# Lossless Linear Analog Compression

Giovanni Alberti<sup>‡</sup>, Helmut Bölcskei<sup>†</sup>, Camillo De Lellis<sup>\*</sup>, Günther Koliander<sup>◊</sup>, and Erwin Riegler<sup>†</sup>

<sup>‡</sup>Dept. of Mathematics, University of Pisa, Italy, Email: galberti1@dm.unipi.it

<sup>†</sup>Dept. of IT & EE, ETH Zurich, Switzerland, Email: {boelcskei, eriegler}@nari.ee.ethz.ch

\*Dept. of Mathematics, University of Zurich, Switzerland, Email: camillo.delellis@math.uzh.ch

<sup>o</sup>Dept. of Telecommunications, Vienna University of Technology, Austria, Email: guenther.koliander@tuwien.ac.at

Abstract-We establish the fundamental limits of lossless linear analog compression by considering the recovery of random vectors  $x \in \mathbb{R}^m$  from the noiseless linear measurements y = Axwith measurement matrix  $\boldsymbol{A} \in \mathbb{R}^{n imes m}$ . Specifically, for a random vector  $\mathbf{x} \in \mathbb{R}^m$  of *arbitrary* distribution we show that  $\mathbf{x}$  can be recovered with zero error probability from  $n > \inf \underline{\dim}_{MB}(\mathcal{U})$ linear measurements, where  $\underline{\dim}_{\mathrm{MB}}(\cdot)$  denotes the lower modified Minkowski dimension and the infimum is over all sets  $\mathcal{U} \subseteq \mathbb{R}^m$ with  $P[x \in U] = 1$ . This achievability statement holds for Lebesgue almost all measurement matrices A. We then show that s-rectifiable random vectors-a stochastic generalization of s-sparse vectors-can be recovered with zero error probability from n > s linear measurements. From classical compressed sensing theory, where x is deterministic, we would expect n > sto be necessary for recovery with zero error probability. Surprisingly, certain classes of s-rectifiable random vectors can be recovered from fewer than s measurements. Imposing an additional regularity condition on the distribution of s-rectifiable random vectors x, we do get the expected converse result of s measurements being necessary. The resulting class of random vectors appears to be new and will be referred to as s-analytic random vectors. Finally, we show that, even for analytic random vectors, the sparsity level in terms of the number of non-zero entries of the realizations of x can be larger than the analyticityparameter s.

#### I. INTRODUCTION

Compressed sensing [1]-[3] deals with the recovery of unknown sparse vectors  $oldsymbol{x} \in \mathbb{R}^m$  from a small (relative to m) number, n, of linear measurements of the form y = Ax, where  $A \in \mathbb{R}^{n \times m}$  is referred to as the measurement matrix. Wu and Verdú [4], [5] developed an information-theoretic framework for compressed sensing, fashioned as an almost lossless analog compression problem. Specifically, [4] presents asymptotic achievability bounds, which show that for almost all (a.a.) measurement matrices A a random i.i.d. vector  $\mathbf{x}$ can be recovered with arbitrarily small probability of error from  $n = \lfloor Rm \rfloor$  linear measurements, provided that  $R > R_{\rm B}$ , where  $R_{\rm B}$  denotes the Minkowski dimension compression rate [4] of x. For the special case of the i.i.d. components in x having a discrete-continuous mixture distribution, this threshold is tight in the sense of  $R \ge R_{\rm B}$  being necessary for the existence of a measurement matrix A such that x can be recovered with probability of error strictly smaller than 1 for m sufficiently large.

Discrete-continuous mixture distributions  $\rho\mu^{c} + (1 - \rho)\mu^{d}$ are relevant as  $\lfloor \rho m \rfloor$ —by the law of large numbers—can be interpreted as the sparsity level of **x** and  $R_{B} = \rho$ . A more direct and non-asymptotic (i.e., fixed-m) statement in [4] says that a.a (with respect to a  $\sigma$ -finite Borel measure) *s*-sparse random vectors can be recovered with zero probability of error provided that m > s. Again, this result holds for Lebesgue a.a. measurement matrices  $\mathbf{A} \in \mathbb{R}^{n \times m}$ . A corresponding converse does, however, not seem to be available.

*Contributions.* We establish the fundamental limits of lossless (i.e., zero probability of error) linear analog compression in the non-asymptotic (i.e., fixed-m) regime for random vectors **x** of arbitrary distribution. In particular, **x** need not be i.i.d. or supported on unions of subspaces (as in classical compressed sensing theory). The formal statement of the problem we consider is as follows. Suppose we have n (noiseless) linear measurements of the random vector  $\mathbf{x} \in \mathbb{R}^m$  in the form of  $\mathbf{y} = A\mathbf{x}$ . For a given  $\varepsilon \in [0, 1)$ , we want to determine whether a decoder, i.e., a Borel measurable map  $g_A : \mathbb{R}^n \to \mathbb{R}^m$  exists such that

$$P[g_{\boldsymbol{A}}(\boldsymbol{A}\boldsymbol{x}) \neq \boldsymbol{x}] \leq \varepsilon.$$
(1)

Specifically, we shall be interested in statements of the following form:

Property P1: For Lebesgue a.a. measurement matrices  $A \in \mathbb{R}^{n \times m}$ , there exists a decoder  $g_A$  satisfying (1) almost surely, i.e., with  $\varepsilon = 0$ .

Property P2: There exists an  $\varepsilon \in [0, 1)$ , an  $\mathbf{A} \in \mathbb{R}^{n \times m}$ , and a decoder  $g_{\mathbf{A}}$  satisfying (1).

Our main achievability result is as follows. For  $\mathbf{x} \in \mathbb{R}^m$  of arbitrary distribution, we show that P1 holds for  $n > \inf \underline{\dim}_{\mathrm{MB}}(\mathcal{U})$ , where  $\underline{\dim}_{\mathrm{MB}}(\cdot)$  denotes the lower modified Minkowski dimension (see Definition 2) and the infimum is over all sets  $\mathcal{U} \subseteq \mathbb{R}^m$  with  $\mathrm{P}[\mathbf{x} \in \mathcal{U}] = 1$ . We emphasize that it is the usage of modified Minkowski dimension, as opposed to Minkowski dimension, that allows us to obtain an achievability result for  $\varepsilon = 0$ . The central conceptual element in its proof is a slightly modified version of the probabilistic null-space property first reported in [6]. The asymptotic achievability bounds in [4] can be recovered in our framework.

We make the connection of our results to classical compressed sensing explicit by considering random vectors  $\mathbf{x} \in \mathbb{R}^m$  that consist of *s* i.i.d. Gaussian entries at positions drawn uniformly at random and that have all other entries equal to zero. This class can be considered a stochastic analogon of *s*-sparse vectors and belongs to the more general class of *s*-rectifiable random vectors, originally introduced in [7] to derive a new concept of entropy that goes beyond classical entropy and differential entropy. An s-rectifiable random vector has a distribution that is supported on an s-rectifiable set [8, Def. 4.1] and is absolutely continuous with respect to the sdimensional Hausdorff measure.<sup>1</sup> Our achievability result particularized to s-rectifiable random vectors shows that P1 holds for n > s. From classical compressed sensing theory, where **x** is deterministic, we would expect  $n \ge s$  to be necessary for recovery with zero error probability. Our information-theoretic framework reveals, however, that this is not the case for certain classes of s-rectifiable random vectors. This will be illustrated by way of an example, which constructs a 2-rectifiable set  $\mathcal{G} \subseteq \mathbb{R}^3$  of positive 2-dimensional Hausdorff measure that can be compressed linearly in a one-to-one fashion into  $\mathbb{R}$ . Operationally, this implies that zero error probability recovery from n = 1 < s = 2 measurement is possible. What renders this result particularly surprising is that  $\mathcal{G}$  contains the image under a continuous differentiable mapping—of a set in  $\mathbb{R}^2$ of positive Lebesgue measure. We then show that imposing a regularity condition on the distribution of x, we do get the expected converse result in the sense of  $n \ge s$  being necessary for P2 to hold. The resulting class of random vectors appears to be new and will be referred to as s-analytic random vectors. Finally, we show that, even for analytic random vectors, the sparsity level in terms of the number of non-zero entries of the realizations of x can be larger than the analyticity-parameter s.

Notation. Capital boldface letters  $A, B, \ldots$  designate deterministic matrices and lower-case boldface letters  $a, b, \ldots$ stand for deterministic vectors. We use sans-serif letters, e.g.  $\mathbf{x}$ , for random quantities and roman letters, e.g.  $\mathbf{x}$ , for deterministic quantities. For measures  $\mu$  and  $\nu$  on the same measure space, we write  $\mu \ll \nu$  to express that  $\mu$  is absolutely continuous with respect to  $\nu$  (i.e., for every measurable set A,  $\nu(\mathcal{A}) = 0$  implies  $\mu(\mathcal{A}) = 0$ ). The product measure of  $\mu$  and  $\nu$ is denoted by  $\mu \times \nu$ . The superscript <sup>T</sup> stands for transposition.  $\|x\|_2 = \sqrt{x^{\mathsf{T}}x}$  is the Euclidean norm of x and  $\|x\|_0$  denotes the number of non-zero entries of x. For the Euclidean space  $(\mathbb{R}^k, \|\cdot\|_2)$ , we let the open ball of radius  $\rho$  centered at  $\boldsymbol{u} \in \mathbb{R}^k$  be  $\mathcal{B}_k(\boldsymbol{u}, \rho)$ , and  $V(k, \rho)$  refers to its volume.  $\mathscr{L}^n$ denotes the Lebesgue measure on  $\mathbb{R}^n$ . If  $f: \mathbb{R}^k \to \mathbb{R}^l$  is differentiable, we write Df(v) for the differential of f at  $oldsymbol{v} \in \mathbb{R}^k$  and we define the  $\min(k,l)$ -dimensional Jacobian Jf(v) at  $v \in \mathbb{R}^k$  by  $Jf(v) = \sqrt{\det(Df(v))}(Df(v))^{\mathsf{T}})$ , if l < k, and  $Jf(\boldsymbol{v}) = \sqrt{\det((Df(\boldsymbol{v}))^{\mathsf{T}} Df(\boldsymbol{v}))}$ , if  $l \geq k$ . For a mapping  $f, f \not\equiv 0$  means that f is not identically zero. For  $f: \mathbb{R}^k \to \mathbb{R}^l$  and  $\mathcal{A} \subseteq \mathbb{R}^k$ ,  $f|_{\mathcal{A}}$  denotes the restriction of f to  $\mathcal{A}$ . A mapping is said to be  $C^1$  if its derivative exists and is continuous. ker(f) stands for the kernel of f.

The definitions of the fractal quantities used in this paper are standard and can be found, along with their basic properties in, e.g., [9], [10]. Throughout the paper, we omit proofs due to space limitations.

#### II. ACHIEVABILITY

We quantify the description complexity of random vectors  $\mathbf{x} \in \mathbb{R}^m$  of general distribution through the infimum over the lower modified Minkowski dimensions of sets  $\mathcal{U} \subseteq \mathbb{R}^m$  with  $P[\mathbf{x} \in \mathcal{U}] = 1$  (referred to as support sets of  $\mathbf{x}$  in the following). We start by defining Minkowski dimension.

Definition 1. (Minkowski dimension<sup>2</sup>) Let  $\mathcal{U}$  be a non-empty bounded set in  $\mathbb{R}^m$ . The lower Minkowski dimension of  $\mathcal{U}$  is defined as

$$\underline{\dim}_{\mathrm{B}}(\mathcal{U}) = \liminf_{\rho \to 0} \frac{\log N_{\mathcal{U}}(\rho)}{\log \frac{1}{\rho}}$$

and the upper Minkowski dimension as

$$\overline{\dim}_{\mathrm{B}}(\mathcal{U}) = \limsup_{\rho \to 0} \frac{\log N_{\mathcal{U}}(\rho)}{\log \frac{1}{\rho}},$$

where

$$N_{\mathcal{U}}(\rho) = \min\left\{k \in \mathbb{N} : \mathcal{U} \subseteq \bigcup_{i \in \{1, \dots, k\}} \mathcal{B}_m(\boldsymbol{u}_i, \rho), \ \boldsymbol{u}_i \in \mathbb{R}^m\right\}$$

is the covering number of  $\mathcal{U}$  for radius  $\rho$ . If  $\underline{\dim}_{\mathrm{B}}(\mathcal{U}) = \overline{\dim}_{\mathrm{B}}(\mathcal{U}) =: \dim_{\mathrm{B}}(\mathcal{U})$ , we say that  $\dim_{\mathrm{B}}(\mathcal{U})$  is the Minkowski dimension of  $\mathcal{U}$ .

Minkowski dimension is a useful measure only for (nonempty) bounded sets, as it equals infinity for unbounded sets. Support sets of random vectors are, however, not bounded in general (see, e.g., the support sets in Examples 1 and 4 below). A measure of description complexity that applies to unbounded sets as well is modified Minkowski dimension.

Definition 2. (Modified Minkowski dimension) Let  $\mathcal{U} \subseteq \mathbb{R}^m$  be a non-empty set. The lower modified Minkowski dimension of  $\mathcal{U}$  is defined as

$$\underline{\dim}_{\mathrm{MB}}(\mathcal{U}) = \inf \left\{ \sup_{i \in \mathcal{I}} \underline{\dim}_{\mathrm{B}}(\mathcal{U}_i) : \mathcal{U} \subseteq \bigcup_{i \in \mathcal{I}} \mathcal{U}_i \right\},\$$

where the infimum is over all countable covers of  $\mathcal{U}$  by nonempty bounded Borel sets.<sup>3</sup> The upper modified Minkowski dimension of  $\mathcal{U}$  is

$$\overline{\dim}_{\mathrm{MB}}(\mathcal{U}) = \inf \left\{ \sup_{i \in \mathcal{I}} \overline{\dim}_{\mathrm{B}}(\mathcal{U}_i) : \mathcal{U} \subseteq \bigcup_{i \in \mathcal{I}} \mathcal{U}_i 
ight\}$$

where, again, the infimum is over all countable covers of  $\mathcal{U}$  by non-empty bounded Borel sets. If  $\underline{\dim}_{MB}(\mathcal{U}) = \overline{\dim}_{MB}(\mathcal{U}) =: \dim_{MB}(\mathcal{U})$ , we say that  $\dim_{MB}(\mathcal{U})$  is the modified Minkowski dimension of  $\mathcal{U}$ .

<sup>&</sup>lt;sup>1</sup>Note that the classical Lebesgue decomposition of measures into continuous, discrete, and singular parts is not useful for *s*-rectifiable random vectors as their distributions are always singular (except for the trivial cases s = m and s = 0). We therefore use the *s*-dimensional Hausdorff measure as reference measure for the ambient space.

 $<sup>^2</sup>Minkowski$  dimension is sometimes also referred to as box-counting dimension, which is the origin of the subscript B in the notation  $\dim_{\rm B}(\cdot)$  used below.

<sup>&</sup>lt;sup>3</sup>The assumption of the sets in the covers being Borel is for technical convenience, as for a non-empty set  $\mathcal{U} \subseteq \mathbb{R}^m$  its closure  $\overline{\mathcal{U}}$  is a Borel set of the same lower and upper Minkowski dimension as  $\mathcal{U}$  [9, Prop. 3.4], a property needed in the proof of our main achievability result.

Upper modified Minkowski dimension, in addition, has the advantage of being countably stable [9, Sec. 3.4], whereas upper Minkowski dimension is only finitely stable. This aspect will turn out to be of key importance in particularizing our achievability result, stated next, for *s*-rectifiable random vectors.

**Theorem 1.** For  $\mathbf{x} \in \mathbb{R}^m$  of arbitrary distribution,  $n > \inf \underline{\dim}_{MB}(\mathcal{U})$  is sufficient for Property P1 to hold, where the infimum is over all support sets  $\mathcal{U} \subseteq \mathbb{R}^m$  of  $\mathbf{x}$ .

This theorem generalizes the achievability result of [4] to random vectors  $\mathbf{x} \in \mathbb{R}^m$  of *arbitrary* distribution. Specifically, neither do the entries of  $\mathbf{x}$  have to be i.i.d. nor does  $\mathbf{x}$  have to be generated according to the finite unions of subspaces model. Finally, perhaps most importantly, the result is non-asymptotic (i.e., for finite m) and pertains to zero error probability.

The central conceptual element in the derivation of Theorem 1 is the following probabilistic null-space property, first reported in [6] for (non-empty) bounded sets and expressed in terms of lower Minkowski dimension. If the lower modified Minkowski dimension of a non-empty (possibly unbounded) set  $\mathcal{U}$  is smaller than n, then, for a.a. measurement matrices A, the set  $\mathcal{U}$  intersects the (m - n)-dimensional kernel of Aat most trivially. What is remarkable here is that the notions of Euclidean dimension (for the kernel of the mapping) and of lower modified Minkowski dimension (for  $\mathcal{U}$ ) are compatible. The formal statement is as follows.

**Proposition 1.** Let s > 0 and  $\mathbf{A} = (\mathbf{a}_1 \dots \mathbf{a}_m)$  with independent columns that are uniformly distributed on  $\mathcal{B}_n(\mathbf{0}, s)$ . Suppose that  $\mathcal{U} \subseteq \mathbb{R}^m$  with  $\underline{\dim}_{MB}(\mathcal{U}) < n$ . Then, we have

$$P[\ker(\mathbf{A}) \cap \mathcal{U} \setminus \{\mathbf{0}\} \neq \emptyset] = 0.$$

#### A. Rectifiable random vectors

We next particularize our achievability result for *s*-rectifiable random vectors  $\mathbf{x}$ —defined below—and start by introducing the central concepts needed, namely, Hausdorff measures, Hausdorff dimension, and (locally) Lipschitz mappings.

Definition 3. (Hausdorff measure) Let  $s \in [0, \infty)$  and  $\mathcal{U} \subseteq \mathbb{R}^m$ . The s-dimensional Hausdorff measure of  $\mathcal{U}$  is given by

$$\mathscr{H}^{s}(\mathcal{U}) = \lim_{\delta \to 0} \mathscr{H}^{s}_{\delta}(\mathcal{U})$$

where, for  $0 < \delta \leq \infty$ ,

$$\mathcal{H}_{\delta}^{s}(\mathcal{U}) = \frac{V(s,1)}{2^{s}} \inf \left\{ \sum_{i \in \mathcal{I}} \operatorname{diam}(\mathcal{U}_{i})^{s} : \operatorname{diam}(\mathcal{U}_{i}) < \delta, \mathcal{U} \subseteq \bigcup_{i \in \mathcal{I}} \mathcal{U}_{i} \right\}$$

for countable covers  $\{\mathcal{U}_i\}_{i\in\mathcal{I}}$  and the diameter of  $\mathcal{U}\subseteq\mathbb{R}^n$  is defined as

$$\operatorname{diam}(\mathcal{U}) = \begin{cases} \sup\{\|\boldsymbol{u} - \boldsymbol{v}\|_2 : \boldsymbol{u}, \boldsymbol{v} \in \mathcal{U}\}, & \text{for } \mathcal{U} \neq \emptyset\\ 0, & \text{for } \mathcal{U} = \emptyset. \end{cases}$$



Fig. 1. ([9, Fig. 2.3]) Graph of  $\mathscr{H}^s(\mathcal{U})$  as a function of  $s \in [0, m]$  for a set  $\mathcal{U} \subseteq \mathbb{R}^m$ .

Definition 4. (Hausdorff dimension) The Hausdorff dimension of  $\mathcal{U} \subseteq \mathbb{R}^m$  is

$$\begin{split} \dim_{\mathrm{H}}(\mathcal{U}) &= \sup\{s \geq 0 : \mathscr{H}^{s}(\mathcal{U}) = \infty\}\\ &= \inf\{s \geq 0 : \mathscr{H}^{s}(\mathcal{U}) = 0\}, \end{split}$$

i.e.,  $\dim_{\mathrm{H}}(\mathcal{U})$  is the value of s for which the sharp transition from  $\infty$  to 0 occurs in Figure 1.

Definition 5. (Locally Lipschitz mapping) We call

(i) a mapping  $f : \mathcal{U} \to \mathbb{R}^l$ , where  $\mathcal{U} \subseteq \mathbb{R}^k$ , Lipschitz if there exists a constant  $L \ge 0$  such that

$$||f(\boldsymbol{u}) - f(\boldsymbol{v})||_2 \le L ||\boldsymbol{u} - \boldsymbol{v}||_2,$$
 (2)

for all  $u, v \in U$ . The smallest constant L for which (2) holds is called the Lipschitz constant of f;

(ii) a mapping  $f : \mathbb{R}^k \to \mathbb{R}^l$  locally Lipschitz if, for each compact set  $\mathcal{K} \subseteq \mathbb{R}^k$ , the mapping  $f|_{\mathcal{K}} : \mathcal{K} \to \mathbb{R}^l$  is Lipschitz.

We are now ready to define the notion of *s*-rectifiable sets and *s*-rectifiable random vectors.

Definition 6. An  $\mathscr{H}^s$ -measurable set  $\mathcal{U} \subseteq \mathbb{R}^m$  is called *s*-rectifiable if there exist a countable set  $\mathcal{I}$ , bounded Borel sets  $\mathcal{A}_i \subseteq \mathbb{R}^s$ ,  $i \in \mathcal{I}$ , and Lipschitz mappings  $\varphi_i : \mathcal{A}_i \to \mathbb{R}^m$ ,  $i \in \mathcal{I}$  such that

$$\mathscr{H}^{s}\Big(\mathcal{U}\setminus\bigcup_{i\in\mathcal{I}}\varphi_{i}(\mathcal{A}_{i})\Big)=0.$$

Definition 7. The random vector  $\mathbf{x} \in \mathbb{R}^m$  is called *s*-rectifiable if it has an *s*-rectifiable support set and  $\mu_{\mathbf{x}} \ll \mathscr{H}^s$ .

The following example speaks to the relevance of the notion of *s*-rectifiable random vectors as it shows that a stochastic analogon of the union of subspaces model—used pervasively in classical compressed sensing—is *s*-rectifiable.

*Example* 1. Suppose that  $\mathbf{x} \in \mathbb{R}^m$  has s i.i.d. Gaussian entries at positions drawn uniformly at random and all other entries are equal to zero. Then, the s-rectifiable set

$$\mathcal{U} = \{ \boldsymbol{x} \in \mathbb{R}^m : \| \boldsymbol{x} \|_0 = s \},$$
(3)

is a support set of **x**. We show in Example 3 that  $\mu_{\mathbf{x}} \ll \mathcal{H}^s$ , which implies *s*-rectifiability of **x**.



Fig. 2. The set  $\mathcal{U} \subseteq \mathbb{R}^m$  contains the image of a set  $\mathcal{A} \subseteq \mathbb{R}^s$  with positive Lebesgue measure  $\mathscr{L}^s(\mathcal{A}) > 0$ . The mapping *h* is one-to-one on  $\mathcal{A}$ .

We next establish an important uniqueness property of *s*-rectifiable random vectors.

#### **Lemma 1.** If x is s-rectifiable and t-rectifiable, then s = t.

Roughly speaking the reason for this uniqueness is the following. If s is too small, then there exists no s-rectifiable support set for x. If s is too large, then  $\mu_x \ll \mathcal{H}^s$  is violated as a consequence of the sharp transition behavior of Hausdorff measure depicted in Figure 1.

We next particularize our achievability result, Theorem 1, to *s*-rectifiable random vectors. To this end, we first establish an upper bound on  $\overline{\dim}_{MB}(\mathcal{U})$  for *s*-rectifiable sets  $\mathcal{U}$ .

**Lemma 2.** For 
$$\mathcal{U} \subseteq \mathbb{R}^m$$
 s-rectifiable we have  $\dim_{\mathrm{MB}}(\mathcal{U}) \leq s$ .

Combining Lemma 2 and Theorem 1 yields the following achievability result for *s*-rectifiable random vectors.

**Corollary 1.** For  $\mathbf{x} \in \mathbb{R}^m$  s-rectifiable, n > s is sufficient for P1 to hold.

#### III. CONVERSE

Our achievability result particularized to s-rectifiable random vectors shows that P1 holds for n > s. From classical compressed sensing theory, where x is deterministic, we would expect  $n \ge s$  to be necessary for recovery with zero error probability. Our information-theoretic framework reveals, however, that this is not the case for certain classes of srectifiable random vectors. This surprising phenomenon will be illustrated through the following example. We construct a 2-rectifiable set  $\mathcal{G} \subset \mathbb{R}^3$  of positive 2-dimensional Hausdorff measure that can be compressed linearly in a one-to-one fashion into  $\mathbb{R}$ . What renders this result surprising is that all this is possible although G contains the image—under a continuous differentiable mapping—of a set in  $\mathbb{R}^2$  of positive Lebesgue measure (see Figure 2). Operationally, this shows that 2-rectifiable random vectors  $\mathbf{x}$  with support set  $\mathcal{G}$  can be recovered from n = 1 < s = 2 linear measurement almost surely.

*Example* 2. We construct a 2-rectifiable set  $\mathcal{G} \subseteq \mathbb{R}^3$  with  $\mathscr{H}^2(\mathcal{G}) > 0$  and a corresponding linear mapping  $f : \mathbb{R}^3 \to \mathbb{R}$  such that f is one-to-one on  $\mathcal{G} = h(\mathcal{A})$ , where  $h : \mathbb{R}^2 \to \mathbb{R}^3$  is  $C^1$ ,  $\mathcal{A} \subseteq \mathbb{R}^2$  has  $\mathscr{L}^2(\mathcal{A}) > 0$ , and h is one-to-one on  $\mathcal{A}$ .

Construction of  $\mathcal{G}$ : It can be shown that there exist a  $C^1$ mapping  $\kappa : \mathbb{R}^2 \to \mathbb{R}$  and a bounded Borel set  $\mathcal{A} \subseteq \mathbb{R}^2$ with  $0 < \mathscr{L}^2(\mathcal{A}) < \infty$  such that  $\kappa$  is one-to-one on  $\mathcal{A}$ . Let  $\mathcal{G} = \{(\mathbf{z} \ \kappa(\mathbf{z}))^{\mathsf{T}} \mid \mathbf{z} \in \mathcal{A}\} \subseteq \mathbb{R}^3$ . Since  $\kappa$  is a  $C^1$ -mapping,

$$egin{aligned} h: \mathbb{R}^2 &
ightarrow \mathbb{R}^3 \ oldsymbol{z} &\mapsto (oldsymbol{z} \; \kappa(oldsymbol{z}))^\mathsf{T} \end{aligned}$$

is also  $C^1$ . It follows from the mean value theorem [11, Thm. 4.1] applied component-wise that h is also locally Lipschitz. We then cover  $\mathbb{R}^2$  by compact sets  $\mathcal{K}_i, i \in \mathcal{I}$ , with  $\mathcal{I}$  countable. The local Lipschitz property of h implies that the mappings  $\varphi_i = h|_{\mathcal{K}_i}, i \in \mathcal{I}$ , are Lipschitz. Therefore, by Definition 6,

$$\mathcal{G} = \bigcup_{i \in \mathcal{I}} \varphi_i(\mathcal{K}_i \cap \mathcal{A})$$

is 2-rectifiable.

 $\mathscr{H}^2(\mathcal{G}) > 0$ : Let  $\pi : \mathbb{R}^3 \to \mathbb{R}^2$ ,  $(x_1 \ x_2 \ x_3)^{\mathsf{T}} \to (x_1 \ x_2)^{\mathsf{T}}$ . Clearly,  $\pi$  is a Lipschitz mapping with Lipschitz constant equal to one. Using [10, Prop. 2.49, Property (iv)] and [10, Thm. 2.53] we get  $\mathscr{H}^2(\mathcal{G}) \geq \mathscr{H}^2(\pi(\mathcal{G})) = \mathscr{H}^2(\mathcal{A}) = \mathscr{L}^2(\mathcal{A}) > 0$ .

Construction of f: The mapping

$$f: \mathbb{R}^3 \to \mathbb{R}$$
$$(x_1 \ x_2 \ x_3)^\mathsf{T} \mapsto x_3$$

is linear and one-to-one on  $\mathcal{G}$ .

The structure theorem in geometric measure theory [10, Thm. 2.65] implies that the 2-rectifiable set  $\mathcal{G}$  we just constructed is "visible" from almost any direction, in the sense of the projection of  $\mathcal{G}$  onto a 2-dimensional linear subspace in general position having positive Lebesgue measure. However, there is not a single 2-dimensional linear subspace whose intersection with  $\mathcal{G}$  has positive Lebesgue measure. For if there was such a subspace  $\mathcal{T}$ , the Steinhaus theorem [4], [12] would imply that  $\mathcal{G} \ominus \mathcal{G} = \{u - v \mid u, v \in \mathcal{G}\}$  contains a 2dimensional ball centered at zero, which would in turn imply that  $\ker(f) \cap (\mathcal{G} \ominus \mathcal{G}) \neq \{0\}$ . Therefore, the linear mapping f would fail to be one-to-one on  $\mathcal{G}$ .

### A. Analytic random vectors

We just demonstrated that, for s-rectifiable random vectors, s is not a recovery threshold in general and additional requirements on  $\mathbf{x}$  are needed to get converse statements of the form of what we would expect from classical compressed sensing theory. This leads us to the new concept of s-analytic measures and s-analytic random vectors. We start with the definition of real analytic mappings.

Definition 8. We call

- (i) a function f : ℝ<sup>k</sup> → ℝ real analytic if, for each x ∈ ℝ<sup>k</sup>, f may be represented by a convergent power series in some neighborhood of x;
- (ii) a mapping  $f : \mathbb{R}^k \to \mathbb{R}^l$ ,  $\boldsymbol{x} \mapsto (f_1(\boldsymbol{x}) \dots f_l(\boldsymbol{x}))^{\mathsf{T}}$  real analytic if each component  $f_i$ ,  $i = 1, \dots, l$ , is a real analytic function.

We are now ready to define the notion of *s*-analytic measures and *s*-analytic random vectors.

Definition 9. We call a Borel measure  $\mu$  on  $\mathbb{R}^m$  s-analytic if for each  $\mathcal{U} \subseteq \mathbb{R}^m$  with  $\mu(\mathcal{U}) > 0$  we can find a real analytic mapping  $h : \mathbb{R}^s \to \mathbb{R}^m$  of s-dimensional Jacobian  $Jh \neq 0$ and a set  $\mathcal{A} \subseteq \mathbb{R}^s$  of positive Lebesgue measure such that  $h(\mathcal{A}) \subseteq \mathcal{U}$ .

Definition 10. The random vector  $\mathbf{x} \in \mathbb{R}^m$  is called *s*-analytic if  $\mu_{\mathbf{x}}$  is *s*-analytic.

It is instructive to compare s-analytic sets  $\mathcal{U}$  with  $\mu(\mathcal{U}) > 0$ to the set  $\mathcal{G}$  in Example 2. Both  $\mathcal{U}$  and  $\mathcal{G}$  contain the image of a set with positive Lebesgue measure under a certain mapping. However, the mapping in Example 2 is  $C^1$ , whereas the mapping in Definition 9 is real-analytic (with  $Jh \neq 0$ ). It turns out that real analyticity is strong enough to prevent  $\mathcal{U}$ from being mapped linearly in a one-to-one fashion into  $\mathbb{R}^t$  for t < s. Since this holds for every set  $\mathcal{U}$  with  $\mu(\mathcal{U}) > 0$ ,  $n \geq s$ is necessary for P2 to hold for s-analytic **x**. For if there exists an  $\varepsilon \in [0, 1)$ , an  $\mathbf{A} \in \mathbb{R}^{n \times m}$ , and a decoder  $g_{\mathbf{A}}$  satisfying (1), there must be a set  $\mathcal{U} \subseteq \mathbb{R}^m$  with  $\mathbb{P}[\mathbf{x} \in \mathcal{U}] \geq 1 - \varepsilon$  such that  $\mathbf{A}$  is one-to-one on  $\mathcal{U}$ , which is not possible for n < sbecause of the analyticity of  $\mu$ .

We are now ready to state our converse result for *s*-analytic random vectors.

**Theorem 2.** Let  $f : \mathbb{R}^m \to \mathbb{R}^n$  be a linear mapping,  $h : \mathbb{R}^s \to \mathbb{R}^m$  a real analytic mapping of s-dimensional Jacobian  $Jh \neq 0$ , and  $A \subseteq \mathbb{R}^s$  of positive Lebesgue measure. Suppose that f is one-to-one on h(A). Then  $n \geq s$ .

**Corollary 2.** For  $\mathbf{x} \in \mathbb{R}^m$  s-analytic,  $n \ge s$  is necessary for P2 to hold.

We close this section by establishing important properties of *s*-analytic measures and *s*-analytic random vectors, which will be used in the examples in the next section.

Lemma 3. Let x be s-analytic. Then, x is

(i) t-analytic for all  $t \in \{1, \ldots, s\}$ ;

(ii) s-rectifiable if it has an s-rectifiable support set.

**Lemma 4.** Suppose that  $\mu$  is s-analytic. Then  $\mu \ll \mathcal{H}^s$ .

## IV. EXAMPLES

*Example* 3. Let  $\mathbf{x} \in \mathbb{R}^m$  be as in Example 1. Using the properties of the Gaussian distribution, a straightforward analysis reveals that  $\mathbf{x}$  is *s*-analytic. Furthermore, the *s*-rectifiable set  $\mathcal{U}$  in (3) is a support set of  $\mathbf{x}$ . Therefore, by (ii) in Lemma 3,  $\mathbf{x}$  is *s*-rectifiable. It follows from Corollary 1 that n > s is sufficient for P1 to hold and from Corollary 2 that  $n \ge s$  is necessary for P2 to hold. The information-theoretic limit we obtain here is best possible in the sense of showing that not knowing the support set a priori does not have an impact on the recovery threshold. Classical compressed sensing recovery thresholds suffer either from the square-root bottleneck or from a  $\log(m)$ -factor. We hasten to add, however, that we do not specify decoders that achieve our threshold, rather we only prove the existence of such decoders.

In the second example, we construct an (r+s-1)-rectifiable and (r+s-1)-analytic random vector with sparsity level in terms of the number of non-zero entries of the vector's realizations— $rs \gg (r+s-1)$ .

*Example* 4. Let  $\mathbf{x} = \mathbf{a} \otimes \mathbf{b} \in \mathbb{R}^{kl}$ , where  $\mathbf{a} \in \mathbb{R}^k$ ,  $\mathbf{b} \in \mathbb{R}^l$ , and  $\mathbf{a}$  and  $\mathbf{b}$  are statistically independent. Suppose that  $\mathbf{a}$  has r i.i.d. Gaussian entries at positions drawn uniformly at random and all other entries equal to zero and  $\mathbf{b}$  has t i.i.d. Gaussian entries at positions drawn uniformly at random and all other entries equal to zero. Lemma 5 below shows that  $\mathbf{x}$  is (r + t - 1)-analytic. Furthermore, a straightforward analysis reveals that  $\mathbf{x}$  has the (r + t - 1)-rectifiable support set

with

$$\tilde{\mathcal{A}}_r = \{ \boldsymbol{a} \in \mathbb{R}^k : \|\boldsymbol{a}\|_0 = r, a_{nz} = 1 \}$$
$$\mathcal{B}_t = \{ \boldsymbol{b} \in \mathbb{R}^l : \|\boldsymbol{b}\|_0 = t \},$$

 $\mathcal{U} = \{ \boldsymbol{a} \otimes \boldsymbol{b} : \boldsymbol{a} \in \tilde{\mathcal{A}}_r, \boldsymbol{b} \in \mathcal{B}_t \}$ 

where  $a_{nz}$  denotes the first non-zero entry of a. By (ii) in Lemma 3,  $\mathbf{x}$  is (r+t-1)-rectifiable. It therefore follows from Corollary 1 that n > (r+t-1) is sufficient for P1 to hold and from Corollary 2 that  $n \ge (r+t-1)$  is necessary for P2 to hold. Note that, for r, t large, we have  $(r+t-1) \ll rt$ . What is interesting here is that the sparsity level of  $\mathbf{x}$ —as quantified by the number of non-zero entries of the realizations of  $\mathbf{x}$ —is rt, yet r+t linear measurements suffice for recovery of  $\mathbf{x}$  with zero probability of error.

**Lemma 5.** Let  $\mathbf{x} = \mathbf{a} \otimes \mathbf{b} \in \mathbb{R}^{kl}$ , where  $\mathbf{a} \in \mathbb{R}^k$  and  $\mathbf{b} \in \mathbb{R}^l$  are random vectors such that  $\mu_{\mathbf{a}} \times \mu_{\mathbf{b}} \ll \mathscr{L}^{k+l}$ . Then,  $\mathbf{x}$  is (k+l-1)-analytic.

#### REFERENCES

- D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
- [2] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.
- [3] S. Foucart and H. Rauhut, A Mathematical Introduction to Compressive Sensing. Basel, Switzerland: Birkhäuser, 2013.
- [4] Y. Wu and S. Verdú, "Rényi information dimension: Fundamental limits of almost lossless analog compression," *IEEE Trans. Inf. Theory*, vol. 56, no. 8, pp. 3721–3748, Aug. 2010.
- [5] —, "Optimal phase transitions in compressed sensing," *IEEE Trans. Inf. Theory*, vol. 58, no. 10, pp. 6241–6263, Oct. 2012.
- [6] D. Stotz, E. Riegler, E. Agustsson, and H. Bölcskei, "Almost lossless analog signal separation and probabilistic uncertainty relations," *IEEE Trans. Inf. Theory*, 2015, submitted, arXiv:1512.01017 [cs.IT].
- [7] G. Koliander, G. Pichler, E. Riegler, and F. Hlawatsch, "Entropy and source coding for integer-dimensional singular random variables," *IEEE Trans. Inf. Theory*, 2015, submitted, arXiv:1505.03337 [cs.IT].
- [8] C. De Lellis, *Rectifiable Sets, Densities and Tangent Measures*. Zurich, Switzerland: European Mathematical Society Publishing House, 2008.
- [9] K. Falconer, Fractal Geometry, 1st ed. New York, NY: Wiley, 1990.
- [10] L. Ambrosio, N. Fusco, and D. Pallara, Functions of Bounded Variation and Free Discontinuity Problems. New York, NY: Oxford Univ. Press, 2000.
- [11] P. K. Sahoo and T. Riedel, *Mean Value Theorems and Functional Equations*. Singapore: World Scientific Pub. Co., 1998.
- [12] K. Stromberg, "An elementary proof of Steinhaus's theorem," in *Proc.* of the American Mathematical Society, vol. 36, no. 1, 1976, p. 308.