherical segment \( S_{d\theta}(\theta) := \{ \hat{v} \in S^{N-1} : \arccos(e_1^T \hat{v}) \in [\theta, \theta + d\theta] \} \), where \( \kappa_N := \frac{\Gamma(\frac{N}{2})}{\sqrt{\pi} \Gamma(\frac{N-1}{2})} = B(\frac{1}{2}, \frac{N-1}{2})^{-1} \) and \( B(k, l) = \Gamma(k)\Gamma(l)/\Gamma(k + l) \) is the Beta function.

It is important to remark that, from [21] [31],

\[
\frac{\sqrt{2}}{\sqrt{\pi}} (N + 1)^{-\frac{1}{2}} \leq \frac{2 \kappa_N}{\sqrt{\pi}} \leq \frac{\sqrt{2}}{\sqrt{\pi}} (N - 1)^{-\frac{1}{2}},
\]

so that, for \( N \geq 2 \),

\[
\frac{1}{\sqrt{2\pi}} (N + 1)^{\frac{1}{2}} - 1 < \kappa_N \leq \frac{1}{\sqrt{2\pi}} (N - 1)^{\frac{1}{2}} \quad \Rightarrow \quad \kappa_N = \Theta(\sqrt{\frac{N}{2\pi}}).
\]
Let us define two angles \( 0 \leq \theta_0 \leq \theta_1 \leq \pi/2 \) such that \( \cos \theta_0 = \min(\delta + 2r, 1) \) and \( \cos \theta_1 = \frac{2r}{L} \), assuming \( 2r \leq L \) (otherwise, \( P_{\alpha'} = 1 \)). The angular integration domain can be split in three intervals: \([0, \theta_0]\), \([\theta_0, \theta_1]\) and \([\theta_1, \pi/2]\). Over the first interval, the integral is always zero since, either we have a zero measure interval \( (\theta_0 = 0) \) or \( \| (w \geq \frac{L}{2} \cos \theta - r) = 0 \) since \( \frac{L}{2} \cos \theta > \frac{L}{2} \cos \theta_0 = \delta/2 + r \) and \( 0 \leq w \leq \delta/2 \). Moreover, over the last interval \([\theta_1, \pi/2]\), \( I(w \geq \frac{L}{2} \cos \theta - r) = 1 \).

Therefore, writing \( a = L/\delta \),

\[
P_{\alpha'}(\alpha, 1|\phi) = \frac{4\kappa_N}{\delta} \int_{0}^{\theta_1} (\sin \theta)^{N-2}(\frac{L}{2} - \frac{L}{2} \cos \theta + r) d\theta + 2\kappa_N \int_{0}^{\theta_1} (\sin \theta)^{N-2} d\theta
\]

\[
= 1 + \frac{4\kappa_N}{\delta} \int_{0}^{\theta_1} (\sin \theta)^{N-2}(\frac{L}{2} - \frac{L}{2} \cos \theta + r) d\theta - 2\kappa_N \int_{0}^{\theta_1} (\sin \theta)^{N-2} d\theta
\]

\[
= 1 - \frac{4\kappa_N}{\delta} \int_{0}^{\theta_1} (\sin \theta)^{N-2}(\frac{L}{2} \cos \theta - r) d\theta - \frac{4\kappa_N}{\delta} \int_{0}^{\theta_1} (\sin \theta)^{N-2} \frac{\delta}{2} d\theta
\]

\[
= 1 - 2\kappa_N a \int_{0}^{1} (1 - v^2)^{\frac{N-3}{2}} \left[ (v - \frac{2r}{L})_+ - (v - \frac{2r+\delta}{L})_+ \right] dv,
\]

applying a variable change \( v = \cos \theta \) on the last line.

Let us study this last integral and the function \( f(v) = (v - \frac{2r}{L})_+ - (v - \frac{2r+\delta}{L})_+ \). We can verify that \( F(v) := \int_{0}^{v} f(v') dv' \) is convex and reads

\[
2F(v) = (v - \frac{2r}{L})_+^2 - (v - \frac{2r+\delta}{L})_+^2 = \begin{cases} 
0, & \text{if } v \leq \frac{2r}{L}, \\
(v - \frac{2r}{L})_+^2, & \text{if } \frac{2r}{L} < v \leq \frac{2r+\delta}{L}, \\
\frac{\delta}{L}(2v - \frac{4r+\delta}{L}), & \text{if } v > \frac{2r+\delta}{L}.
\end{cases}
\]

Moreover, by integrating by part,

\[
\int_{0}^{1} (1 - v^2)^{\frac{N-3}{2}} f(v) dv = \int_{0}^{1} (N-3) v(1 - v^2)^{\frac{N-5}{2}} F(v) dv,
\]

The positive measure \( \mu(v) = (N-3) v(1 - v^2)^{\frac{N-5}{2}} \) has unit mass over \([0, 1]\) so that, by convexity of \( F \) and using the Jensen inequality,

\[
\int_{0}^{1} F(v) \mu(v) dv \geq F\left( \int_{0}^{1} v \mu(v) dv \right).
\]

However, since

\[
(N-3) \int_{0}^{1} (1 - v^2)^{\frac{N-5}{2}} v^q dv = \frac{N-3}{2} B\left(\frac{q+1}{2}, \frac{N-3}{2}\right) = \frac{\Gamma\left(\frac{q+1}{2}\right) \Gamma\left(\frac{N-3}{2}\right)}{\Gamma\left(\frac{N+q-2}{2}\right)},
\]

we find

\[
\int_{0}^{1} v \mu(v) dv = (N-3) \int_{0}^{1} (1 - v^2)^{\frac{N-5}{2}} v^2 dv = \frac{\sqrt{\pi} \Gamma\left(\frac{N-1}{2}\right)}{2 \Gamma\left(\frac{N}{2}\right)} = \frac{1}{2\kappa_N}.
\]

and

\[
P_{\alpha'}(\alpha, 1|\phi) \leq 1 - 2\kappa_N a F\left(\frac{1}{2\kappa_N}\right).
\]
From the definition of $F$ above, if $\frac{1}{2\kappa_N} \leq \frac{2r}{L}$, $F = 0$ and we cannot show anything. Let us thus set $2r = \frac{\lambda}{2\kappa_N} L$, where $\lambda \in (0, 1)$ will be determined later. Notice that, since $s' = \phi r$ and $\|\tilde{p} - \tilde{q}\| = \phi L$, we implicitly impose $2r/L = 2s'/\|\tilde{p} - \tilde{q}\| = \frac{\lambda}{2\kappa_N}$.

Then,

$$2F\left(\frac{1}{2\kappa_N}\right) = \begin{cases} \frac{1}{4\kappa_N}(1 - \lambda)^2, & \text{if } a \leq \frac{2\kappa_N}{1 - \lambda}, \\ \frac{1}{4\kappa_N}(1 - \lambda) - \frac{1}{a}, & \text{if } a \geq \frac{2\kappa_N}{1 - \lambda}, \end{cases}$$

so that, writing $\phi_0 = \frac{2\kappa_N}{1 - \lambda}$, we have

$$P_{\phi'}(\alpha, 1|\phi) \leq \begin{cases} 1 - \frac{\alpha}{\phi_0}(1 - \lambda), & \text{if } a \leq \phi_0, \\ \lambda + \frac{\phi}{\phi_0}(1 - \lambda), & \text{if } a \geq \phi_0. \end{cases}$$

Let us recall that $P_{\phi'}(\alpha, 1|\phi)$ is defined conditionally to $\phi = \|\varphi\|$ with $\phi \sim \chi(N)$. Moreover, $a = \|\tilde{p} - \tilde{q}\|/\delta = \phi\|\tilde{p} - \tilde{q}\|/\delta = \alpha \phi$ with $\alpha = \|\tilde{p} - \tilde{q}\|/\delta$. Denoting the pdf of $\chi(N)$ by $\gamma_N(\phi) = c_N \phi^{N-1} \exp(-\phi^2/2)$ and $c_N = 2^{1 - \frac{N}{2}}/\Gamma(\frac{N}{2})$, we can develop $P_{\phi'}(\alpha, 1) = \int_0^{+\infty} P_{\phi'}(\alpha, 1|\phi) \gamma_N(\phi) \, d\phi$ as follows

$$P_{\phi'}(\alpha, 1) \leq \int_0^{\phi_0/\alpha} (1 - \frac{\phi}{\phi_0}(1 - \lambda)) \gamma_N(\phi) \, d\phi + \int_{\phi_0/\alpha}^{+\infty} (\lambda + \frac{\phi}{\phi_0}(1 - \lambda)) \gamma_N(\phi) \, d\phi$$

$$= \lambda + (1 - \lambda) \int_0^{\phi_0/\alpha} (1 - \frac{\phi}{\phi_0}) \gamma_N(\phi) \, d\phi + (1 - \lambda) \int_{\phi_0/\alpha}^{+\infty} \frac{\phi}{\phi_0} \gamma_N(\phi) \, d\phi$$

$$= \lambda + (1 - \lambda) \int_0^{+\infty} \phi(\alpha \phi) \gamma_N(\phi) \, d\phi,$$

with $\varphi(t) = 1 - \frac{t}{\lambda}$ if $0 \leq t < 1$ and $\varphi(t) = \frac{1}{2t}$ if $t \geq 1$.

We can notice that $t \varphi(t)$, which is equal to $\frac{1}{4} t (2 - t)$ over $[0, 1]$ and to $\frac{1}{2}$ for $t \geq 1$, is a concave function. Therefore, by Jensen inequality,

$$\int_0^{+\infty} \varphi(\alpha \phi) \gamma_N(\phi) \, d\phi = \frac{c_N}{c_N - 1} \int_0^{+\infty} \varphi(\phi) \gamma_{1-N}(\phi) \, d\phi$$

$$\leq \frac{c_N}{c_N - 1} (E_{N-1} \phi) \varphi(\frac{\alpha E_{N-1} \phi}{\phi_0})$$

We have also $c_N/c_{N-1} = \Gamma(\frac{N-1}{2})/(\sqrt{2\pi} \Gamma(\frac{N}{2}))$ and $E_{N-1} \phi = \sqrt{2\pi}(\frac{N}{2})/\Gamma(\frac{N-1}{2}) = \sqrt{2\pi} \kappa_N$, so that $\frac{c_N}{c_{N-1}} (E_{N-1} \phi) = 1$ and

$$\frac{\alpha E_{N-1} \phi}{\phi_0} = \alpha \frac{1 - \lambda}{2\kappa_N} E_{N-1} \phi = \sqrt{\frac{2}{\pi}} (1 - \lambda) \alpha.$$

Consequently, since $\varphi(t) \leq \frac{2}{\pi} t$,

$$P_{\phi'}(\alpha, 1) \leq \lambda + (1 - \lambda) \varphi\left(\frac{\sqrt{2}}{\pi} (1 - \lambda) \alpha\right)$$

$$\leq \lambda + (1 - \lambda) \frac{2}{2 + \sqrt{2} (1 - \lambda) \alpha}$$

$$= 1 - \frac{\sqrt{2} (1 - \lambda) \alpha}{2 + \sqrt{2} (1 - \lambda) \alpha} = 1 - \frac{\sqrt{2} \alpha}{2 + \alpha} < 1 - \frac{3\alpha}{8 + 4\alpha}$$

taking $(1 - \lambda) = \sqrt{\frac{2}{\pi}} > 3/4$.  

16
Moreover, from the bounds on $\kappa_N$ given in [15], this shows also that
\[
\frac{2s}{\|p-q\|} = \frac{2r}{L} = \frac{\lambda}{2\kappa_N} \geq (1 - \sqrt{\frac{2}{\pi}})\sqrt{\frac{\pi}{2}} \frac{1}{(N-1)^{1/2}} > \frac{1}{4\sqrt{N}},
\]
as stated at the beginning of Lemma [1].

Acknowledgements

Laurent Jacques is a Research Associate funded by the Belgian F.R.S.-FNRS.

References

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

[2] T. Blumensath and M. Davies. Iterative hard thresholding for compressive sensing. *Appl. Comput. Harmon. Anal.*, 27(3):265–274, 2009.

[3] P. T. Boufounos. *Quantization and Erasures in Frame Representations*. D.Sc. Thesis, MIT EECS, Cambridge, MA, January 2006.

[4] P. T. Boufounos. Universal rate-efficient scalar quantization. *IEEE Trans. Info. Theory*, 58(3):1861–1872, March 2012.

[5] P. T. Boufounos and R. G. Baraniuk. 1-bit compressive sensing. In *Proc. Conf. Inform. Science and Systems (CISS)*, Princeton, NJ, March 19-21 2008.

[6] P. T. Boufounos and S. Rane. Efficient coding of signal distances using universal quantized embeddings. In *Proc. Data Compression Conference (DCC)*, Snowbird, UT, March 20-22 2013.

[7] P. T. Boufounos, L. Jacques, F. Krahmer, and R. Saab. Quantization and compressive sensing. In *book "Compressed Sensing and its Applications"*, Springer, 2014.

[8] G. C. Buffon. *Essai d’arithmétique morale*, volume 4 of *Supplément à l’histoire naturelle*, 1777. [http://www.buffon.cnrs.fr](http://www.buffon.cnrs.fr)

[9] P. Bürgisser, F. Cucker, M. Lotz, et al. Coverage processes on spheres and condition numbers for linear programming. *The Annals of Probability*, 38(2):570–604, 2010.

[10] E. Candès, J. Romberg, and T. Tao. Stable signal recovery from incomplete and inaccurate measurements. *Comm. Pure Appl. Math*, 59(8):1207–1223, 2006.

[11] E. J. Candès and T. Tao. Near-optimal signal recovery from random projections: Universal encoding strategies? *Information Theory, IEEE Transactions on*, 52(12):5406–5425, 2006.

[12] S. S. Chen, D. L. Donoho, and M. A. Saunders. Atomic Decomposition by Basis Pursuit. *SIAM Journal on Scientific Computing*, 20(1):33–61, 1998.

[13] W. Dai, H. V. Pham, and O. Milenkovic. Distortion-Rate Functions for Quantized Compressive Sensing. Technical Report arXiv:0901.0749, 2009.

[14] V. K Goyal, M. Vetterli, and N. T. Thao. Quantized overcomplete expansions in $\mathbb{R}^N$: Analysis, synthesis, and algorithms. *IEEE Trans. Info. Theory*, 44(1):16–31, 1998.

[15] V. K. Goyal, J. Kovacević, and J. A. Kelner. Quantized frame expansions with erasures. *Applied and Computational Harmonic Analysis*, 10(3):203–233, 2001.

[16] R.M. Gray and D.L. Neuhoff. Quantization. *IEEE Trans. Info. Theory*, 44(6):2325–2383, Oct 1998.
[17] C. S. Güntürk, M. Lammers, A. M. Powell, R. Saab, and Ö. Yılmaz. Sobolev duals for random frames and σδ quantization of compressed sensing measurements. *Foundations of Computational Mathematics*, 13(1):1–36, 2013.

[18] J. D. Hey, T. M. Neugebauer, and C. M. Pasca. Georges-Louis Leclerc de Buffon's Essays on Moral Arithmetic. In *The Selten School of Behavioral Economics*, pages 245–282. Springer, 2010.

[19] L. Jacques, D. K. Hammond, and M. J. Fadili. Dequantizing Compressed Sensing: When Oversampling and Non-Gaussian Constraints Combine. *IEEE Trans. Inf. Theory*, 57(1):559–571, January 2011.

[20] L. Jacques, D. K. Hammond, and M. J. Fadili. Stabilizing Nonuniformly Quantized Compressed Sensing With Scalar Companders. *IEEE Trans. Inf. Theory*, 59(12):7969 – 7984, January 2013.

[21] L. Jacques. A Quantized Johnson Lindenstrauss Lemma: The Finding of Buffon’s Needle. *arXiv preprint arXiv:1309.1507*, 2013.

[22] L. Jacques, K. Degraux, and C. De Vleeschouwer. Quantized iterative hard thresholding: Bridging 1-bit and high-resolution quantized compressed sensing. In *10th international conference on Sampling Theory and Applications (SampTA 2013)*, *arXiv:1305.1786*, pages 105–108, Bremen, Germany, 2013.

[23] L. Jacques, J. N. Laska, P. T. Boufounos, and R. G. Baraniuk. Robust 1-bit compressive sensing via binary stable embeddings of sparse vectors. *IEEE Transactions on Information Theory*, 59(4):2082–2102, 2013.

[24] D. E. Knuth. Big omicron and big omega and big theta. *ACM Sigact News*, 8(2):18–24, 1976.

[25] F. Krahmer, R. Saab, and R. Ward. Root-exponential accuracy for coarse quantization of finite frame expansions. *IEEE Trans. Inform. Theory*, 58(2):1069–1079, 2012.

[26] B. K. Natarajan. Sparse approximate solutions to linear systems. *SIAM journal on computing*, 24(2):227–234, 1995.

[27] H. Q. Nguyen, V.K. Goyal, and L.R. Varshney. Frame Permutation Quantization. *Applied and Computational Harmonic Analysis (ACHA)*, Nov. 2010.

[28] Y. Plan and R. Vershynin. Dimension reduction by random hyperplane tessellations. *arXiv preprint arXiv:1111.4452*, 2011.

[29] Y. Plan and R. Vershynin. One-bit compressed sensing by linear programming. *Communications on Pure and Applied Mathematics*, 66(8):1275–1297, 2013.

[30] A. M. Powell and J. T. Whitehouse. Error bounds for consistent reconstruction: Random polytopes and coverage processes. *arXiv preprint arXiv:1405.7094*, 2014.

[31] F. Qi, Q.-M. Luo, et al. Bounds for the ratio of two gamma functions from Wendels and related inequalities to logarithmically completely monotonic functions. *Banach J. Math. Anal.*, 6(2):132–158, 2012.

[32] J. A. Tropp. Just relax: Convex programming methods for identifying sparse signals in noise. *Information Theory, IEEE Transactions on*, 52(3):1030–1051, 2006.