Robust One-bit Compressed Sensing With Manifold Data

Sjoerd Dirksen
Utrecht University
s.dirksen@uu.nl

Mark Iwen
Michigan State University
markiwen@math.msu.edu

Sara Krause-Solberg
Technical University of Munich
sara.krause-solberg@ma.tum.de

Johannes Maly
RWTH Aachen University
maly@mathc.rwth-aachen.de

***Abstract*—We study one-bit compressed sensing for signals on a low-dimensional manifold. We introduce two computationally efficient reconstruction algorithms that only require access to a geometric multi-resolution analysis approximation of the manifold. We derive rigorous reconstruction guarantees for these methods in the scenario that the measurements are subgaussian and show that they are robust with respect to both pre- and post-quantization noise. Our results substantially improve upon earlier work in this direction.**

I. Introduction

We consider a union $\mathcal{M}$ of low-dimensional C^1-manifolds of dimension d in the Euclidean ball $\mathcal{B}(\mathbf{0}, R)$ in a high-dimensional space $\mathbb{R}^D$, $d \ll D$. We imagine that we do not know $\mathcal{M}$ perfectly, and instead only have access to a certain structured approximation for $\mathcal{M}$, called a Geometric Multi Resolution Analysis (GMRA) approximation [1]. Given this data, our goal is to recover an unknown signal $\mathbf{x} \in \mathcal{M}$ from m memoryless one-bit quantized measurements

$$\mathbf{y} = \operatorname{sign}(\mathbf{A}\mathbf{x} + \boldsymbol{\nu} + \boldsymbol{\tau}), \tag{1}$$

where $\mathbf{A} \in \mathbb{R}^{m\times D}$ is the measurement matrix, $\boldsymbol{\nu} \in \mathbb{R}^m$ is a noise vector, and $\boldsymbol{\tau} \in \mathbb{R}^m$ is a vector of quantization thresholds. We wish to recover $\mathbf{x}$ accurately using as few measurements as possible, using a computationally efficient algorithm. The recent work [2] introduced a recovery algorithm for this problem in the setting where $\mathbf{A}$ is standard Gaussian, $\boldsymbol{\nu} = \mathbf{0}$, $\boldsymbol{\tau} = \mathbf{0}$ and $\mathcal{M} \subset \mathbb{S}^{D-1}$. The purpose of this note is to show that if one uses dithering in the quantizer, i.e., uses a suitably chosen random threshold vector $\boldsymbol{\tau}$, then superior results can be obtained. We introduce two simple, computationally efficient reconstruction algorithms and derive recovery guarantees for subgaussian matrices and manifolds located in $\mathcal{B}(\mathbf{0}, R)$. Both methods are robust to noise before and during quantization. In addition, our bounds on the required number of measurements for accurate signal recovery exhibit better parameter dependencies than [2].

II. Preliminaries

Let us first fix some notation and terminology that will be used throughout our presentation. We let $\mathcal{B}(\mathbf{z}, r)$ denote the Euclidean ball in $\mathbb{R}^D$ with center $\mathbf{z}$ and radius r. Throughout $\mathcal{M}$ denotes a union of finitely many d-dimensional C^1-manifolds in $\mathbb{R}^D$. We let

$$\operatorname{tube}_r(\mathcal{M}) := \{\mathbf{x} \in \mathbb{R}^D \ : \ \inf_{\mathbf{y}\in\mathcal{M}} \|\mathbf{x} - \mathbf{y}\|_2 \le r\}$$

denote the closed r-neighbourhood (or tube) around $\mathcal{M}$. For any $T \subset \mathbb{R}^n$ we let $|T|$ denote its cardinality. We use $\operatorname{star}(T) = \{\rho\mathbf{x} \ : \ \rho \in [0,1], \ \mathbf{x} \in T\}$ to denote the star-shaped hull of T. A random vector $X \in \mathbb{R}^n$ is called L-subgaussian if $\|\langle X, \mathbf{x}\rangle\|_{L_p} \le L\sqrt{p}\|\langle X, \mathbf{x}\rangle\|_{L_2}$ for all $\mathbf{x} \in \mathbb{R}^n$, and $1 \le p < \infty$. We use $d_H(\mathbf{z}, \mathbf{z}') = |\{i \in [n] \ : \ z_i \neq z_i'\}|$ to denote the Hamming distance of two bit strings $\mathbf{z}, \mathbf{z}' \in \{-1,1\}^n$. Let $\mathbb{P}_S(\mathbf{x})$ denote the Euclidean projection of $\mathbf{x}$ onto a given convex set $S \subset \mathbb{R}^D$. We use c_α, C_α to denote positive constants depending only on α, which may change from line to line. Finally, we write $a \lesssim_\alpha b$ if $a \le C_\alpha b$ and $a \simeq_\alpha b$ if both $a \lesssim_\alpha b$ and $b \lesssim_\alpha a$ hold.

To start our development, let us recall the definition of a GMRA approximation of $\mathcal{M}$.

Definition II.1 (GMRA Approximation to $\mathcal{M}$ [1], [3])**.** *Let $J \in \mathbb{N}$ and $K_1, ..., K_J \in \mathbb{N}$. Then a* Geometric Multi Resolution Analysis (GMRA) *approximation of $\mathcal{M}$ is a collection $\{(\mathcal{C}_j, \mathcal{P}_j)\}$, $j \in [J] := \{1, ..., J\}$, of sets $\mathcal{C}_j = \{\mathbf{c}_{j,k}\}_{k=1}^{K_j} \subset \mathbb{R}^D$ of centers and*

$$\mathcal{P}_j = \left\{\mathbb{P}_{j,k} : \mathbb{R}^D \to \mathbb{R}^D \ \middle| \ k \in [K_j]\right\}$$

of affine projectors which approximate $\mathcal{M}$ at scale j, such that the following assumptions hold.

(1) ***Affine Projections:*** *Every $\mathbb{P}_{j,k} \in \mathcal{P}_j$ has an associated center $\mathbf{c}_{j,k} \in \mathcal{C}_j$ and a matrix $\Phi_{j,k} \in \mathbb{R}^{d\times D}$ satisfying $\Phi_{j,k}\Phi_{j,k}^T = \mathrm{Id}_d$, such that*

$$\mathbb{P}_{j,k}(\mathbf{z}) = \Phi_{j,k}^T\Phi_{j,k}(\mathbf{z} - \mathbf{c}_{j,k}) + \mathbf{c}_{j,k},$$

i.e., $\mathbb{P}_{j,k}$ is the projector $\mathbb{P}_{P_{j,k}}$ onto a certain affine d-dimensional linear subspace $P_{j,k}$ containing $\mathbf{c}_{j,k}$.

(2) ***Centers:*** *The number of centers is $|\mathcal{C}_j| = K_j$ and $K_j \le C2^{dj}$ for an absolute constant $C \ge 1$.*

(3) ***Multiscale Approximation:*** *There exist absolute constants $C_1, C_2 > 0$ such that*

(a) There exists a $j_0 \in [J-1]$, such that $\mathbf{c}_{j,k} \in \operatorname{tube}_{C_1\cdot 2^{-j-2}}(\mathcal{M})$, for all $j > j_0 \ge 1$ and $k \in [K_j]$.

(b) *For each $j \in [J]$ and $\mathbf{z} \in \mathbb{R}^D$ let $\mathbf{c}_{j,k_j(\mathbf{z})}$ be one of the centers closest to $\mathbf{z}$, i.e.,*

$$k_j(\mathbf{z}) \in \arg\min_{k \in [K_j]} \|\mathbf{z} - \mathbf{c}_{j,k}\|_2. \tag{2}$$

Then, for each $\mathbf{z} \in \mathcal{M}$ there exists a constant $C_\mathbf{z} > 0$ such that, for all $j \in [J]$,

$$\|\mathbf{z} - \mathbb{P}_{j,k_j(\mathbf{z})}(\mathbf{z})\|_2 \le C_\mathbf{z} \cdot 2^{-2j}.$$

(c) *For each $\mathbf{z} \in \mathcal{M}$ there exists $\tilde{C}_\mathbf{z} > 0$ such that*

$$\|\mathbf{z} - \mathbb{P}_{j,k'}(\mathbf{z})\|_2 \le \tilde{C}_\mathbf{z} \cdot 2^{-j/2}, \tag{3}$$

for all $j \in [J]$ and $k' \in [K_j]$ satisfying

$$\begin{aligned} \|\mathbf{z} - \mathbf{c}_{j,k'}\|_2 &\le C_2 \sqrt{j + \log(R)} \\ &\cdot \max\left\{\|\mathbf{z} - \mathbf{c}_{j,k_j(\mathbf{z})}\|_2, C_1 \cdot 2^{-j-1}\right\}. \end{aligned} \tag{4}$$

Remark II.2. *Definition II.1 differs slightly from the original GMRA axioms as proposed in [3] in two respects. First, in [3] it is additionally assumed that the centers are well-separated and organized in a tree-like structure. Second, instead of* (3c) *it is assumed that if $\|\mathbf{z} - \mathbf{c}_{j,k'}\|_2 \le C_2 \cdot \max\left\{\|\mathbf{z} - \mathbf{c}_{j,k_j(\mathbf{z})}\|_2, C_1 \cdot 2^{-j-1}\right\}$ (a stronger requirement), then $\|\mathbf{z} - \mathbb{P}_{j,k'}(\mathbf{z})\|_2 \le \tilde{C}_\mathbf{z} \cdot 2^{-j}$ (a stronger property).*

For each level j, the union of manifolds $\mathcal{M}$ is approximated by a union of d-dimensional affine subspaces $P_{j,k}$ described by the centers $\mathbf{c}_{j,k}$ and the matrices $\Phi_{j,k}$. The centers are not required to lie on $\mathcal{M}$ but their distance to $\mathcal{M}$ is controlled by property (3a). If we let j_* be the smallest integer exceeding j_0 so that $\mathrm{tube}_{C_1 2^{-j-2}}(\mathcal{M}) \subset \mathcal{B}(\mathbf{0}, 2R)$, then for all $j \ge j_*$ property (3a) ensures that $\mathcal{C}_j \subset \mathcal{B}(\mathbf{0}, 2R)$ and in particular that $P_{j,k} \cap \mathcal{B}(\mathbf{0}, 2R) \neq \emptyset$ for all $k \in [K_j]$. Below we will use

$$U_j = \bigcup_{k \in [K_j]} P_{j,k} \cap \mathcal{B}(\mathbf{0}, 2R) \tag{5}$$

to denote the part of the j-th level of the GMRA approximation contained in $\mathcal{B}(\mathbf{0}, 2R)$.

Let us finally recall two standard notions to describe the complexity of a set $K \subset \mathbb{R}^D$. We let $\mathcal{N}(K, \varepsilon)$ denote the Euclidean covering numbers of K. Second, letting $\mathbf{g} \in \mathbb{R}^D$ be standard Gaussian, the Gaussian complexity of K is denoted by $\gamma(K) = \mathbb{E}\sup_{\mathbf{x} \in K} |\langle \mathbf{g}, \mathbf{x}\rangle|$. The following Gaussian complexity bounds will be used in our results below. The proof is a standard application of Dudley's inequality and therefore omitted.

Lemma II.3. *For any $j > j_*$, $\gamma(\mathcal{C}_j) \lesssim R\sqrt{jd}$. For any $\delta > 0$ and $k \in [K_j]$,*

$$\gamma(\mathrm{star}(P_{j,k} - P_{j,k}) \cap \mathcal{B}(\mathbf{0}, \delta)) \lesssim \delta\sqrt{d},$$

$$\gamma(\mathrm{star}(U_j - U_j) \cap \mathcal{B}(\mathbf{0}, \delta)) \lesssim \delta(\sqrt{jd} + \sqrt{d\log(eR/\delta)}).$$

III. Main Results

We present two approaches to recover signals in $\mathcal{M}$, assuming that we have access to a GMRA approximation of $\mathcal{M}$ and to a possibly corrupted vector of quantized measurements $\mathbf{y}_c$ satisfying $d_H(\mathbf{y}_c, \mathbf{y}) \le \beta m$ for some $0 \le \beta < 1$. The first approach is based on the following observation. By GMRA property (3b), the vector $\tilde{\mathbf{x}} := \mathbb{P}_{j,k_j(\mathbf{x})}(\mathbf{x}) \in U_j$ is close to $\mathbf{x} \in \mathcal{M}$ in terms of the Euclidean distance and as a consequence the binary vectors $\mathbf{y} = \mathrm{sign}(\mathbf{A}\mathbf{x} + \boldsymbol{\nu} + \boldsymbol{\tau})$ and $\tilde{\mathbf{y}} = \mathrm{sign}(\mathbf{A}\tilde{\mathbf{x}} + \boldsymbol{\nu} + \boldsymbol{\tau})$ differ only in few entries. Hence, we can view the observed vector $\mathbf{y}_c$ as a corrupted version of $\mathrm{sign}(\mathbf{A}\tilde{\mathbf{x}} + \boldsymbol{\nu} + \boldsymbol{\tau})$. We can therefore recover $\tilde{\mathbf{x}}$ by using a reconstruction program for signals in U_j which is robust to post-quantization noise, i.e., bit corruptions of the quantized measurements. This reconstruction of $\tilde{\mathbf{x}}$ will then be an accurate reconstruction of $\mathbf{x}$ as well.

The following two results from [4] form the foundation for the two steps in this reasoning. Here and below we will assume that $\mathbf{A}$ has i.i.d. symmetric, isotropic, L-subgaussian rows, that $\boldsymbol{\nu}$ has i.i.d. L-subgaussian entries with mean zero and variance σ^2, that $\boldsymbol{\tau}$ has i.i.d. entries which are uniformly distributed in $[-\lambda, \lambda]$, and that $\mathbf{A}$, $\boldsymbol{\nu}$ and $\boldsymbol{\tau}$ are independent. The following two results are immediate from the proofs of [4, Theorem 2.9] and [4, Theorem 1.7], respectively.

Theorem III.1. *There exist constants $a_1, ..., a_4$ depending only on L such that the following holds. Let $T, T' \subset \mathcal{B}(\mathbf{0}, R)$. Fix $\rho, \lambda > 0$ and let $r' = a_1\rho/\sqrt{\log(e\lambda/\rho)}$. For $S = T$ and $S = T'$ set $S_{r'} = (S - S) \cap \mathcal{B}(\mathbf{0}, r')$ and assume*

$$m \ge a_2 \frac{\lambda}{\rho^3}\gamma^2(S_{r'}) + a_2\frac{\lambda}{\rho}\log\mathcal{N}(S, r'). \tag{6}$$

Then with probability at least $1 - 8\exp(-a_3 m\rho/\lambda)$, for every $\mathbf{z}, \mathbf{z}' \in T \cup T'$ with $\|\mathbf{z} - \mathbf{z}'\|_2 \le r'/2$,

$$\frac{1}{m} d_H(\mathrm{sign}(\mathbf{A}\mathbf{z} + \boldsymbol{\nu} + \boldsymbol{\tau}), \mathrm{sign}(\mathbf{A}\mathbf{z}' + \boldsymbol{\nu} + \boldsymbol{\tau})) \le a_4\frac{\rho}{\lambda}.$$

Theorem III.2. *There exist constants $b_1, ..., b_5$ depending only on L such that the following holds. Fix $\delta > 0$, let $T \subset \mathcal{B}(\mathbf{0}, R)$ and $U_\delta = \mathrm{star}(T - T) \cap \mathcal{B}(\mathbf{0}, \delta)$, set $\lambda \ge b_1(\sigma + R)\sqrt{\log(b_1(\sigma + R)/\delta)}$ and let $r = b_2\delta/\log(e\lambda/\delta)$. If m and β satisfy*

$$m \ge b_3\Big(\Big(\frac{\lambda\gamma(U_\delta)}{\delta^2}\Big)^2 + \lambda^2\frac{\log\mathcal{N}(T, r)}{\delta^2}\Big), \tag{7}$$

and $\beta\sqrt{\log(e/\beta)} \le b_4\frac{\delta}{\lambda}$, then, with probability at least $1 - 8\exp(-b_5 m\delta^2/\lambda^2)$ the following holds: for any $\mathbf{x} \in T$ and any $\mathbf{y}_c \in \{-1, 1\}^m$ satisfying $d_H(\mathbf{y}_c, \mathbf{y}) \le \beta m$, any Euclidean projection $\mathbf{x}^\# = \mathbb{P}_T(\frac{\lambda}{m}\mathbf{A}^T\mathbf{y}_c)$ satisfies $\|\mathbf{x}^\# - \mathbf{x}\|_2 \le \delta$.

Using the above two results, we can derive a guarantee for our first recovery algorithm.

Algorithm 1: Exhaustive search

Input: The GMRA, j, $\mathbf{A}$, λ, m, $\mathbf{y}_c$, and R

I. For $k = 1, \ldots, K_j$ compute $\mathbb{P}_{j,k}\Big(\frac{\lambda}{m}\mathbf{A}^T\mathbf{y}_c\Big)$.

II. Select $k^\#$ to be

$$\underset{k\in[K_j]}{\arg\min} \left\| \left(\mathbb{P}_{P_{j,k'}\cap\mathcal{B}(\mathbf{0},2R)} - \mathrm{Id}\right)\left(\frac{\lambda}{m}\mathbf{A}^T\mathbf{y}_c\right)\right\|_2$$

and output $\mathbf{x}^\# = \mathbb{P}_{P_{j,k^\#}\cap\mathcal{B}(\mathbf{0},2R)}(\frac{\lambda}{m}\mathbf{A}^T\mathbf{y}_c)$.

Theorem III.3. *There exist constants $c_0, \ldots, c_5$ depending only on L such that the following holds. Fix $\delta, C_* > 0$, set $r = c_0\delta/\log(e\lambda/\delta)$ and suppose that the GMRA approximation level j satisfies $j > j_*$ (as defined in Section II) and $2^{2j} \geq c_1 C^*/r$. Suppose that*

$$\lambda \geq c_2(\sigma + R)\sqrt{\log(c_2(\sigma+R)/\delta)},$$

$$m \geq c_3\Big(\frac{\lambda^2 d(j + \log(eR/r))}{\delta^2} + \frac{\lambda\log^{3/2}(e\lambda/\delta)\gamma^2(\mathcal{M}_r)}{\delta^3} + \frac{\lambda\log^{1/2}(e\lambda/\delta)}{\delta}\log\mathcal{N}(\mathcal{M}, r)\Big),$$

and $\tilde{\beta} \leq \frac{c_4\delta}{\lambda\sqrt{\log(e\lambda/\delta)}}$. Then, with probability at least $1 - 16\exp(-c_5 m\delta^2/\lambda^2)$, for any $\mathbf{x} \in \mathcal{M}$ with $C_{\mathbf{x}} \leq C^*$ and any $\mathbf{y}_c \in \{-1,1\}^m$ satisfying $d_H(\mathbf{y}_c, \mathbf{y}) \leq \tilde{\beta}m$ the output $\mathbf{x}^\#$ of Algorithm 1 satisfies

$$\|\mathbf{x} - \mathbf{x}^\#\|_2 \leq \delta + C_{\mathbf{x}}2^{-2j}. \quad (8)$$

Proof: Condition on the events E_1 and E_2 of Theorems III.1 for $T = U_j$, $T' = \mathcal{M}$ and III.2 for $T = U_j$ (with U_j as defined in (5)), and respective parameters ρ and δ that will be determined below. By GMRA property (3b), the vector $\tilde{\mathbf{x}} := \mathbb{P}_{j,k_j(\mathbf{x})}(\mathbf{x})$ satisfies $\|\mathbf{x} - \tilde{\mathbf{x}}\| \leq C_{\mathbf{x}}2^{-2j}$. Let us choose ρ so that

$$C^* 2^{-2j} \leq a_1\rho/\sqrt{\log(e\lambda/\rho)}.$$

Then E_1 guarantees that $\frac{1}{m}d_H(\mathrm{sign}(\mathbf{A}\mathbf{x} + \boldsymbol{\nu} + \boldsymbol{\tau}), \mathrm{sign}(\mathbf{A}\tilde{\mathbf{x}} + \boldsymbol{\nu} + \boldsymbol{\tau})) \leq a_4\frac{\rho}{\lambda}$. Clearly, $\mathbf{x}^\# = \mathbb{P}_T(\frac{\lambda}{m}\mathbf{A}^T\mathbf{y}_c)$ for $T = U_j$. Hence, if c_4 and ρ are small enough to ensure that $\beta := \tilde{\beta} + a_4(\rho/\lambda)$ satisfies $\beta\sqrt{\log(e/\beta)} \leq b_4\frac{\delta}{\lambda}$, then the event E_2 implies that $\|\mathbf{x}^\# - \tilde{\mathbf{x}}\|_2 \leq \delta$ and so (8) holds. One readily verifies that

$$\rho = \delta/(C_L\sqrt{\log(e\lambda/\delta)}), \quad (9)$$

where C_L is a large enough constant depending only on L, satisfies all the restrictions. It remains to verify that E_1 and E_2 happen with high probability under the stated assumption on m. By Lemma II.3, for any $0 < \eta < R$,

$$\gamma^2(\mathrm{star}(U_j - U_j)\cap\mathcal{B}(\mathbf{0},\eta)) \lesssim \eta^2(jd + d\log(eR/\eta)).$$

One can now verify that (for our choice of ρ) our assumption on m ensures that the conditions (6) and (7) are satisfied. Hence, by Theorems III.1 and III.2 we find $P(E_1^c \cup E_2^c) \leq 16\exp(-c_L m\delta^2/\lambda^2)$. This completes the proof. ■

Remark III.4. *By defining $\mathbf{c} = \mathbb{P}_{j,k}(\mathbf{0})$, $c = \|\mathbf{c}\|_2$, and noting that the radius of the d-dimensional disk $P_{j,k}\cap\mathcal{B}(\mathbf{0},2R)$ is given by $\sqrt{(2R)^2 - c^2}$, the projection $\mathbb{P}_{P_{j,k}\cap\mathcal{B}(\mathbf{0},2R)}(\mathbf{z})$ can be computed in time $O(dD)$ as*

$$\min\left\{\left\|\Phi_{j,k}^T\Phi_{j,k}\mathbf{z}\right\|_2, \sqrt{(2R)^2 - c^2}\right\}\frac{\Phi_{j,k}^T\Phi_{j,k}\mathbf{z}}{\left\|\Phi_{j,k}^T\Phi_{j,k}\mathbf{z}\right\|_2} + \mathbf{c}.$$

Algorithm 2: Two step method

Input: The GMRA, j, $\mathbf{A}$, λ, m, $\mathbf{y}_c$, R, and $\boldsymbol{\tau}$

I. Identify a center $\mathbf{c}_{j,k'}$ close to $\mathbf{x}$ via

$$\mathbf{c}_{j,k'} \in \underset{\mathbf{c}_{j,k}\in\mathcal{C}_j}{\arg\min}\ d_H(\mathrm{sign}(\mathbf{A}\mathbf{c}_{j,k} + \boldsymbol{\tau}), \mathbf{y}_c).$$

If $d_H(\mathrm{sign}(\mathbf{A}\mathbf{c}_{j,k'} + \boldsymbol{\tau}), \mathbf{y}_c) = 0$, directly choose $\mathbf{x}^* = \mathbf{c}_{j,k'}$ and omit step II.

II. If there is no center lying in the same cell as $\mathbf{x}$, set

$$\mathbf{x}^* = \mathbb{P}_{P_{j,k'}\cap\mathcal{B}(\mathbf{0},2R)}\left(\frac{\lambda}{m}\mathbf{A}^T\mathbf{y}_c\right).$$

The reconstruction method in Theorem III.3 is very robust to noise on the analog measurements and to post-quantization noise. On the other hand, the computational effort needed to compute $\mathbf{x}^\#$ for a given signal $\mathbf{x}$ consists in computing $\mathbb{P}_{P_{j,k}\cap\mathcal{B}(\mathbf{0},2R)}(\frac{\lambda}{m}\mathbf{A}^T\mathbf{y}_c)$ for all $k \in [K_j]$ and determining $k^\#$, leading to a total computational cost of $O(|K_j|mD)$. Since in the worst case $|K_j|$ scales as 2^{jd}, this can quickly become prohibitive for applications involving higher dimensional manifolds and/or high accuracy demands.

To improve on this, we follow the main idea of [2] and introduce a pre-processing step, leading to Algorithm 2. As we will rigorously prove in Lemma III.6, under appropriate conditions the first step of the algorithm will identify a center $\mathbf{c}_{j,k'}$ satisfying (4). By GMRA property (3b), this ensures that $\mathbb{P}_{j,k'}(\mathbf{x})$ will be close to $\mathbf{x}$ in terms of the Euclidean distance. Hence, our prior considerations show that $\mathbf{x}^*$ will be an accurate reconstruction of $\mathbf{x}$.

Clearly, the computational cost of the second step is $O(mD)$. To execute the first step, one needs to compute $\mathrm{sign}(\mathbf{A}\mathbf{c}_{j,k} + \boldsymbol{\tau})$ for all centers, which can be performed offline in time $O(|K_j|mD)$. The online computation cost amounts to nearest neighbor search (with respect to the Hamming metric) of $|K_j|$ binary vectors, which can be performed efficiently (see, e.g., [5], [6]). Hence, Algorithm 2 is substantially faster than our first algorithm if we are interested in recovering multiple signals. On the downside, the recovery guarantees for the second algorithm will be worse. In particular, Algorithm 2 will only be robust with respect to a small amount of noise on the analog measurements (with variance of the order $O(2^{-j})$). In what follows we will assume for simplicity that $\boldsymbol{\nu} = 0$ and that no bit corruptions occur during quantization (i.e., $\mathbf{y}_c = \mathbf{y}$). It is, however, straightforward to adapt the proofs to accommodate bit corruptions.

We need a variant of [4, Theorem 1.1] which can be deduced from [4, Theorem 2.3]

Theorem III.5. *There exist constants $d_0, ..., d_4$ depending only on L such that the following holds. Set*

$$d_{\mathbf{A}}(\mathbf{z}, \mathbf{z}') = \frac{1}{m} d_H(\text{sign}(\mathbf{A}\mathbf{z} + \boldsymbol{\tau}), \text{sign}(\mathbf{A}\mathbf{z}' + \boldsymbol{\tau}))$$

Fix $0 < \delta < R$. If $T \subset \mathcal{B}(\mathbf{0}, R)$, $\lambda \geq d_0 R$, and

$$m \geq d_1 \lambda \log(e\lambda/\delta)\delta^{-3}\gamma^2(T), \tag{10}$$

then with probability at least $1 - 8\exp(-d_2 m\delta/\lambda)$, for any $\mathbf{z}, \mathbf{z}' \in \text{conv}(T)$ such that $\|\mathbf{z} - \mathbf{z}'\|_2 \geq \delta$, one has

$$d_3 \|\mathbf{z} - \mathbf{z}'\|_2 \leq \lambda d_{\mathbf{A}}(\mathbf{z}, \mathbf{z}') \leq d_4 \sqrt{\log(e\lambda/\delta)}\, \|\mathbf{z} - \mathbf{z}'\|_2 .$$

Lemma III.6. *There exist e_0, e_1, e_2, e_3 depending only on L such that the following holds for any $\theta > 0$. If $\lambda \geq e_0 R$ and $m \geq e_1 \lambda\theta^{-3}\log(e\lambda/\theta)(\gamma^2(\mathcal{M}) + R^2 jd)$, then with probability at least $1 - 16\exp(-e_2 m\theta/\lambda)$ for all $\mathbf{x} \in \mathcal{M}$ the center $\mathbf{c}_{j,k'}$ chosen in step I of Algorithm 2 fulfills $\|\mathbf{c}_{j,k'} - \mathbf{x}\|_2 \leq \theta$ if $d_H(\text{sign}(\mathbf{A}\mathbf{c}_{j,k'} + \boldsymbol{\tau}), \mathbf{y}) = 0$ and otherwise $\|\mathbf{x} - \mathbf{c}_{j,k'}\|_2$ is bounded by*

$$e_3 \sqrt{\log(e\lambda/\theta)} \cdot \max\left\{\|\mathbf{x} - \mathbf{c}_{j,k_j(\mathbf{x})}\|_2, \theta\right\}.$$

Proof: Condition on the events E_1 and E_2 of Theorems III.5 and III.1 for $T = \mathcal{M} \cup \mathcal{C}_j$ and parameters $\delta = \theta$ and $\rho = c_L \theta\sqrt{\log(e\lambda/\theta)}$ for a suitably large constant c_L depending only on L. By E_1, $\|\mathbf{c}_{j,k'} - \mathbf{x}\|_2 \leq \theta$ if $d_H(\text{sign}(\mathbf{A}\mathbf{c}_{j,k'} + \boldsymbol{\tau}), \mathbf{y}) = 0$.

Suppose now that $d_H(\text{sign}(\mathbf{A}\mathbf{c}_{j,k'} + \boldsymbol{\tau}), \mathbf{y}) > 0$. If $\|\mathbf{c}_{j,k'} - \mathbf{x}\|_2 < \theta$ then the claim is trivial, so we may assume $\|\mathbf{c}_{j,k'} - \mathbf{x}\|_2 \geq \theta$. If $\left\|\mathbf{c}_{j,k_j(\mathbf{x})} - \mathbf{x}\right\|_2 \geq \theta$, then by the event E_1 and the definition of $\mathbf{c}_{j,k'}$ we find

$$\begin{aligned} d_3 \|\mathbf{c}_{j,k'} - \mathbf{x}\|_2 &\leq \lambda d_{\mathbf{A}}(\mathbf{c}_{j,k'}, \mathbf{x}) \leq \lambda d_{\mathbf{A}}(\mathbf{c}_{j,k_j(\mathbf{x})}, \mathbf{x}) \\ &\leq d_4 \sqrt{\log(e\lambda/\theta)}\left\|\mathbf{c}_{j,k_j(\mathbf{x})} - \mathbf{x}\right\|_2 . \end{aligned}$$

Similarly, if $\left\|\mathbf{c}_{j,k_j(\mathbf{x})} - \mathbf{x}\right\|_2 \leq \theta$, then

$$d_3 \|\mathbf{c}_{j,k'} - \mathbf{x}\|_2 \leq c_L a_4 \sqrt{\log(e\lambda/\theta)}\theta$$

by E_1 and E_2. Finally, $\gamma^2(\mathcal{M} \cup \mathcal{C}_j) \lesssim \gamma^2(\mathcal{M}) + R^2 jd$ by Lemma II.3. Hence our conditions on λ and m ensure that E_1 and E_2 happen with the stated probability. ■

We can now derive a performance guarantee for Algorithm 2.

Theorem III.7. *There exist constants $f_0, f_1, ..., f_5$ depending only on L such that the following holds. Fix $\delta > 0$, let $r = f_0\delta/\log(e\lambda/\delta)$ and suppose that the GMRA approximation level j satisfies $j > j_*$ and $2^{j/2} \geq f_1 \tilde{C}^*/r$. Let C_1, C_2 be the GMRA constants. Set $\theta = C_1 2^{-j-1}$ and assume that C_1, C_2 are large enough to ensure that $e_3\sqrt{\log(e\lambda/\theta)} \leq C_2\sqrt{j + \log(R)}$. Suppose that $\lambda \geq f_2 R\sqrt{\log(f_2 R/\delta)}$ and m exceeds*

$$f_3\Big[\frac{\lambda^2\log(eR/r)d}{\delta^2} + \frac{\lambda\log(e\lambda/\theta)}{\theta^3}(\gamma^2(\mathcal{M}) + R^2 jd)\Big].$$

Then, with probability at least $1 - 16\exp(-f_4 m\delta^2/\lambda^2) - 16\exp(-f_4 m\theta/\lambda)$, for any $\mathbf{x} \in \mathcal{M}$ with $\tilde{C}_{\mathbf{x}} \leq \tilde{C}^*$ the following holds: if Algorithm 2 receives $\mathbf{y}_c = \mathbf{y} = \text{sign}(\mathbf{A}\mathbf{x} + \boldsymbol{\tau})$, then its output $\mathbf{x}^*$ satisfies

$$\|\mathbf{x} - \mathbf{x}^*\|_2 \leq \theta + \delta + \tilde{C}_{\mathbf{x}} 2^{-j/2}. \tag{11}$$

Proof: Condition on the event E_1 of Lemma III.6 for $\theta = C_1 2^{-j-1}$. Then $\|\mathbf{c}_{j,k'} - \mathbf{x}\|_2 \leq \theta$ if $d_H(\text{sign}(\mathbf{A}\mathbf{c}_{j,k'} + \boldsymbol{\tau}), \mathbf{y}) = 0$ and so $\|\mathbf{x}^* - \mathbf{x}\|_2 \leq \theta$ in this case. Otherwise, $\|\mathbf{x} - \mathbf{c}_{j,k'}\|_2$ is bounded by

$$e_3\sqrt{\log(e\lambda/\theta)} \cdot \max\{\|\mathbf{x} - \mathbf{c}_{j,k_j(\mathbf{x})}\|_2, \theta\}.$$

and, since $e_3\sqrt{\log(e\lambda/\theta)} \leq C_2\sqrt{j + \log(R)}$, GMRA property (3c) implies

$$\|\mathbf{x} - \mathbb{P}_{j,k'}(\mathbf{x})\|_2 \leq \tilde{C}_{\mathbf{x}} 2^{-j/2}.$$

Analogously to the proof of Theorem III.3, (11) is valid if we condition on the events E_2 and E_3 of Theorems III.1 and III.2 for $T = T' = P_{j,k'}$ with respective parameters ρ and δ satisfying (9).

It remains to verify that E_1, E_2, and E_3 happen with the stated probability. Clearly, $P(E_1^c) \leq 16\exp(-c_L m\theta/\lambda)$. By Lemma II.3, for any $0 < \eta < R$, $\gamma^2(\text{star}(P_{j,k} - P_{j,k}) \cap \mathcal{B}(\mathbf{0}, \eta)) \lesssim \eta^2 d$ and it is clear that $\log\mathcal{N}(P_{j,k}, \eta) \lesssim d\log(eR/\eta)$. One can now readily verify that our assumption on m ensures that the conditions (6) and (7) are satisfied. Hence, $P(E_2^c \cup E_3^c) \leq 16\exp(-c_L m\delta^2/\lambda^2)$. ■

IV. Acknowledgments

The authors were supported in part by DFG SFB Transregio 109 (Sara Krause-Solberg), DFG SPP 1798 (Sjoerd Dirksen, Johannes Maly), and by NSF CCF 1615489 (Mark Iwen).

References

[1] W. Allard, G. Chen, and M. Maggioni, "Multi-scale geometric methods for data sets II: Geometric multi-resolution analysis," *Appl. Comput. Harmon. Anal.*, vol. 32, no. 3, pp. 435–462, 2012, iSSN 1063-5203.

[2] M. A. Iwen, F. Krahmer, S. Krause-Solberg, and J. Maly, "On recovery guarantees for one-bit compressed sensing on manifolds," *ArXiv:1807.06490*, 2018.

[3] M. A. Iwen and M. Maggioni, "Approximation of points on low-dimensional manifolds via random linear projections," *Inform. Inf.*, vol. 2, no. 1, pp. 1–31, 2013.

[4] S. Dirksen and S. Mendelson, “Non-gaussian hyperplane tessellations and robust one-bit compressed sensing,” *ArXiv:1805.09409*, 2018.
[5] A. Beygelzimer, S. Kakade, and J. Langford, “Cover trees for nearest neighbor,” in *Proceedings of the 23rd International Conference on Machine Learning*, ser. ICML ’06. ACM, 2006, pp. 97–104.
[6] M. Norouzi, A. Punjani, and D. J. Fleet, “Fast search in hamming space with multi-index hashing,” in *2012 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2012, pp. 3108–3115.