Lossless Analog Compression

Giovanni Alberti, Helmut Bölcskei, Camillo De Lellis,
Günther Koliander, and Erwin Riegler

***Abstract*—We establish the fundamental limits of lossless analog compression by considering the recovery of arbitrary random vectors $\mathsf{x} \in \mathbb{R}^m$ from the noiseless linear measurements $\mathsf{y} = \boldsymbol{A}\mathsf{x}$ with measurement matrix $\boldsymbol{A} \in \mathbb{R}^{n \times m}$. Our theory is inspired by the groundbreaking work of Wu and Verdú (2010) on almost lossless analog compression, but applies to the nonasymptotic, i.e., fixed-m case, and considers zero error probability. Specifically, our achievability result states that, for Lebesgue-almost all $\boldsymbol{A}$, the random vector x can be recovered with zero error probability provided that $n > K(\mathsf{x})$, where $K(\mathsf{x})$ is given by the infimum of the lower modified Minkowski dimensions over all support sets $\mathcal{U}$ of x (i.e., sets $\mathcal{U} \subseteq \mathbb{R}^m$ with $\mathrm{P}[\mathsf{x} \in \mathcal{U}] = 1$). We then particularize this achievability result to the class of s-rectifiable random vectors as introduced in Koliander *et al.* (2016); these are random vectors of absolutely continuous distribution—with respect to the s-dimensional Hausdorff measure—supported on countable unions of s-dimensional C^1-submanifolds of $\mathbb{R}^m$. Countable unions of C^1-submanifolds include essentially all signal models used in the compressed sensing literature such as the standard union of subspaces model underlying much of compressed sensing theory and spectrum-blind sampling, smooth manifolds, block-sparsity, and low-rank matrices as considered in the matrix completion problem. Specifically, we prove that, for Lebesgue-almost all $\boldsymbol{A}$, s-rectifiable random vectors x can be recovered with zero error probability from $n > s$ linear measurements. This threshold is, however, found not to be tight as exemplified by the construction of an s-rectifiable random vector that can be recovered with zero error probability from $n < s$ linear measurements. Motivated by this observation, we introduce the new class of s-analytic random vectors, which admit a strong converse in the sense of $n \geq s$ being necessary for recovery with probability of error smaller than one. The central conceptual tools in the development of our theory are geometric measure theory and the theory of real analytic functions.**

I. Introduction

Compressed sensing [2]–[6] deals with the recovery of unknown sparse vectors $\boldsymbol{x} \in \mathbb{R}^m$ from a small (relative to m) number, n, of linear measurements of the form $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$, where $\boldsymbol{A} \in \mathbb{R}^{n \times m}$ is the measurement matrix.[1] Known recovery guarantees can be categorized as deterministic, probabilistic, and information-theoretic. The literature in all three categories is abundant and the ensuing overview is hence necessarily highly incomplete, yet representative.

Deterministic results, such as those in [2], [6]–[10], are uniform in the sense of applying to all s-sparse vectors $\boldsymbol{x} \in \mathbb{R}^m$, i.e., vectors $\boldsymbol{x}$ that are supported on a finite union of s-dimensional linear subspaces of $\mathbb{R}^m$, and for a fixed measurement matrix $\boldsymbol{A}$. Typical guarantees say that s-sparse vectors $\boldsymbol{x}$ can be recovered through convex optimization algorithms or greedy algorithms provided that [11, Chapter 3] $s < \frac{1}{2}(1 + 1/\mu)$, where μ denotes the coherence of $\boldsymbol{A}$, i.e., the largest (in absolute value) inner product of any two different columns of $\boldsymbol{A}$. The Welch bound [12] implies that the minimum number of linear measurements, n, required for uniform recovery is of order s^2, a result known as the "square-root bottleneck" all coherence-based recovery thresholds suffer from.

Probabilistic results are either based on random measurement matrices $\mathbf{A}$ ([3]–[5], [13]–[16]) or deterministic $\boldsymbol{A}$ and random s-sparse vectors x ([14], [15]), and typically state that s-sparse vectors can be recovered, again using convex optimization algorithms or greedy algorithms, with high probability, provided that n is of order $s \log m$.

An information-theoretic framework for compressed sensing, fashioned as an almost lossless analog compression problem, was developed by Wu and Verdú [17], [18]. Specifically, [17] derives asymptotic (in m) achievability results and converses for linear encoders and measurable decoders, measurable encoders and Lipschitz continuous decoders, and continuous encoders and continuous decoders. For the particular case of linear encoders and measurable decoders, [17] shows that, asymptotically in m, for Lebesgue almost all (a.a.) measurement matrices $\boldsymbol{A}$, the random vector x can be recovered with arbitrarily small probability of error from $n = \lfloor Rm \rfloor$ linear measurements, provided that $R > R_{\mathrm{B}}$, where R_{B} denotes the Minkowski dimension compression rate [17, Definition 10] of the random process generating x. For the special case of x with independent and identically distributed (i.i.d.) discrete-continuous mixture entries, a matching converse exists. Discrete-continuous mixture distributions $\rho\mu^{\mathrm{c}} + (1-\rho)\mu^{\mathrm{d}}$ are relevant as they mimic sparse vectors for large m. In particular, if the discrete part μ^{d} is a Dirac measure at 0, then the nonzero entries of x can be generated only by the continuous part μ^{c}, and the fraction of nonzero entries in x converges—in probability—to ρ as m tends to infinity. A nonasymptotic, i.e., fixed-m, statement in [17] says that a.a. (with respect to a σ-finite Borel measure on $\mathbb{R}^m$) s-sparse random vectors can be recovered with zero error probability provided that $n > s$. Again, this result holds for Lebesgue a.a. measurement matrices $\boldsymbol{A} \in \mathbb{R}^{n \times m}$. A corresponding converse does not seem to be available. For recent work on

G. Alberti is with the Dept. of Mathematics, University of Pisa, Italy (e-mail: galberti1@dm.unipi.it).

H. Bölcskei is with the Dept. of IT & EE and the Dept. of Mathematics, ETH Zurich, Switzerland (e-mail: hboelcskei@ethz.ch).

C. De Lellis is with the IAS, Princeton, New Jersey and the University of Zurich, Switzerland (e-mail: camillo.delellis@math.ias.edu).

G. Koliander is with the ARI, Austrian Academy of Sciences, Vienna, Austria (e-mail: guenther.koliander@kfs.oeaw.ac.at). His work was supported by the Vienna Science and Technology Fund (WWTF): MA16-053.

E. Riegler is with the Dept. of IT & EE, ETH Zurich, Switzerland (e-mail: eriegler@mins.ee.ethz.ch).

The material in this paper was presented in part at the 2016 IEEE International Symposium on Information Theory [1].

[1] Throughout the paper, we assume that $n \leq m$.

the connection between lossy data compression of stochastic processes under distortion constraints and mean dimension theory for dynamical systems, we refer the interested reader to [19]–[21].

Contributions. We establish the fundamental limits of lossless, i.e., zero error probability, analog compression in the nonasymptotic, i.e., fixed-m, regime for arbitrary random vectors $\mathsf{x} \in \mathbb{R}^m$. Specifically, we show that x can be recovered with zero error probability provided that $n > K(\mathsf{x})$, with $K(\mathsf{x})$ given by the infimum of the lower modified Minkowski dimensions over all support sets $\mathcal{U}$ of x, i.e., all sets $\mathcal{U} \subseteq \mathbb{R}^m$ with $\mathrm{P}[\mathsf{x} \in \mathcal{U}] = 1$. This statement holds for Lebesgue-a.a. measurement matrices. Lower modified Minkowski dimension vastly generalizes the notion of s-sparsity, and allows for arbitrary support sets that are not necessarily finite unions of s-dimensional linear subspaces. For s-sparse vectors, we get the recovery guarantee $n > s$ showing that our information-theoretic thresholds suffer neither from the square-root bottleneck [12] nor from a $\log m$-factor [14], [15]. We hasten to add, however, that we do not specify explicit decoders that achieve these thresholds, rather we provide existence results absent computational considerations. The central conceptual element in the proof of our achievability result is the probabilistic null-space property first reported in [22]. We emphasize that it is the usage of modified Minkowski dimension, as opposed to Minkowski dimension [17], [22], that allows us to obtain achievability results for zero error probability. The asymptotic achievability result for linear encoders in [17] can be recovered in our framework.

We particularize our achievability result to s-rectifiable random vectors x as introduced in [23]; these are random vectors supported on countable unions of s-dimensional C^1-submanifolds of $\mathbb{R}^m$ and of absolutely continuous—with respect to s-dimensional Hausdorff measure—distribution. Countable unions of C^1-submanifolds include numerous signal models prevalent in the compressed sensing literature, namely, the standard union of subspaces model underlying much of compressed sensing theory [24], [25] and spectrum-blind sampling [26], [27], smooth manifolds [28], block-sparsity [29]–[31], and low-rank matrices as considered in the matrix completion problem [32]–[34]. Our achievability result shows that s-rectifiable random vectors can be recovered with zero error probability provided that $n > s$. Again, this statement holds for Lebesgue-a.a. measurement matrices. Absolute continuity with respect to s-dimensional Hausdorff measure is a regularity condition ensuring that the distribution is not too concentrated; in particular, sets of Hausdorff dimension $t < s$ are guaranteed to carry zero probability mass. One would therefore expect $n \geq s$ to be necessary for zero error recovery of s-rectifiable random vectors. It turns out, however, that, perhaps surprisingly, this is not the case in general. An example elucidating this phenomenon constructs a set $\mathcal{G} \subseteq \mathbb{R}^2$ of positive 2-dimensional Hausdorff measure that can be compressed linearly in a one-to-one fashion into $\mathbb{R}$. This will then be seen to lead to the statement that every 2-rectifiable random vector of distribution absolutely continuous with respect to 2-dimensional Hausdorff measure restricted to $\mathcal{G}$ can be recovered with zero error probability from a single linear measurement. What renders this result surprising is that all this is possible even though $\mathcal{G}$ contains the image of a Borel set in $\mathbb{R}^2$ of positive Lebesgue measure under a C^∞-embedding. The picture changes completely when the embedding is real analytic. Specifically, we show that if a set $\mathcal{U} \subseteq \mathbb{R}^m$ contains the real analytic embedding of a Borel set in $\mathbb{R}^s$ of positive Lebesgue measure, it cannot be compressed linearly (in fact, not even through a nonlinear real analytic mapping) in a one-to-one fashion into $\mathbb{R}^n$ with $n < s$. This leads to the new concept of s-analytic random vectors, which allows a strong converse in the sense of $n \geq s$ being necessary for recovery of x with probability of error smaller than one. The qualifier "strong" refers to the fact that recovery from $n < s$ linear measurements is not possible even if we allow an arbitrary positive error probability strictly smaller than one. The only strong converse available in the literature applies to random vectors x with i.i.d. discrete-continuous mixture entries [17].

Organization of the paper. In Section II, we present our achievability results, with the central statement in Theorem II.1. Section III particularizes these results to s-rectifiable random vectors, and presents an example of a 2-rectifiable random vector that can be recovered from a single linear measurement with zero error probability. In Section IV, we introduce and characterize the new class of s-analytic random vectors and we derive a corresponding strong converse, stated in Theorem IV.1. Sections V–VII contain the proofs of the main technical results stated in Sections II–IV. Appendices A–G contain proofs of further technical results stated in the main body of the paper. In Appendices H–K, we summarize concepts and basic results from (geometric) measure theory, the theory of set-valued functions, sequences of functions in several variables, and real analytic mappings, all needed throughout the paper. The reader not familiar with these results is advised to consult the corresponding appendices before studying the proofs of our main results. These appendices also contain new results, which are highly specific and would disrupt the flow of the paper if presented in the main body.

Notation. We use capital boldface roman letters $\boldsymbol{A}, \boldsymbol{B}, \dots$ to denote deterministic matrices and lower-case boldface roman letters $\boldsymbol{a}, \boldsymbol{b}, \dots$ to designate deterministic vectors. Random matrices and vectors are set in sans-serif font, e.g., A and x. The $m \times m$ identity matrix is denoted by $\boldsymbol{I}_m$. We write $\mathrm{rank}(\boldsymbol{A})$ and $\ker(\boldsymbol{A})$ for the rank and the kernel of $\boldsymbol{A}$, respectively. The superscript $^\mathsf{T}$ stands for transposition. The i-th unit vector is denoted by $\boldsymbol{e}_i$. For a vector $\boldsymbol{x} \in \mathbb{R}^m$, $\|\boldsymbol{x}\|_2 = \sqrt{\boldsymbol{x}^\mathsf{T}\boldsymbol{x}}$ is its Euclidean norm and $\|\boldsymbol{x}\|_0$ denotes the number of nonzero entries of $\boldsymbol{x}$. For the set $\mathcal{A}$, we write $\mathrm{card}(\mathcal{A})$ for its cardinality, $\overline{\mathcal{A}}$ for its closure, $2^{\mathcal{A}}$ for its power set, and $\mathbb{1}_{\mathcal{A}}$ for the indicator function on $\mathcal{A}$. With $\mathcal{A} \subseteq \mathbb{R}^m$ and $\mathcal{B} \subseteq \mathbb{R}^n$, we let $\mathcal{A} \times \mathcal{B} = \{(\boldsymbol{a}, \boldsymbol{b}) : \boldsymbol{a} \in \mathcal{A}, \boldsymbol{b} \in \mathcal{B}\}$ and $\mathcal{A} \otimes \mathcal{B} = \{\boldsymbol{a} \otimes \boldsymbol{b} : \boldsymbol{a} \in \mathcal{A}, \boldsymbol{b} \in \mathcal{B}\}$, where $\otimes$ denotes the Kronecker product (see [35, Definition 4.2.1]). For $\mathcal{A}, \mathcal{B} \subseteq \mathbb{R}^m$, we write $\mathcal{A} \subsetneq \mathcal{B}$ to express strict inclusion according to $\mathcal{A} \subseteq \mathcal{B}$ with $\mathcal{A} \neq \mathcal{B}$ and we let $\mathcal{A} - \mathcal{B} = \{\boldsymbol{a} - \boldsymbol{b} : \boldsymbol{a} \in \mathcal{A}, \boldsymbol{b} \in \mathcal{B}\}$. We set $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$. For the Euclidean space $(\mathbb{R}^k, \|\cdot\|_2)$, we designate the open ball of radius ρ centered at $\boldsymbol{u} \in \mathbb{R}^k$ by $\mathcal{B}_k(\boldsymbol{u}, \rho)$. We write $\mathscr{S}(\mathcal{X})$ for a general σ-algebra on $\mathcal{X}$,

$\mathscr{B}(\mathcal{X})$ for the Borel σ-algebra on a topological space $\mathcal{X}$, and $\mathscr{L}(\mathbb{R}^m)$ for the Lebesgue σ-algebra on $\mathbb{R}^m$. The product σ-algebra of $\mathscr{S}(\mathcal{X})$ and $\mathscr{S}(\mathcal{Y})$ is denoted by $\mathscr{S}(\mathcal{X})\otimes\mathscr{S}(\mathcal{Y})$. For measures μ and ν on the same measurable space, we denote absolute continuity of μ with respect to ν by $\mu \ll \nu$. We write $\mu\times\nu$ for the product measure of μ and ν. Throughout we assume, without loss of generality (w.l.o.g.), that a measure on a measurable space is defined on all subsets of the measurable space (see [36, Remark 1.2.6]). The Lebesgue measure on $\mathbb{R}^k$ and $\mathbb{R}^{k\times l}$ is designated as λ^k and $\lambda^{k\times l}$, respectively. The distribution of a random vector x is denoted by μ_{x}. If $f\colon \mathbb{R}^k \to \mathbb{R}^l$ is differentiable, we write $Df(\boldsymbol{v}) \in \mathbb{R}^{l\times k}$ for its differential at $\boldsymbol{v} \in \mathbb{R}^k$ and define the $\min\{k,l\}$-dimensional Jacobian $Jf(\boldsymbol{v})$ at $\boldsymbol{v} \in \mathbb{R}^k$ by

$$Jf(\boldsymbol{v}) = \begin{cases} \sqrt{\det((Df(\boldsymbol{v}))^{\mathsf{T}} Df(\boldsymbol{v}))} & \text{if } l \geq k \\ \sqrt{\det(Df(\boldsymbol{v})(Df(\boldsymbol{v}))^{\mathsf{T}})} & \text{else.} \end{cases} \tag{1}$$

For an open set $\mathcal{U} \subseteq \mathbb{R}^k$, a differentiable mapping $f\colon \mathcal{U} \to \mathbb{R}^l$, where $l \geq k$, is called an immersion if $Jf(\boldsymbol{v}) > 0$ for all $\boldsymbol{v} \in \mathcal{U}$. A one-to-one immersion is referred to as an embedding. For a mapping f, we write $f \equiv \mathbf{0}$ if it is identically zero and $f \not\equiv \mathbf{0}$ otherwise. For $f\colon \mathcal{U} \to \mathcal{V}$ and $g\colon \mathcal{V} \to \mathcal{W}$, the composition $g\circ f\colon \mathcal{U} \to \mathcal{W}$ is defined as $(g\circ f)(x) = g(f(x))$ for all $x \in \mathcal{U}$. For $f\colon \mathcal{U} \to \mathcal{V}$ and $\mathcal{A} \subseteq \mathcal{U}$, $f|_{\mathcal{A}}$ denotes the restriction of f to $\mathcal{A}$. For $f\colon \mathcal{U} \to \mathcal{V}$ and $\mathcal{B} \subseteq \mathcal{V}$, we set $f^{-1}(\mathcal{B}) = \{x \in \mathcal{U} : f(x) \in \mathcal{B}\}$.

II. Achievability

In classical compressed sensing theory [3]–[6], one typically deals with the recovery of s-sparse vectors $\boldsymbol{x} \in \mathbb{R}^m$, i.e., vectors $\boldsymbol{x}$ that are supported on a finite union of s-dimensional linear subspaces of $\mathbb{R}^m$. The purpose of this paper is the development of a comprehensive theory of signal recovery in the sense of allowing arbitrary support sets $\mathcal{U}$, which are not necessarily unions of (a finite number of) linear subspaces of $\mathbb{R}^m$. Formalizing this idea requires a suitable dimension measure for general nonempty sets. There is a rich variety of dimension measures available in the literature [37]–[39]. Our choice will be guided by the requirement of information-theoretic operational significance. Specifically, the dimension measure should allow the formulation of nonasymptotic, i.e., fixed-m, achievability results with zero error probability. The modified Minkowski dimension will turn out to meet these requirements.

We first recall the definitions of Minkowski dimension and modified Minkowski dimension, compare the two concepts, and state the basic properties of modified Minkowski dimension needed in the remainder of the paper.

Definition II.1. (Minkowski dimension[2]) [38, Equivalent definitions 2.1] For $\mathcal{U} \subseteq \mathbb{R}^m$ nonempty, the lower and upper Minkowski dimension of $\mathcal{U}$ is defined as

$$\underline{\dim}_{\mathrm{B}}(\mathcal{U}) = \liminf_{\rho\to 0} \frac{\log N_{\mathcal{U}}(\rho)}{\log \frac{1}{\rho}} \tag{2}$$

and

$$\overline{\dim}_{\mathrm{B}}(\mathcal{U}) = \limsup_{\rho\to 0} \frac{\log N_{\mathcal{U}}(\rho)}{\log \frac{1}{\rho}}, \tag{3}$$

respectively, where

$$N_{\mathcal{U}}(\rho) = \min\left\{ k \in \mathbb{N} : \mathcal{U} \subseteq \bigcup_{i\in\{1,\dots,k\}} \mathcal{B}_m(\boldsymbol{u}_i,\rho),\ \boldsymbol{u}_i \in \mathbb{R}^m \right\} \tag{4}$$

is the covering number of $\mathcal{U}$ for radius ρ. If $\underline{\dim}_{\mathrm{B}}(\mathcal{U}) = \overline{\dim}_{\mathrm{B}}(\mathcal{U})$, this common value, denoted by $\dim_{\mathrm{B}}(\mathcal{U})$, is the Minkowski dimension of $\mathcal{U}$.

Definition II.2. (Modified Minkowski dimension) [38, p. 37] For $\mathcal{U} \subseteq \mathbb{R}^m$ nonempty, the lower and upper modified Minkowski dimension of $\mathcal{U}$ is defined as

$$\underline{\dim}_{\mathrm{MB}}(\mathcal{U}) = \inf\left\{ \sup_{i\in\mathbb{N}} \underline{\dim}_{\mathrm{B}}(\mathcal{U}_i) : \mathcal{U} \subseteq \bigcup_{i\in\mathbb{N}} \mathcal{U}_i \right\} \tag{5}$$

and

$$\overline{\dim}_{\mathrm{MB}}(\mathcal{U}) = \inf\left\{ \sup_{i\in\mathbb{N}} \overline{\dim}_{\mathrm{B}}(\mathcal{U}_i) : \mathcal{U} \subseteq \bigcup_{i\in\mathbb{N}} \mathcal{U}_i \right\}, \tag{6}$$

respectively, where in (5) and (6) the infima are over all possible coverings $\{\mathcal{U}_i\}_{i\in\mathbb{N}}$ of $\mathcal{U}$ by nonempty bounded sets $\mathcal{U}_i$. If $\underline{\dim}_{\mathrm{MB}}(\mathcal{U}) = \overline{\dim}_{\mathrm{MB}}(\mathcal{U})$, this common value, denoted by $\dim_{\mathrm{MB}}(\mathcal{U})$, is the modified Minkowski dimension of $\mathcal{U}$.

The main properties of modified Minkowski dimension are summarized in Lemma H.15. In particular,

$$\underline{\dim}_{\mathrm{MB}}(\cdot) \leq \underline{\dim}_{\mathrm{B}}(\cdot) \tag{7}$$

and

$$\overline{\dim}_{\mathrm{MB}}(\cdot) \leq \overline{\dim}_{\mathrm{B}}(\cdot). \tag{8}$$

Both lower and upper modified Minkowski dimension have the advantage of being countably stable, a key property we will use frequently. In contrast, upper Minkowski dimension is only finitely stable, and lower Minkowski dimension is not even finitely stable (see [38, p. 34]). For example, all countable subsets of $\mathbb{R}^m$ have modified Minkowski dimension equal to zero (the Minkowski dimension of a single point in $\mathbb{R}^m$ equals zero), but there exist infinitely countable sets with nonzero Minkowski dimension:

Example II.1. [38, Example 2.7] Let $\mathcal{F} = \{0, 1/2, 1/3, \dots\}$. Then, $\dim_{\mathrm{MB}}(\mathcal{F}) = 0 < \dim_{\mathrm{B}}(\mathcal{F}) = 1/2$.

Minkowski dimension and modified Minkowski dimension also behave differently for unbounded sets. Specifically, by monotonicity of (upper) modified Minkowski dimension, $\underline{\dim}_{\mathrm{MB}}(\mathcal{A}) \leq \overline{\dim}_{\mathrm{MB}}(\mathcal{A}) \leq \dim_{\mathrm{MB}}(\mathbb{R}^m) = m$ for all $\mathcal{A} \subseteq \mathbb{R}^m$, in particular also for unbounded sets, whereas $\underline{\dim}_{\mathrm{B}}(\mathcal{A}) = \overline{\dim}_{\mathrm{B}}(\mathcal{A}) = \infty$ for all unbounded sets $\mathcal{A}$ as a consequence of $N_{\mathcal{A}}(\rho) = \infty$ for all $\rho \in (0,\infty)$. Working with lower modified Minkowski dimension will allow us to consider arbitrary random vectors, regardless of whether they admit bounded support sets or not.

The following example shows that the modified Minkowski dimension agrees with the sparsity notion used in classical compressed sensing theory.

[2]Minkowski dimension is sometimes also referred to as box-counting dimension, which is the origin of the subscript B in the notation $\dim_{\mathrm{B}}(\cdot)$ used henceforth.

Example II.2. For $\mathcal{I}$ finite or countable infinite, let $\mathcal{T}_i$, $i \in \mathcal{I}$, be linear subspaces with their Euclidean dimensions $\dim(\mathcal{T}_i)$ satisfying

$$\max_{i\in\mathcal{I}} \dim(\mathcal{T}_i) = s, \tag{9}$$

and consider the union of subspaces

$$\mathcal{U} = \bigcup_{i\in\mathcal{I}} \mathcal{T}_i. \tag{10}$$

As every linear subspace is a smooth submanifold of $\mathbb{R}^m$, it follows from Properties ii) and vi) of Lemma H.15 that $\dim_{\mathrm{MB}}(\mathcal{U}) = s$. In the union of subspaces model, prevalent in compressed sensing theory [3]–[6], $|\mathcal{I}| = \binom{m}{s}$ and the subspaces $\mathcal{T}_i$ correspond to different sparsity patterns, each of cardinality equal to the sparsity level s.

The aim of the present paper is to develop a theory for lossless analog compression of arbitrary random vectors $\mathbf{x} \in \mathbb{R}^m$. An obvious choice for the stochastic equivalent of the sparsity level s is the stochastic sparsity level $S(\mathbf{x})$ defined as

$$S(\mathbf{x}) = \min\left\{ s : \exists \mathcal{T}_1, \dots, \mathcal{T}_k \text{ with } \mathrm{P}\left[\mathbf{x} \in \bigcup_{i=1}^{k} \mathcal{T}_i \right] = 1 \right\}, \tag{11}$$

where every $\mathcal{T}_i$ is a linear subspace of $\mathbb{R}^m$ of dimension $\dim(\mathcal{T}_i) \le s$ and $k \in \mathbb{N}$. This definition is, however, specific to the (finite) union of linear subspaces structure. The theory we develop here requires a more general notion of description complexity, which we define in terms of the lower modified Minkowski dimension according to

$$K(\mathbf{x}) = \inf\{\underline{\dim}_{\mathrm{MB}}(\mathcal{U}) : \mathcal{U} \subseteq \mathbb{R}^m, \mathrm{P}[\mathbf{x} \in \mathcal{U}] = 1\}. \tag{12}$$

Sets satisfying $\mathrm{P}[\mathbf{x} \in \mathcal{U}] = 1$ are hereafter referred to as support sets of $\mathbf{x}$. While the definition of $S(\mathbf{x})$ involves minimization of Euclidean dimensions of linear subspaces, $K(\mathbf{x})$ is defined by minimizing lower modified Minkowski dimensions of general support sets. Definitions (11) and (12) imply directly that $K(\mathbf{x}) \le S(\mathbf{x})$ for all random vectors $\mathbf{x} \in \mathbb{R}^m$ (see Example II.2). We will see in Section III that this inequality can actually be strict.

Next, we show that application of a locally Lipschitz mapping cannot increase a random vector's description complexity. This result will allow us to construct random vectors with low description complexity out of existing ones simply by applying locally Lipschitz mappings. The formal statement is as follows.

Lemma II.1. Let $\mathbf{x} \in \mathbb{R}^k$ and $f\colon \mathbb{R}^k \to \mathbb{R}^m$ be locally Lipschitz. Then, $K(f(\mathbf{x})) \le K(\mathbf{x})$.

Proof.

$$
\begin{aligned}
&K(f(\mathbf{x})) && (13)\\
&= \inf\{\underline{\dim}_{\mathrm{MB}}(\mathcal{V}) : \mathcal{V} \subseteq \mathbb{R}^m,\ \mathrm{P}[f(\mathbf{x}) \in \mathcal{V}] = 1\} && (14)\\
&= \inf\{\underline{\dim}_{\mathrm{MB}}(f(\mathcal{U})) : \mathcal{U} \subseteq \mathbb{R}^k,\ \mathrm{P}[\mathbf{x} \in \mathcal{U}] = 1\} && (15)\\
&\le \inf\{\underline{\dim}_{\mathrm{MB}}(\mathcal{U}) : \mathcal{U} \subseteq \mathbb{R}^k,\ \mathrm{P}[\mathbf{x} \in \mathcal{U}] = 1\} && (16)\\
&= K(\mathbf{x}), && (17)
\end{aligned}
$$

where (16) follows from Property vii) of Lemma H.15. □

When the mapping f is invertible and both f and f^{-1} are locally Lipschitz, the description complexity remains unchanged:

Corollary II.1. Let $\mathbf{x} \in \mathbb{R}^m$ and consider an invertible mapping $f\colon \mathbb{R}^m \to \mathbb{R}^m$. Suppose that f and f^{-1} are both locally Lipschitz. Then, $K(\mathbf{x}) = K(f(\mathbf{x}))$.

Proof. $K(\mathbf{x}) = K((f^{-1} \circ f)(\mathbf{x})) \le K(f(\mathbf{x})) \le K(\mathbf{x})$, where we applied Lemma II.1 twice. □

As a consequence of Corollary II.1, the description complexity $K(\mathbf{x})$ is invariant under a basis change. Our main achievability result can now be formulated as follows.

Theorem II.1. (Achievability) For $\mathbf{x} \in \mathbb{R}^m$, $n > K(\mathbf{x})$ is sufficient for the existence of a Borel measurable mapping $g\colon \mathbb{R}^{n\times m} \times \mathbb{R}^n \to \mathbb{R}^m$, referred to as (measurable) decoder, satisfying

$$\mathrm{P}[g(\boldsymbol{A}, \boldsymbol{A}\mathbf{x}) \neq \mathbf{x}] = 0 \quad \text{for } \lambda^{n\times m}\text{-a.a. } \boldsymbol{A} \in \mathbb{R}^{n\times m}. \tag{18}$$

Proof. See Section V. □

Theorem II.1 generalizes the achievability result for linear encoders in [17] in the sense of being nonasymptotic (i.e., it applies for finite m) and guaranteeing zero error probability.

The central conceptual element in the proof of Theorem II.1 is the following probabilistic null-space property for arbitrary (possibly unbounded) nonempty sets, first reported in [22] for bounded sets and expressed in terms of lower Minkowski dimension. If the lower modified Minkowski dimension of a nonempty set $\mathcal{U}$ is smaller than n, then, for $\lambda^{n\times m}$-a.a. measurement matrices $\boldsymbol{A} \in \mathbb{R}^{n\times m}$, the set $\mathcal{U}$ intersects the $(m-n)$-dimensional kernel of $\boldsymbol{A}$ at most trivially. What is remarkable here is that the notions of Euclidean dimension (for the kernel of the linear mapping induced by $\boldsymbol{A}$) and of lower modified Minkowski dimension (for $\mathcal{U}$) are compatible. The formal statement is as follows.

Proposition II.1. Let $\mathcal{U} \subseteq \mathbb{R}^m$ be nonempty with $\underline{\dim}_{\mathrm{MB}}(\mathcal{U}) < n$. Then,

$$\ker(\boldsymbol{A}) \cap (\mathcal{U}\backslash\{\mathbf{0}\}) = \emptyset \quad \text{for } \lambda^{n\times m}\text{-a.a. } \boldsymbol{A} \in \mathbb{R}^{n\times m}. \tag{19}$$

Proof. By definition of lower modified Minkowski dimension, there exists a covering $\{\mathcal{U}_i\}_{i\in\mathbb{N}}$ of $\mathcal{U}$ by nonempty compact sets $\mathcal{U}_i$ with $\underline{\dim}_{\mathrm{B}}(\mathcal{U}_i) < n$ for all $i \in \mathbb{N}$. The countable subadditivity of Lebesgue measure now implies that

$$
\begin{aligned}
&\lambda^{n\times m}\{\boldsymbol{A} \in \mathbb{R}^{n\times m} : \ker(\boldsymbol{A}) \cap (\mathcal{U}\backslash\{\mathbf{0}\}) \neq \emptyset\} && (20)\\
&\le \sum_{i\in\mathbb{N}} \lambda^{n\times m}\{\boldsymbol{A} \in \mathbb{R}^{n\times m} : \ker(\boldsymbol{A}) \cap (\mathcal{U}_i\backslash\{\mathbf{0}\}) \neq \emptyset\}. && (21)
\end{aligned}
$$

The proof is concluded by noting that [22, Proposition 1] with $\underline{\dim}_{\mathrm{B}}(\mathcal{U}_i) < n$ for all $i \in \mathbb{N}$ implies that every term in the sum of (21) equals zero. □

We close this section by elucidating the level of generality of our theory through particularization of the achievability result Theorem II.1 to random vectors supported on attractor sets of systems of contractions as defined below. The formal definition is as follows. Let $\mathcal{A} \subseteq \mathbb{R}^m$ be closed. For $i = 1, \dots, k$, consider $s_i : \mathcal{A} \to \mathcal{A}$ and $c_i \in (0,1)$ such that

$$\|s_i(\boldsymbol{u}) - s_i(\boldsymbol{v})\|_2 \le c_i \|\boldsymbol{u} - \boldsymbol{v}\|_2 \quad \text{for all } \boldsymbol{u}, \boldsymbol{v} \in \mathcal{A}. \tag{22}$$

Such mappings are called contractions. By [38, Theorem 9.1], there exists a unique compact set $\mathcal{K} \subseteq \mathcal{A}$, referred to as an attractor set, such that

$$\mathcal{K} = \bigcup_{i=1}^{k} s_i(\mathcal{K}). \tag{23}$$

Thanks to [38, Proposition 9.6], $\overline{\dim}_{\mathrm{B}}(\mathcal{K}) \le d$, where $d > 0$ is the unique solution of

$$\sum_{i=1}^{k} c_i^d = 1. \tag{24}$$

If, in addition, $\mathcal{K}$ satisfies the open set condition [38, Equation (9.12)], then $\dim_{\mathrm{B}}(\mathcal{K}) = d$ by [38, Theorem 9.3]. The middle-third Cantor set [38, Example 9.1], Sierpiński gaskets [38, Example 9.4], and the modified von Koch curves [38, Example 9.5] are all attractor sets (with their underlying contractions) that meet the open set condition [38, Chapter 9]. As $\underline{\dim}_{\mathrm{MB}}(\mathcal{K}) \le \overline{\dim}_{\mathrm{B}}(\mathcal{K})$ by Property iv) of Lemma H.15, every $\mathbf{x} \in \mathbb{R}^m$ such that $\mathrm{P}[\mathbf{x} \in \mathcal{K}] = 1$ has description complexity $K(\mathbf{x}) \le d$ (see (12)). For an excellent in-depth treatment of attractor sets of systems of contractions, the interested reader is referred to [38, Chapter 9]. We finally note that for self-similar distributions [17, Section 3.E] on attractor sets $\mathcal{K}$ satisfying the open set condition [38, Equation (9.12)], an achievability result in terms of information dimension was reported in [17, Theorem 7].

III. Rectifiable Sets and Rectifiable Random Vectors

The signal models employed in classical compressed sensing [3], [4], model-based compressed sensing [25], and block-sparsity [29]–[31] all fall under the rubric of finite unions of linear subspaces. More general prevalent signal models in the theory of sparse signal recovery include finite unions of smooth manifolds, either in explicit form as in [28] or implicitly in the context of low-rank matrix recovery [32]–[34]. All these models are subsumed by the countable unions of C^1-manifolds structure, formalized next using the notion of rectifiable sets. We start with the definition of rectifiable sets.

Definition III.1. (Rectifiable sets) [40, Definition 3.2.14] Let $s \in \mathbb{N}$, and consider a measure μ on $\mathbb{R}^m$. A nonempty set $\mathcal{U} \subseteq \mathbb{R}^m$ is

i) s-rectifiable if there exist a compact set $\mathcal{A} \subseteq \mathbb{R}^s$ and a Lipschitz mapping $\varphi : \mathcal{A} \to \mathbb{R}^m$ such that $\mathcal{U} = \varphi(\mathcal{A})$,
ii) countably s-rectifiable if it is the countable union of s-rectifiable sets,
iii) countably (μ, s)-rectifiable if it is μ-measurable and there exists a countably s-rectifiable set $\mathcal{V}$ such that $\mu(\mathcal{U} \backslash \mathcal{V}) = 0$.
iv) (μ, s)-rectifiable if it is countably (μ, s)-rectifiable and $\mu(\mathcal{U}) < \infty$.

Our definitions of s-rectifiability and countably s-rectifiability differ from those of [40, Definition 3.2.14] as we require the s-rectifiable set to be the Lipschitz image of a compact rather than a bounded set. The more restrictive definitions i) and ii) above have the advantage of s-rectifiable sets and countably s-rectifiable sets being guaranteed to be Borel. (This holds since the image of a compact set under a continuous mapping is a compact set, and a countable union of compact sets is Borel.) We note however that, by Lemma H.2, our definitions of (μ, s)-rectifiable and countably (μ, s)-rectifiable (see Definition III.1, Items iii) and iv)) are nevertheless equivalent to those in [40, Definition 3.2.14].

In what follows, we only need Items iii) and iv) in Definition III.1 for the specific case of $\mu = \mathscr{H}^s$ (see Definition H.3), in which measurability of $\mathcal{U}$ is guaranteed for all Borel sets. Therefore, s-rectifiable sets and countably s-rectifiable sets are also $\mathscr{H}^s$-measurable, which leads to the following chain of implications:

$\mathcal{U}$ is s-rectifiable $\Rightarrow$ $\mathcal{U}$ is countably s-rectifiable $\Rightarrow$ $\mathcal{U}$ is countably $(\mathscr{H}^s, s)$-rectifiable.

The following result collects properties of (countably) s-rectifiable sets for later use.

Lemma III.1.

i) If $\mathcal{U} \subseteq \mathbb{R}^m$ is s-rectifiable, then it is t-rectifiable for all $t \in \mathbb{N}$ with $t > s$.
ii) For locally Lipschitz φ_i, $i \in \mathbb{N}$, the set
$$\mathcal{V} = \bigcup_{i \in \mathbb{N}} \varphi_i(\mathbb{R}^s) \tag{25}$$
is countably s-rectifiable.
iii) If $\mathcal{U} \subseteq \mathbb{R}^m$ is countably s-rectifiable and $\mathcal{V} \subseteq \mathbb{R}^n$ is countably t-rectifiable, then
$$\mathcal{W} = \{(\boldsymbol{u}^\mathsf{T}\ \boldsymbol{v}^\mathsf{T})^\mathsf{T} : \boldsymbol{u} \in \mathcal{U}, \boldsymbol{v} \in \mathcal{V}\} \subseteq \mathbb{R}^{m+n} \tag{26}$$
is countably $(s+t)$-rectifiable.
iv) Every s-dimensional C^1-submanifold [36, Definition 5.3.1] of $\mathbb{R}^m$ is countably s-rectifiable. In particular, every s-dimensional affine subspace of $\mathbb{R}^m$ is countably s-rectifiable.
v) For $\mathcal{A}_i$ countably s_i-rectifiable and $s_i \le s$, $i \in \mathbb{N}$, the set
$$\mathcal{A} = \bigcup_{i \in \mathbb{N}} \mathcal{A}_i \tag{27}$$
is countably s-rectifiable.

Proof. See Appendix A. □

Countable unions of s-dimensional C^1-submanifolds of $\mathbb{R}^m$ are countably s-rectifiable by Properties iv) and v) of Lemma III.1. For countably $(\mathscr{H}^s, s)$-rectifiable sets we even get an equivalence result, namely:

Theorem III.1. [40, Theorem 3.2.29] A set $\mathcal{U} \subseteq \mathbb{R}^m$ is countably $(\mathscr{H}^s, s)$-rectifiable if and only if $\mathscr{H}^s$-a.a. of $\mathcal{U}$ is contained in the countable union of s-dimensional C^1-submanifolds of $\mathbb{R}^m$.

We now show that the upper modified Minkowski dimension of a countably s-rectifiable set is upper-bounded by s. This will allow us to conclude that, for a random vector x admitting a countably s-rectifiable support set, $K(\mathsf{x}) \le s$. The formal statement is as follows.

Lemma III.2. If $\mathcal{U} \subseteq \mathbb{R}^m$ is countably s-rectifiable, then $\overline{\dim}_{\mathrm{MB}}(\mathcal{U}) \le s$.

Proof. Suppose that $\mathcal{U} \subseteq \mathbb{R}^m$ is countably s-rectifiable. Then, Definition III.1 implies that there exist nonempty compact sets $\mathcal{A}_i \subseteq \mathbb{R}^s$ and Lipschitz mappings $\varphi_i \colon \mathcal{A}_i \to \mathbb{R}^m$, with $i \in \mathbb{N}$, such that

$$\mathcal{U} = \bigcup_{i\in\mathbb{N}} \varphi_i(\mathcal{A}_i). \tag{28}$$

Thus,

$$\begin{aligned}\overline{\dim}_{\mathrm{MB}}(\mathcal{U}) &= \sup_{i\in\mathbb{N}} \overline{\dim}_{\mathrm{MB}}(\varphi_i(\mathcal{A}_i)) &(29)\\ &\le \sup_{i\in\mathbb{N}} \overline{\dim}_{\mathrm{MB}}(\mathcal{A}_i) &(30)\\ &\le \sup_{i\in\mathbb{N}} \overline{\dim}_{\mathrm{MB}}(\mathbb{R}^s) &(31)\\ &= s, &(32)\end{aligned}$$

where the individual steps follow from properties of Lemma H.15, namely, (29) is by Property vi), (30) by Property v), (31) by Property iii), and (32) by Property ii). □

We next investigate the effect of locally Lipschitz mappings on (countably) s-rectifiable and countably $(\mathscr{H}^s, s)$-rectifiable sets.

Lemma III.3. Let $\mathcal{U} \subseteq \mathbb{R}^m$, and consider a locally Lipschitz $f \colon \mathbb{R}^m \to \mathbb{R}^n$. If $\mathcal{U}$ is

i) s-rectifiable, then $f(\mathcal{U})$ is s-rectifiable,
ii) countably s-rectifiable, then $f(\mathcal{U})$ is countably s-rectifiable,
iii) countably $(\mathscr{H}^s, s)$-rectifiable and Borel, then $f(\mathcal{U}) = \mathcal{A} \cup \mathcal{B}$, where $\mathcal{A}$ is a countably $(\mathscr{H}^s, s)$-rectifiable Borel set and $\mathscr{H}^s(\mathcal{B}) = 0$.

Proof. See Appendix B. □

A slightly weaker version of this statement, valid for Lipschitz mappings, was derived previously in [23, Lemma 4].

We are now ready to define rectifiable random vectors.

Definition III.2. (Rectifiable random vectors) [23, Definition 11] A random vector $\mathsf{x} \in \mathbb{R}^m$ is s-rectifiable if there exists a countably $(\mathscr{H}^s, s)$-rectifiable set $\mathcal{U} \subseteq \mathbb{R}^m$ such that $\mu_{\mathsf{x}} \ll \mathscr{H}^s|_{\mathcal{U}}$. The corresponding value s is the rectifiability parameter.

It turns out that an s-rectifiable random vector x always admits a countably s-rectifiable support set and, therefore, has description complexity $K(\mathsf{x}) \le s$ by Lemma III.2. The formal statement is as follows.

Lemma III.4. Every s-rectifiable random vector $\mathsf{x} \in \mathbb{R}^m$ admits a countably s-rectifiable support set. In particular, every s-rectifiable random vector x has description complexity $K(\mathsf{x}) \le s$.

Proof. Suppose that x is s-rectifiable. Then, there exists a countably $(\mathscr{H}^s, s)$-rectifiable set $\mathcal{U} \subseteq \mathbb{R}^m$ such that $\mu_{\mathsf{x}} \ll \mathscr{H}^s|_{\mathcal{U}}$. As $\mathcal{U}$ is countably $(\mathscr{H}^s, s)$-rectifiable, by Definition III.1 there exists a countably s-rectifiable set $\mathcal{V} \subseteq \mathbb{R}^m$ such that $\mathscr{H}^s(\mathcal{U} \setminus \mathcal{V}) = 0$. Set $\mathcal{W} = (\mathcal{U} \setminus \mathcal{V}) \cup \mathcal{V}$ and note that $\mathcal{U} \subseteq \mathcal{W}$ implies $\mathscr{H}^s|_{\mathcal{U}} \ll \mathscr{H}^s|_{\mathcal{W}}$ by monotonicity of $\mathscr{H}^s$. Moreover, from the definition of $\mathcal{W}$, the countable additivity of $\mathscr{H}^s$, and $\mathscr{H}^s(\mathcal{U} \setminus \mathcal{V}) = 0$, it follows that $\mathscr{H}^s|_{\mathcal{W}} = \mathscr{H}^s|_{\mathcal{V}}$. Thus, $\mathscr{H}^s|_{\mathcal{U}} \ll \mathscr{H}^s|_{\mathcal{V}}$, and $\mu_{\mathsf{x}} \ll \mathscr{H}^s|_{\mathcal{U}}$ implies $\mu_{\mathsf{x}} \ll \mathscr{H}^s|_{\mathcal{V}}$. Therefore, as $\mathscr{H}^s|_{\mathcal{V}}(\mathbb{R}^m \setminus \mathcal{V}) = 0$, we can conclude that $\mathrm{P}[\mathsf{x} \in \mathbb{R}^m \setminus \mathcal{V}] = 0$, which is equivalent to $\mathrm{P}[\mathsf{x} \in \mathcal{V}] = 1$, that is, the countably s-rectifiable set $\mathcal{V}$ is a support set of x. The proof is now concluded by noting that $K(\mathsf{x}) \le \overline{\dim}_{\mathrm{MB}}(\mathcal{V}) \le s$, where the latter inequality is by Lemma III.2. □

In light of Theorem III.1, an s-rectifiable random vector $\mathsf{x} \in \mathbb{R}^m$ is supported on a countable union of s-dimensional C^1-submanifolds of $\mathbb{R}^m$. Absolute continuity of μ_{x} with respect to the s-dimensional Hausdorff measure is a regularity condition guaranteeing that x cannot have positive probability measure on sets of Hausdorff dimension $t < s$ (see Property i) of Lemma H.3). This regularity condition together with the sharp transition behavior of Hausdorff measures (see Figure 3 in Appendix H) implies uniqueness of the rectifiability parameter. The corresponding formal statement is as follows.

Lemma III.5. If $\mathsf{x} \in \mathbb{R}^m$ is both s and t-rectifiable, then $s = t$.

Proof. See Appendix C. □

We are now ready to particularize our achievability result to s-rectifiable random vectors.

Corollary III.1. For s-rectifiable $\mathsf{x} \in \mathbb{R}^m$, $n > s$ is sufficient for the existence of a Borel measurable mapping $g \colon \mathbb{R}^{n\times m} \times \mathbb{R}^n \to \mathbb{R}^m$ satisfying

$$\mathrm{P}[g(\boldsymbol{A}, \boldsymbol{A}\mathsf{x}) \neq \mathsf{x}] = 0 \quad \text{for } \lambda^{n\times m}\text{-a.a. } \boldsymbol{A} \in \mathbb{R}^{n\times m}. \tag{33}$$

Proof. Follows from the general achievability result Theorem II.1 and Lemma III.4. □

Again, Corollary III.1 is nonasymptotic (i.e., it applies for finite m) and guarantees zero error probability. Next, we present three examples of s-rectifiable random vectors aimed at illustrating the relationship between the rectifiability parameter and the stochastic sparsity level defined in (11). Specifically, in the first example, the random vector's rectifiability parameter will be shown to agree with its stochastic sparsity level. The second example constructs an $(r + t - 1)$-rectifiable random vector of stochastic sparsity level $S(\mathsf{x}) = rt$ for general r, t. In this case, the stochastic sparsity level rt can be much larger than the rectifiability parameter $r + t - 1$, and hence Corollary III.1 implies that this random vector can be recovered with zero error probability from a number of linear measurements that is much smaller than its stochastic sparsity level. The third example constructs a random vector that is uniformly distributed on a manifold, namely, the unit circle. In this case, the random vector's rectifiability parameter equals the dimension of the manifold, whereas its stochastic sparsity level equals the dimension of the ambient space, i.e., the random vector is not sparse at all.

Example III.1. Suppose that $\mathbf{x} = (\boldsymbol{e}_{\mathsf{k}_1} \dots \boldsymbol{e}_{\mathsf{k}_s})\mathbf{z} \in \mathbb{R}^m$, where $\mathbf{z} \in \mathbb{R}^s$ with $\mu_{\mathbf{z}} \ll \lambda^s$ and $\mathbf{k} = (\mathsf{k}_1 \dots \mathsf{k}_s)^\mathsf{T} \in \{1, \dots, m\}^s$ satisfies $\mathsf{k}_1 < \dots < \mathsf{k}_s$. We first show that $S(\mathbf{x}) = s$. To this end, let

$$\mathcal{U} = \{\boldsymbol{x} \in \mathbb{R}^m : \|\boldsymbol{x}\|_0 \leq s\}. \tag{34}$$

Since $\mathrm{P}[\mathbf{x} \in \mathcal{U}] = 1$ by construction, it follows that $S(\mathbf{x}) \leq s$. To establish that $S(\mathbf{x}) \geq s$, and hence $S(\mathbf{x}) = s$, towards a contradiction, assume that there exists a linear subspace $\mathcal{T} \subseteq \mathbb{R}^m$ of dimension $d < s$ such that $\mathrm{P}[\mathbf{x} \in \mathcal{T}] > 0$. Since

$$\begin{aligned} 0 &< \mu_{\mathbf{x}}(\mathcal{T}) && (35)\\ &= \mathrm{P}[(\boldsymbol{e}_{\mathsf{k}_1} \dots \boldsymbol{e}_{\mathsf{k}_s})\mathbf{z} \in \mathcal{T}] && (36)\\ &\leq \sum_{1 \leq i_1 < \dots < i_s \leq m} \mathrm{P}[(\boldsymbol{e}_{i_1} \dots \boldsymbol{e}_{i_s})\mathbf{z} \in \mathcal{T}], && (37) \end{aligned}$$

there must exist a set of indices $\{i_1, \dots, i_s\} \subseteq \{1, \dots, m\}$ with $i_1 < \dots < i_s$ such that

$$\mathrm{P}[\boldsymbol{E}\mathbf{z} \in \mathcal{T}] > 0, \tag{38}$$

where $\boldsymbol{E} = (\boldsymbol{e}_{i_1} \dots \boldsymbol{e}_{i_s})$. Next, consider the linear subspace

$$\widetilde{\mathcal{T}} = \{\boldsymbol{z} \in \mathbb{R}^s : \boldsymbol{E}\boldsymbol{z} \in \mathcal{T}\} \subseteq \mathbb{R}^s, \tag{39}$$

which, by (38), satisfies $\mathrm{P}[\mathbf{z} \in \widetilde{\mathcal{T}}] > 0$. As $d < s$ by assumption and $\dim(\widetilde{\mathcal{T}}) \leq \dim(\mathcal{T}) = d$, there must exist a nonzero vector $\boldsymbol{b}_0 \in \mathbb{R}^s$ such that $\boldsymbol{b}_0^\mathsf{T}\boldsymbol{z} = 0$ for all $\boldsymbol{z} \in \widetilde{\mathcal{T}}$. It follows that $\mathrm{P}[\boldsymbol{b}_0^\mathsf{T}\mathbf{z} = 0] \geq \mathrm{P}[\mathbf{z} \in \widetilde{\mathcal{T}}] > 0$, which stands in contradiction to $\mu_{\mathbf{z}} \ll \lambda^s$ because $\lambda^s(\{\boldsymbol{z} : \boldsymbol{b}_0^\mathsf{T}\boldsymbol{z} = 0\}) = 0$. Thus, we indeed have $S(\mathbf{x}) = s$. It follows from Properties iv) and v) in Lemma III.1 that $\mathcal{U}$ in (34) is countably s-rectifiable and, therefore, also countably $(\mathscr{H}^s, s)$-rectifiable. We will see in Example IV.3 that $\mathbf{x}$ is in fact s-analytic, which, by Property ii) of Lemma IV.3, implies $\mu_{\mathbf{x}} \ll \mathscr{H}^s$. Finally, by $\mathrm{P}[\mathbf{x} \in \mathcal{U}] = 1$, we get $\mu_{\mathbf{x}} \ll \mathscr{H}^s|_{\mathcal{U}}$, which, thanks to the countable $(\mathscr{H}^s, s)$-rectifiability of $\mathcal{U}$ establishes the s-rectifiability of $\mathbf{x}$.

Example III.2. Let $\mathbf{x} = \mathbf{a} \otimes \mathbf{b}$ with $\mathbf{a} \in \mathbb{R}^k$ and $\mathbf{b} \in \mathbb{R}^l$. Suppose that $\mathbf{a} = (\boldsymbol{e}_{\mathsf{p}_1} \dots \boldsymbol{e}_{\mathsf{p}_r})\mathbf{u}$ and $\mathbf{b} = (\boldsymbol{e}_{\mathsf{q}_1} \dots \boldsymbol{e}_{\mathsf{q}_t})\mathbf{v}$, where $\mathbf{u} \in \mathbb{R}^r$ and $\mathbf{v} \in \mathbb{R}^t$ with $\mu_{\mathbf{u}} \times \mu_{\mathbf{v}} \ll \lambda^{r+t}$, and $\mathbf{p} = (\mathsf{p}_1 \dots \mathsf{p}_r)^\mathsf{T} \in \{1, \dots, k\}^r$ and $\mathbf{q} = (\mathsf{q}_1 \dots \mathsf{q}_t)^\mathsf{T} \in \{1, \dots, l\}^t$ satisfy $\mathsf{p}_1 < \dots < \mathsf{p}_r$ and $\mathsf{q}_1 < \dots < \mathsf{q}_t$, respectively. We first show that $S(\mathbf{x}) = rt$. Since $\mathrm{P}[\|\mathbf{x}\|_0 \leq rt] = 1$ by construction, it follows that $S(\mathbf{x}) \leq rt$. To establish that $S(\mathbf{x}) \geq rt$, and hence $S(\mathbf{x}) = rt$, towards a contradiction, assume that there exists a linear subspace $\mathcal{T} \subseteq \mathbb{R}^m$ of dimension $d < rt$ such that $\mathrm{P}[\mathbf{x} \in \mathcal{T}] > 0$. Since

$$\begin{aligned} &0 < \mu_{\mathbf{x}}(\mathcal{T}) && (40)\\ &= \mathrm{P}[((\boldsymbol{e}_{\mathsf{p}_1} \dots \boldsymbol{e}_{\mathsf{p}_r})\mathbf{u}) \otimes ((\boldsymbol{e}_{\mathsf{q}_1} \dots \boldsymbol{e}_{\mathsf{q}_t})\mathbf{v}) \in \mathcal{T}] && (41)\\ &= \mathrm{P}[((\boldsymbol{e}_{\mathsf{p}_1} \dots \boldsymbol{e}_{\mathsf{p}_r}) \otimes (\boldsymbol{e}_{\mathsf{q}_1} \dots \boldsymbol{e}_{\mathsf{q}_t}))(\mathbf{u} \otimes \mathbf{v}) \in \mathcal{T}] && (42)\\ &\leq \sum_{\substack{1 \leq i_1 < \dots < i_r \leq k\\ 1 \leq j_1 < \dots < j_t \leq l}} \mathrm{P}[((\boldsymbol{e}_{i_1} \dots \boldsymbol{e}_{i_r}) \otimes (\boldsymbol{e}_{j_1} \dots \boldsymbol{e}_{j_t}))(\mathbf{u} \otimes \mathbf{v}) \in \mathcal{T}], && (43) \end{aligned}$$

where (42) relies on [35, Lemma 4.2.10], there must exist a set of indices $\{i_1, \dots, i_r\} \subseteq \{1, \dots, k\}$ with $i_1 < \dots < i_r$ and a set of indices $\{j_1, \dots, j_t\} \subseteq \{1, \dots, l\}$ with $j_1 < \dots < j_t$ such that

$$\mathrm{P}[(\boldsymbol{E}_1 \otimes \boldsymbol{E}_2)(\mathbf{u} \otimes \mathbf{v}) \in \mathcal{T}] > 0, \tag{44}$$

where $\boldsymbol{E}_1 = (\boldsymbol{e}_{i_1} \dots \boldsymbol{e}_{i_r})$ and $\boldsymbol{E}_2 = (\boldsymbol{e}_{j_1} \dots \boldsymbol{e}_{j_t})$. Next, consider the linear subspace

$$\widetilde{\mathcal{T}} = \{\boldsymbol{z} \in \mathbb{R}^{rt} : (\boldsymbol{E}_1 \otimes \boldsymbol{E}_2)\boldsymbol{z} \in \mathcal{T}\} \subseteq \mathbb{R}^{rt}, \tag{45}$$

which, by (44), satisfies $\mathrm{P}[\mathbf{u} \otimes \mathbf{v} \in \widetilde{\mathcal{T}}] > 0$. As $d < rt$ by assumption and $\dim(\widetilde{\mathcal{T}}) \leq \dim(\mathcal{T}) = d$, there must exist a nonzero vector $\boldsymbol{b}_0 \in \mathbb{R}^{rt}$ such that $\boldsymbol{b}_0^\mathsf{T}\boldsymbol{z} = 0$ for all $\boldsymbol{z} \in \widetilde{\mathcal{T}}$. It follows that $\mathrm{P}[\boldsymbol{b}_0^\mathsf{T}(\mathbf{u} \otimes \mathbf{v}) = 0] \geq \mathrm{P}[\mathbf{u} \otimes \mathbf{v} \in \widetilde{\mathcal{T}}] > 0$. As $\mu_{\mathbf{u}} \times \mu_{\mathbf{v}} \ll \lambda^{r+t}$ by assumption, we also have

$$\lambda^{r+t}\big(\{(\boldsymbol{u}^\mathsf{T}\ \boldsymbol{v}^\mathsf{T})^\mathsf{T} : \boldsymbol{u} \in \mathbb{R}^r, \boldsymbol{v} \in \mathbb{R}^t, \boldsymbol{b}_0^\mathsf{T}(\boldsymbol{u} \otimes \boldsymbol{v}) = 0\}\big) > 0. \tag{46}$$

We now view $\boldsymbol{b}_0^\mathsf{T}(\boldsymbol{u} \otimes \boldsymbol{v})$ as a polynomial in the entries of $\boldsymbol{u}$ and $\boldsymbol{v}$. Since a polynomial vanishes either on a set of Lebesgue measure zero or is identically zero (see Corollary K.1 and Lemma K.5), it follows that

$$\boldsymbol{b}_0^\mathsf{T}(\boldsymbol{u} \otimes \boldsymbol{v}) = 0 \quad \text{for all } \boldsymbol{u} \in \mathbb{R}^r \text{ and } \boldsymbol{v} \in \mathbb{R}^t, \tag{47}$$

which stands in contradiction to $\boldsymbol{b}_0 \neq \mathbf{0}$. Thus, we indeed have $S(\mathbf{x}) = rt$. We next construct a countably $(r+t-1)$-rectifiable support set $\mathcal{U}$ of $\mathbf{x}$. To this end, we let

$$\begin{aligned} \mathcal{A} &= \{\boldsymbol{a} \in \mathbb{R}^k : \|\boldsymbol{a}\|_0 \leq r\}, && (48)\\ \mathcal{B} &= \{\boldsymbol{b} \in \mathbb{R}^l : \|\boldsymbol{b}\|_0 \leq t\}, && (49) \end{aligned}$$

and set $\mathcal{U} = \mathcal{A} \otimes \mathcal{B}$. Since $\mathcal{A}$ is a support set of $\mathbf{a}$ and $\mathcal{B}$ is a support set of $\mathbf{b}$, $\mathcal{U}$ is a support set of $\mathbf{x} = \mathbf{a} \otimes \mathbf{b}$. Note that $\mathcal{U} = (\mathcal{A} \backslash \{\mathbf{0}\}) \otimes \mathcal{B}$. For $\boldsymbol{a} \in \mathcal{A} \backslash \{\mathbf{0}\}$, let $\bar{a}$ denote the first nonzero entry of $\boldsymbol{a}$. We can now write

$$\boldsymbol{a} \otimes \boldsymbol{b} = \left(\frac{\boldsymbol{a}}{\bar{a}}\right) \otimes (\bar{a}\,\boldsymbol{b}) \quad \text{for all } \boldsymbol{a} \in \mathcal{A} \backslash \{\mathbf{0}\}, \boldsymbol{b} \in \mathcal{B}. \tag{50}$$

This allows us to decompose $\mathcal{U}$ according to $\mathcal{U} = \tilde{\mathcal{A}} \otimes \mathcal{B}$, where

$$\tilde{\mathcal{A}} = \{\boldsymbol{a} \in \mathcal{A} \backslash \{\mathbf{0}\} : \|\boldsymbol{a}\|_0 \leq r, \bar{a} = 1\}. \tag{51}$$

Now, since $\tilde{\mathcal{A}}$ is a finite union of affine subspaces all of dimension $r-1$, it is countably $(r-1)$-rectifiable by Properties iv) and v) in Lemma III.1. By the same token, $\mathcal{B}$ as a finite union of linear subspaces all of dimension t is countably t-rectifiable. Therefore, the set

$$\mathcal{C} = \left\{\left(\boldsymbol{a}^\mathsf{T}\ \boldsymbol{b}^\mathsf{T}\right)^\mathsf{T} : \boldsymbol{a} \in \tilde{\mathcal{A}}, \boldsymbol{b} \in \mathcal{B}\right\} \subseteq \mathbb{R}^{k+l} \tag{52}$$

is countably $(r+t-1)$-rectifiable thanks to Property iii) of Lemma III.1. Now, the multivariate mapping $\sigma\colon \mathbb{R}^{k+l} \to \mathbb{R}^{kl}$, $(\boldsymbol{a}^\mathsf{T}\ \boldsymbol{b}^\mathsf{T})^\mathsf{T} \mapsto \boldsymbol{a} \otimes \boldsymbol{b}$ is bilinear and as such locally Lipschitz. Moreover, since $\mathcal{U} = \sigma(\mathcal{C})$ with $\mathcal{C}$ countably $(r+t-1)$-rectifiable, it follows from Property ii) of Lemma III.3 that $\mathcal{U}$ is countably $(r+t-1)$-rectifiable and, therefore, also countably $(\mathscr{H}^{r+t-1}, r+t-1)$-rectifiable. We will see in Example IV.4 that $\mathbf{x}$ is in fact $(r+t-1)$-analytic, which, by Property ii) of Lemma IV.3, implies $\mu_{\mathbf{x}} \ll \mathscr{H}^{r+t-1}$. With $\mathrm{P}[\mathbf{x} \in \mathcal{U}] = 1$ this yields $\mu_{\mathbf{x}} \ll \mathscr{H}^{r+t-1}|_{\mathcal{U}}$ and, in turn, thanks to countable

$(\mathscr{H}^{r+t-1}, r+t-1)$-rectifiability of $\mathcal{U}$, establishes $(r+t-1)$-rectifiability of x.

Example III.3. Let $\mathcal{S}^1$ denote the unit circle in $\mathbb{R}^2$, $\mathsf{z} \in \mathbb{R}$ with $\mu_{\mathsf{z}} \ll \lambda^1$, and $g\colon \mathbb{R} \to \mathcal{S}^1$, $z \mapsto (\cos(z)\ \ \sin(z))^\mathsf{T}$. Set $\mathsf{x} = g(\mathsf{z})$, and note that this implies $\mathrm{P}[\mathsf{x} \in \mathcal{S}^1] = 1$. We first establish that $S(\mathsf{x}) = 2$. Since $\mathrm{P}[\mathsf{x} \in \mathbb{R}^2] = 1$, it follows that $S(\mathsf{x}) \leq 2$. To establish that $S(\mathsf{x}) \geq 2$, and hence $S(\mathsf{x}) = 2$, towards a contradiction, assume that there exists a linear subspace $\mathcal{T} \subseteq \mathbb{R}^2$ of dimension one such that $\mathrm{P}[\mathsf{x} \in \mathcal{T}] > 0$. Set $\mathcal{A} = \mathcal{T} \cap \mathcal{S}^1$, which consists of two antipodal points on $\mathcal{S}^1$ (see Figure 1). Now, $0 < \mathrm{P}[\mathsf{x} \in \mathcal{T}] = \mathrm{P}[\mathsf{x} \in \mathcal{A}] = \mathrm{P}[\mathsf{z} \in g^{-1}(\mathcal{A})]$, which constitutes a contradiction to $\mu_{\mathsf{z}} \ll \lambda^1$ because $g^{-1}(\mathcal{A})$—as a countable set—must have Lebesgue measure zero. Therefore, $S(\mathsf{x}) = 2$. Finally, x is 1-rectifiable by [23, Section III.D].

Since s-rectifiable random vectors cannot have positive probability measure on sets of Hausdorff dimension $t < s$, it is natural to ask whether taking $n \geq s$ linear measurements is necessary for zero error recovery of s-rectifiable random vectors. Surprisingly, it turns out that this is not the case. This will be demonstrated by first constructing a 2-rectifiable (and therefore also $(\mathscr{H}^2, 2)$-rectifiable) set $\mathcal{G} \subseteq \mathbb{R}^3$ of strictly positive 2-dimensional Hausdorff measure with the property that $\boldsymbol{e}_3^\mathsf{T}\colon \mathbb{R}^3 \to \mathbb{R}$, $(x_1\ x_2\ x_3)^\mathsf{T} \mapsto x_3$ is one-to-one on $\mathcal{G}$. Then, we show that every 2-rectifiable random vector x satisfying $\mu_{\mathsf{x}} \ll \mathscr{H}^2|_{\mathcal{G}}$ can be recovered with zero error probability from one linear measurement, specifically from $\mathsf{y} = \boldsymbol{e}_3^\mathsf{T}\mathsf{x}$. Moreover, all this is possible even though $\mathcal{G}$ contains the image of a Borel set in $\mathbb{R}^2$ of positive Lebesgue measure under a C^∞-embedding.

The construction of our example is based on the following result.

Theorem III.2. There exist a compact set $\mathcal{A} \in \mathscr{B}(\mathbb{R}^2)$ with $\lambda^2(\mathcal{A}) = 1/4$, and a C^∞-function $\kappa\colon \mathbb{R}^2 \to \mathbb{R}$ such that κ is one-to-one on $\mathcal{A}$.

Proof. See Section VI for an explicit construction of κ and $\mathcal{A}$. □

We now proceed to the construction of our example demonstrating that $n \geq s$ is in general not a necessary condition for zero error recovery of s-rectifiable random vectors.

Example III.4. Let κ and $\mathcal{A}$ be as constructed in the proof of Theorem III.2 and consider the mapping

$$h\colon \mathbb{R}^2 \to \mathbb{R}^3 \tag{53}$$
$$\boldsymbol{z} \mapsto (\boldsymbol{z}^\mathsf{T}\ \kappa(\boldsymbol{z}))^\mathsf{T}. \tag{54}$$

We set $\mathcal{G} = h(\mathcal{A})$ and show the following:

i) h is a C^∞-embedding;
ii) $\mathcal{G}$ is 2-rectifiable;
iii) $0 < \mathscr{H}^2(\mathcal{G}) < \infty$;
iv) $\boldsymbol{e}_3^\mathsf{T}\colon \mathbb{R}^3 \to \mathbb{R}$, $(x_1\ x_2\ x_3)^\mathsf{T} \mapsto x_3$ is one-to-one on $\mathcal{G}$.
v) For every 2-rectifiable random vector $\mathsf{x} \in \mathbb{R}^3$ with $\mu_{\mathsf{x}} \ll \mathscr{H}^2|_{\mathcal{G}}$, there exists a Borel measurable mapping $g\colon \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}^3$ satisfying $\mathrm{P}\big[g\big(\boldsymbol{e}_3, \boldsymbol{e}_3^\mathsf{T}\mathsf{x}\big) \neq \mathsf{x}\big] = 0$.

It follows immediately that h is one-to-one. Thus, to establish Property i), it suffices to prove that h is a C^∞-immersion. Since κ is C^∞, so is h. Furthermore,

$$Jh(\boldsymbol{z}) = \sqrt{\det((Dh(\boldsymbol{z}))^\mathsf{T} Dh(\boldsymbol{z}))} \tag{55}$$
$$= \sqrt{\det(\boldsymbol{I}_2 + \boldsymbol{a}(\boldsymbol{z})\boldsymbol{a}(\boldsymbol{z})^\mathsf{T})} \quad \text{for all } \boldsymbol{z} \in \mathbb{R}^2, \tag{56}$$

where

$$\boldsymbol{a}(z) = \begin{pmatrix} \frac{\partial\kappa(\boldsymbol{z})}{\partial z_1} \\ \frac{\partial\kappa(\boldsymbol{z})}{\partial z_2} \end{pmatrix}. \tag{57}$$

Since $\boldsymbol{a}(\boldsymbol{z})\boldsymbol{a}(\boldsymbol{z})^\mathsf{T}$ is positive semidefinite, $Jh(z) \geq \sqrt{\det(\boldsymbol{I}_2)} = 1$ for all $\boldsymbol{z} \in \mathbb{R}^2$, which establishes that h is an immersion and completes the proof of i).

To prove ii), note that h is C^∞ and as such locally Lipschitz. As $\mathcal{A}$ is compact, Lemma H.12 implies that $h|_{\mathcal{A}}$ is Lipschitz. The set $\mathcal{G} = h(\mathcal{A})$ is hence the Lipschitz image of a compact set in $\mathbb{R}^2$ and as such 2-rectifiable.

To establish iii), we first note that

$$\mathscr{H}^2(\mathcal{G}) = \mathscr{H}^2(h(\mathcal{A})) \tag{58}$$
$$\leq L^2 \mathscr{H}^2(\mathcal{A}) \tag{59}$$
$$= L^2 \lambda^2(\mathcal{A}) \tag{60}$$
$$< \infty, \tag{61}$$

where the individual steps follow from Properties of Lemma H.3, namely, (59) from Property ii) with L denoting the Lipschitz constant of $h|_{\mathcal{A}}$, and (60) from Property iii).

To establish $\mathscr{H}^2(\mathcal{G}) > 0$, consider the linear mapping

$$\pi\colon \mathbb{R}^3 \to \mathbb{R}^2 \tag{62}$$
$$(x_1\ x_2\ x_3)^\mathsf{T} \mapsto (x_1\ x_2)^\mathsf{T}. \tag{63}$$

Clearly, π is Lipschitz with Lipschitz constant equal to 1. Therefore,

$$\mathscr{H}^2(\mathcal{G}) \geq \mathscr{H}^2(\pi(\mathcal{G})) \tag{64}$$
$$= \mathscr{H}^2(\mathcal{A}) \tag{65}$$
$$= \lambda^2(\mathcal{A}) \tag{66}$$
$$= \tfrac{1}{4}, \tag{67}$$

where (64) follows from Property ii) of Lemma H.3, (65) from $\pi(\mathcal{G}) = \mathcal{A}$, and (66) is by Property iii) of Lemma H.3.

To show iv), let $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{G}$ with $\boldsymbol{x}_1 \neq \boldsymbol{x}_2$. Thus, $\boldsymbol{x}_1 = (\boldsymbol{z}_1^\mathsf{T}\ \kappa(\boldsymbol{z}_1))^\mathsf{T}$ and $\boldsymbol{x}_2 = (\boldsymbol{z}_2^\mathsf{T}\ \kappa(\boldsymbol{z}_2))^\mathsf{T}$ with $\boldsymbol{z}_1, \boldsymbol{z}_2 \in \mathcal{A}$ and $\boldsymbol{z}_1 \neq \boldsymbol{z}_2$. As κ is one-to-one on $\mathcal{A}$, we conclude that $\boldsymbol{e}_3^\mathsf{T}$ is one-to-one on $\mathcal{G}$.

It remains to establish v). Since $\mu_{\mathsf{x}} \ll \mathscr{H}^2|_{\mathcal{G}}$ by assumption, it follows that $\mathrm{P}[\mathsf{x} \in \mathcal{G}] = 1$. We show that there exists a Borel measurable mapping $g\colon \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}^3$ such that

$$g(\boldsymbol{a}, y) \begin{cases} \in \big\{\boldsymbol{v} \in \mathcal{G} : \boldsymbol{a}^\mathsf{T}\boldsymbol{v} = y\big\} & \text{if } \exists \boldsymbol{v} \in \mathcal{G} : \boldsymbol{a}^\mathsf{T}\boldsymbol{v} = y \\ = \boldsymbol{e} & \text{else,} \end{cases} \tag{68}$$

where $\boldsymbol{e}$ is an arbitrary vector not in $\mathcal{G}$, used to declare a decoding error. Since $\mathrm{P}[\mathsf{x} \in \mathcal{G}] = 1$ and $\boldsymbol{e}_3^\mathsf{T}$ is one-to-one on

$\mathcal{G}$, this then implies that $\mathrm{P}\big[g\big(\boldsymbol{e}_3, \boldsymbol{e}_3^\mathsf{T}\mathbf{x}\big) \neq \mathbf{x}\big] = 0$. To construct g in (68), consider first the mapping

$$f\colon \mathbb{R}^3 \times \mathbb{R} \times \mathbb{R}^3 \to \mathbb{R} \tag{69}$$

$$(\boldsymbol{a}, y, \boldsymbol{u}) \mapsto |y - \boldsymbol{a}^\mathsf{T}\boldsymbol{u}|. \tag{70}$$

Since f is continuous, Lemma I.3 implies that f is a normal integrand (see Definition I.4) with respect to $\mathscr{B}(\mathbb{R}^3 \times \mathbb{R})$. Let

$$\mathcal{T} = \{(\boldsymbol{a}, y) \in \mathbb{R}^3 \times \mathbb{R} : \exists \boldsymbol{u} \in \mathcal{G} \text{ with } f(\boldsymbol{a}, y, \boldsymbol{u}) \leq 0\} \tag{71}$$

$$= \{(\boldsymbol{a}, y) \in \mathbb{R}^3 \times \mathbb{R} : \exists \boldsymbol{u} \in \mathcal{G} \text{ with } \boldsymbol{a}^\mathsf{T}\boldsymbol{u} = y\}. \tag{72}$$

Note that $\mathcal{G}$ as the Lipschitz image of the compact set $\mathcal{A}$ is compact (see Lemma H.13). It now follows from Properties ii) and iii) of Lemma I.5 (with $\mathcal{T} = \mathbb{R}^3 \times \mathbb{R}$, $\alpha = 0$, $\mathcal{K} = \mathcal{G}$, and f as in (69)–(70), which is a normal integrand with respect to $\mathscr{B}(\mathbb{R}^3 \times \mathbb{R})$) that i) $\mathcal{T} \in \mathscr{B}(\mathbb{R}^3 \times \mathbb{R})$ and ii) there exists a Borel measurable mapping

$$p\colon \mathcal{T} \to \mathbb{R}^3 \tag{73}$$

$$\big(\boldsymbol{a}, y\big) \mapsto p\big(\boldsymbol{a}, y\big) \in \big\{\boldsymbol{u} \in \mathcal{G} : \boldsymbol{a}^\mathsf{T}\boldsymbol{u} = y\big\}. \tag{74}$$

This mapping can then be extended to a mapping $g\colon \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}^3$ by setting

$$g|_{\mathcal{T}} = p \tag{75}$$

$$g|_{(\mathbb{R}^3 \times \mathbb{R}) \setminus \mathcal{T}} = \boldsymbol{e}. \tag{76}$$

Finally, g is Borel measurable owing to Lemma H.8 as p is Borel measurable and $\mathcal{T} \in \mathscr{B}(\mathbb{R}^3 \times \mathbb{R})$.

IV. Strong Converse

Example III.4 in the previous section demonstrates that $n \geq s$ is not necessary for zero error recovery of s-rectifiable random vectors in general. In this section, we introduce the class of s-analytic random vectors $\mathbf{x}$, which will be shown to allow for a strong converse in the sense of $n \geq s$ being necessary for recovery of $\mathbf{x}$ with probability of error smaller than one. The adjective "strong" refers to the fact that $n < s$ linear measurements are necessarily insufficient even if we allow a recovery error probability that is arbitrarily close to one. We prove that an s-analytic random vector is s-rectifiable if and only if it admits a support set $\mathcal{U}$ that is "not too rich" (in terms of σ-finiteness of $\mathscr{H}^s|_{\mathcal{U}}$), show that the s-rectifiable random vectors considered in Examples III.1–III.3 are all s-analytic, and discuss examples of s-analytic random vectors that fail to be s-rectifiable. Random vectors that are both s-analytic and s-rectifiable can be recovered with zero error probability from $n > s$ linear measurements, and $n \geq s$ linear measurements are necessary for recovery with error probability smaller than one. The border case $n = s$ remains open.

We now make our way towards developing the strong converse and the formal definition of s-analyticity. The following auxiliary result will turn out to be useful.

Lemma IV.1. For $\mathbf{x} \in \mathbb{R}^m$ and $\boldsymbol{A} \in \mathbb{R}^{n \times m}$, consider the following statements:

i) There exists a Borel measurable mapping $g\colon \mathbb{R}^{n \times m} \times \mathbb{R}^n \to \mathbb{R}^m$ satisfying $\mathrm{P}[g(\boldsymbol{A}, \boldsymbol{A}\mathbf{x}) \neq \mathbf{x}] < 1$.
ii) There exists a set $\mathcal{U} \in \mathscr{B}(\mathbb{R}^m)$ with $\mathrm{P}[\mathbf{x} \in \mathcal{U}] > 0$ such that $\boldsymbol{A}$ is one-to-one on $\mathcal{U}$.

Then, i) implies ii).

Proof. See Appendix D. □

We first establish a strong converse for the class of random vectors considered in Example III.1. This will guide us to the crucial defining property of s-analytic random vectors.

Lemma IV.2. Let $\mathbf{x} = (\boldsymbol{e}_{\mathsf{k}_1} \dots \boldsymbol{e}_{\mathsf{k}_s})\mathbf{z} \in \mathbb{R}^m$, where $\mathbf{z} \in \mathbb{R}^s$ with $\mu_{\mathbf{z}} \ll \lambda^s$ and $\mathbf{k} = (\mathsf{k}_1 \dots \mathsf{k}_s)^\mathsf{T} \in \{1, \dots, m\}^s$ satisfies $\mathsf{k}_1 < \dots < \mathsf{k}_s$. If there exist a measurement matrix $\boldsymbol{A} \in \mathbb{R}^{n \times m}$ and a Borel measurable mapping $g\colon \mathbb{R}^{n \times m} \times \mathbb{R}^n \to \mathbb{R}^m$ such that $\mathrm{P}[g(\boldsymbol{A}, \boldsymbol{A}\mathbf{x}) \neq \mathbf{x}] < 1$, then $n \geq s$.

Proof. Towards a contradiction, suppose that there exist a measurement matrix $\boldsymbol{A} \in \mathbb{R}^{n \times m}$ and a Borel measurable mapping $g\colon \mathbb{R}^{n \times m} \times \mathbb{R}^n \to \mathbb{R}^m$ so that $\mathrm{P}[g(\boldsymbol{A}, \boldsymbol{A}\mathbf{x}) \neq \mathbf{x}] < 1$ for $n < s$. By Lemma IV.1, there must exist a $\mathcal{U} \in \mathscr{B}(\mathbb{R}^m)$ with $\mathrm{P}[\mathbf{x} \in \mathcal{U}] > 0$ such that $\boldsymbol{A}$ is one-to-one on $\mathcal{U}$. Since

$$\mathrm{P}[\mathbf{x} \in \mathcal{U}] = \mathrm{P}[(\boldsymbol{e}_{\mathsf{k}_1} \dots \boldsymbol{e}_{\mathsf{k}_s})\mathbf{z} \in \mathcal{U}] \tag{77}$$

$$\leq \sum_{1 \leq i_1 < \dots < i_s \leq m} \mathrm{P}[(\boldsymbol{e}_{i_1} \dots \boldsymbol{e}_{i_s})\mathbf{z} \in \mathcal{U}], \tag{78}$$

there must exist a set of indices $\{i_1, \dots, i_s\} \subseteq \{1, \dots, m\}$ with $i_1 < \dots < i_s$ such that the rank-s matrix $\boldsymbol{H} := (\boldsymbol{e}_{i_1} \dots \boldsymbol{e}_{i_s})$ satisfies $\mathrm{P}[\boldsymbol{H}\mathbf{z} \in \mathcal{U}] > 0$. Setting $\mathcal{A} = \{\boldsymbol{z} \in \mathbb{R}^s : \boldsymbol{H}\boldsymbol{z} \in \mathcal{U}\}$ yields $\mu_{\mathbf{z}}(\mathcal{A}) = \mathrm{P}[\boldsymbol{H}\mathbf{z} \in \mathcal{U}] > 0$. Furthermore, $\mathcal{A}$ as the inverse image of the Borel set $\mathcal{U}$ under a linear mapping is Borel. Finally, since $\mu_{\mathbf{z}} \ll \lambda^s$ by assumption, we conclude that $\lambda^s(\mathcal{A}) > 0$. Summarizing, there exist a set $\mathcal{A} \in \mathscr{B}(\mathbb{R}^s)$ with $\lambda^s(\mathcal{A}) > 0$ and a matrix $\boldsymbol{H} \in \mathbb{R}^{m \times s}$ such that $\mathcal{U}$ contains the one-to-one image of $\mathcal{A}$ under $\boldsymbol{H}$. We now follow the line of argumentation used in the proof of the converse part of [17, Theorem 6]. Specifically, as $\boldsymbol{A}$ is one-to-one on $\mathcal{U}$ and $\boldsymbol{H}$ is one-to-one on $\mathbb{R}^s$, it follows that $\boldsymbol{A}\boldsymbol{H}$ is one-to-one on $\mathcal{A}$, i.e.,

$$\ker(\boldsymbol{A}\boldsymbol{H}) \cap (\mathcal{A} - \mathcal{A}) = \{\mathbf{0}\}. \tag{79}$$

As $\lambda^s(\mathcal{A}) > 0$, the Steinhaus Theorem [41] implies the existence of an $r > 0$ such that $\mathcal{B}_s(\mathbf{0}, r) \subseteq \mathcal{A} - \mathcal{A} \subseteq \mathbb{R}^s$. Since $\dim \ker(\boldsymbol{A}\boldsymbol{H}) \geq s - n > 0$, we conclude that the linear subspace $\ker(\boldsymbol{A}\boldsymbol{H})$ must have a nontrivial intersection with $\mathcal{A} - \mathcal{A}$, which stands in contradiction to (79). □

The strong converse just derived hinges critically on the specific structure of the s-rectifiable random vector $\mathbf{x}$ considered. Concretely, we used the fact that, for every $\mathcal{U} \in \mathscr{B}(\mathbb{R}^m)$ with $\mathrm{P}[\mathbf{x} \in \mathcal{U}] > 0$, there exist a set $\mathcal{A} \in \mathscr{B}(\mathbb{R}^s)$ with $\lambda^s(\mathcal{A}) > 0$ and a matrix $\boldsymbol{H} \in \mathbb{R}^{m \times s}$ such that $\mathcal{U}$ contains the one-to-one image of $\mathcal{A}$ under $\boldsymbol{H}$. The following example demonstrates, however, that this property is too strong for our purposes as it fails to hold for random vectors on general manifolds like, e.g., the unit circle:

Example IV.1. Let $\mathcal{S}^1 \subseteq \mathbb{R}^2$ denote the unit circle and consider $\mathbf{x} \in \mathbb{R}^2$ supported on $\mathcal{S}^1$, i.e., $\mathrm{P}[\mathbf{x} \in \mathcal{S}^1] = 1$. Towards a contradiction, suppose that there exist a set $\mathcal{A} \in \mathscr{B}(\mathbb{R})$ with $\lambda^1(\mathcal{A}) > 0$ and a vector $\boldsymbol{h} \in \mathbb{R}^2$ such that $\{\boldsymbol{h}z : z \in \mathcal{A}\} \subseteq \mathcal{S}^1$ and $\boldsymbol{h}$ is one-to-one on $\mathcal{A}$. Since $\boldsymbol{h}$ is one-to-one on $\mathcal{A}$ and $\lambda^1(\mathcal{A}) > 0$, it follows that $\boldsymbol{h} \neq \mathbf{0}$.

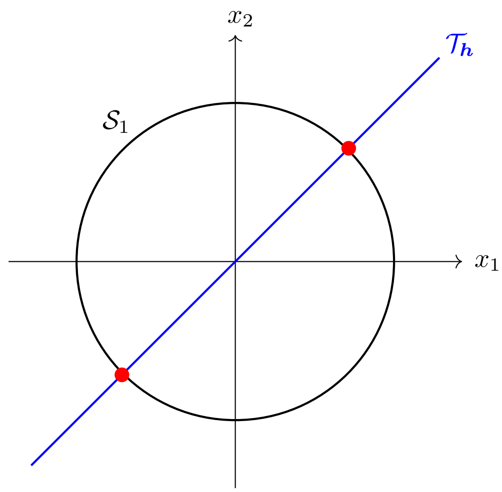

Figure 1. For every vector $\boldsymbol{h} \subseteq \mathbb{R}^2 \backslash \{\mathbf{0}\}$, the linear subspace $\mathcal{T}_{\boldsymbol{h}} := \{\boldsymbol{h}z : z \in \mathbb{R}\}$ intersects the unit circle $\mathcal{S}_1$ in the two antipodal points $\pm\boldsymbol{h}/\|\boldsymbol{h}\|_2$.

Noting that $\{\boldsymbol{h}z : z \in \mathcal{A}\} \subseteq \{\boldsymbol{h}/\|\boldsymbol{h}\|_2, -\boldsymbol{h}/\|\boldsymbol{h}\|_2\}$ (see Figure 1), $\mathcal{A}$ necessarily satisfies $\mathcal{A} \subseteq \{1/\|\boldsymbol{h}\|_2, -1/\|\boldsymbol{h}\|_2\}$. Thus, $\lambda^1(\mathcal{A}) = 0$, which is a contradiction to $\lambda^1(\mathcal{A}) > 0$.

The reason for this failure is that every $\boldsymbol{h} \in \mathbb{R}^2$ maps into a 1-dimensional linear subspace in $\mathbb{R}^2$, and 1-dimensional linear subspaces in $\mathbb{R}^2$ intersect the unit circle in two antipodal points only. To map a set $\mathcal{A} \in \mathscr{B}(\mathbb{R})$ to a set in $\mathbb{R}^2$ that is not restricted to be a subset of a 1-dimensional linear subspace, we have to employ a nonlinear mapping. But this puts us into the same dilemma as in Example III.4, where we demonstrated that even the requirement of every $\mathcal{U} \in \mathscr{B}(\mathbb{R}^m)$ with $\mathrm{P}[\mathbf{x} \in \mathcal{U}] > 0$ containing the embedded image—under a C^∞-mapping—of a set $\mathcal{A} \in \mathscr{B}(\mathbb{R}^s)$ of positive Lebesgue measure is not sufficient to obtain a strong converse for general $\mathbf{x}$. We therefore need to impose additional constraints on the mapping. It turns out that requiring real analyticity is enough. Examples of real analytic mappings include, e.g., multivariate polynomials, the exponential function, or trigonometric mappings. This finally leads us to the new concept of s-analytic measures and s-analytic random vectors.

Definition IV.1. (Analytic measures) A Borel measure μ on $\mathbb{R}^m$ is s-analytic, with $s \in \{1, \ldots, m\}$, if, for each $\mathcal{U} \in \mathscr{B}(\mathbb{R}^m)$ with $\mu(\mathcal{U}) > 0$, there exist a set $\mathcal{A} \in \mathscr{B}(\mathbb{R}^s)$ of positive Lebesgue measure and a real analytic mapping (see Definition K.1) $h \colon \mathbb{R}^s \to \mathbb{R}^m$ of s-dimensional Jacobian $Jh \not\equiv 0$ such that $h(\mathcal{A}) \subseteq \mathcal{U}$.

Note that the only requirement on the real analytic mappings in Definition IV.1 is that their s-dimensional Jacobians do not vanish identically. Since the s-dimensional Jacobian of a real analytic mapping is a real analytic function, it vanishes either identically or on a set of Lebesgue measure zero (see Lemma K.7). By Lemma K.8, this guarantees that, for an analytic measure μ, every $\mathcal{U} \in \mathscr{B}(\mathbb{R}^m)$ with $\mu(\mathcal{U}) > 0$ contains the real analytic *embedding* of a set $\mathcal{A} \in \mathscr{B}(\mathbb{R}^s)$ of positive Lebesgue measure.

We have the following properties of s-analytic measures.

Lemma IV.3. If μ is an s-analytic measure on $\mathbb{R}^m$, then the following holds:

i) μ is t-analytic for all $t \in \{1, \ldots, s-1\}$;
ii) $\mu \ll \mathscr{H}^s$;
iii) there exists a set $\mathcal{U} \subseteq \mathbb{R}^m$ such that $\mu = \mu|_{\mathcal{U}}$ and $\mathscr{H}^s|_{\mathcal{U}}$ is σ-finite if and only if there exists a countably $(\mathscr{H}^s, s)$-rectifiable set $\mathcal{W} \subseteq \mathbb{R}^m$ such that $\mu = \mu|_{\mathcal{W}}$.

Proof. See Appendix E. □

We are now ready to define s-analytic random vectors.

Definition IV.2. (Analytic random vectors) A random vector $\mathbf{x} \in \mathbb{R}^m$ is s-analytic if $\mu_{\mathbf{x}}$ is s-analytic. The corresponding value s is the analyticity parameter.

We have the following immediate consequence of Lemma IV.3.

Corollary IV.1. Let $\mathbf{x}$ be s-analytic. Then, $\mathbf{x}$ is s-rectifiable if and only if it admits a support set $\mathcal{U}$ such that $\mathscr{H}^s|_{\mathcal{U}}$ is σ-finite.

Proof. Follows from Properties ii) and iii) in Lemma IV.3 and Definition III.2. □

By Corollary IV.1, an s-analytic random vector is s-rectifiable if and only if it admits a support set $\mathcal{U}$ that is "not too rich" (in terms of σ-finiteness of $\mathscr{H}^s|_{\mathcal{U}}$). As an example of an s-analytic random vector that is not s-rectifiable, consider an $(s+1)$-analytic random vector $\mathbf{x}$ with $s > 0$. By Property i) in Lemma IV.3, this $\mathbf{x}$ is also s-analytic, but it cannot be s-rectifiable, as shown next. Towards a contradiction, suppose that $\mathbf{x}$ is s-rectifiable. Then, by Lemma III.4, $\mathbf{x}$ has a countably s-rectifiable support set $\mathcal{U}$, which by Property i) in Lemma III.1 is also countably $(s+1)$-rectifiable. As, by assumption, $\mathbf{x}$ is $(s+1)$-analytic, Property ii) in Lemma IV.3 implies $\mu_{\mathbf{x}} \ll \mathscr{H}^{s+1}$. Thus, $\mu_{\mathbf{x}} \ll \mathscr{H}^{s+1}|_{\mathcal{U}}$ with $\mathcal{U}$ countably $(s+1)$-rectifiable, and we conclude that $\mathbf{x}$ would also be $(s+1)$-rectifiable, which contradicts uniqueness of the rectifiability parameter, as guaranteed by Lemma III.5.

We just demonstrated that s-analytic random vectors cannot be s-rectifiable if they are also $(s+1)$-analytic. The question now arises whether s-analytic random vectors that fail to be $(s+1)$-analytic (and, therefore, fail to be t-analytic for all $t > s$ by Property i) in Lemma IV.3) are necessarily s-rectifiable. The next example shows that this is not the case.

Example IV.2. Let $\mathcal{C}$ be the middle third Cantor set [38, pp. xvi–xvii] and consider $\mathcal{U} = \{(c\ t)^\mathsf{T} : c \in \mathcal{C}, t \in [0,1]\} \subseteq \mathbb{R}^2$. Since $0 < \mathscr{H}^{\ln 2/\ln 3}(\mathcal{C}) < \infty$ [38, Example 4.5], it follows that the random vector $\mathbf{x}$ with distribution $\mu_{\mathbf{x}} = \pi \times (\lambda^1|_{[0,1]})$, where

$$\pi = \frac{\mathscr{H}^{\ln 2/\ln 3}|_{\mathcal{C}}}{\mathscr{H}^{\ln 2/\ln 3}(\mathcal{C})} \tag{80}$$

is the normalized Hausdorff measure on $\mathcal{C}$, is well defined. We now show that

i) $\mathbf{x}$ is 1-analytic;
ii) $\mathbf{x}$ is not 2-analytic;
iii) $\mathbf{x}$ is not 1-rectifiable.

To establish i), consider $\mathcal{B} \in \mathscr{B}(\mathbb{R}^2)$ with $\mathrm{P}[\mathbf{x} \in \mathcal{B}] > 0$. Now,

$$\begin{aligned} 0 &< \mu_{\mathbf{x}}(\mathcal{B}) && (81)\\ &= \int_{\mathcal{C}} \lambda^1\big(\{t \in [0,1] : (c\ t)^{\mathsf{T}} \in \mathcal{B}\}\big)\,\mathrm{d}\pi(c) && (82)\\ &\leq \int_{\mathcal{C}} \lambda^1\big(\{t \in \mathbb{R} : (c\ t)^{\mathsf{T}} \in \mathcal{B}\}\big)\,\mathrm{d}\pi(c), && (83) \end{aligned}$$

where in (82) we applied Corollary H.1 (with the finite measure spaces $(\mathbb{R}, \mathscr{B}(\mathbb{R}), \pi)$ and $(\mathbb{R}, \mathscr{B}(\mathbb{R}), \lambda^1|_{[0,1]})$ and (83) is by monotonicity of Lebesgue measure. Thus, by Lemma H.4, there must exist a $c_0 \in \mathcal{C}$ such that $\mathcal{A} := \{t \in \mathbb{R} : (c_0\ t)^{\mathsf{T}} \in \mathcal{B}\}$ satisfies $\lambda^1(\mathcal{A}) > 0$. Now, define the mapping $h\colon \mathbb{R} \to \mathbb{R}^2$, $t \mapsto (c_0\ t)^{\mathsf{T}}$ and note that this mapping is (trivially) real analytic with $Jh \equiv 1$. Moreover, $h(\mathcal{A}) \subseteq \mathcal{B}$ by construction, and $\mathcal{A}$ is Borel measurable as the inverse image of the Borel set $\mathcal{B}$ under the real analytic and, therefore, continuous mapping h. Thus, $\mathbf{x}$ is 1-analytic.

We next show that $\mathbf{x}$ is not 2-analytic. Towards a contradiction, suppose that $\mathbf{x}$ is 2-analytic. Since $\mu_{\mathbf{x}}(\mathcal{U}) = 1$, by 2-analyticity of $\mathbf{x}$, there must exist a set $\mathcal{D} \in \mathscr{B}(\mathbb{R}^2)$ with $\lambda^2(\mathcal{D}) > 0$ and a real-analytic mapping $g\colon \mathbb{R}^2 \to \mathbb{R}^2$ of 2-dimensional Jacobian $Jg \not\equiv 0$ such that $g(\mathcal{D}) \subseteq \mathcal{U}$. By Property ii) in Lemma K.8, we can assume, w.l.o.g., that $g|_{\mathcal{D}}$ is an embedding. It follows that

$$\begin{aligned} \mathscr{H}^2(g(\mathcal{D})) &= \int_{\mathcal{D}} Jg(\boldsymbol{z})\,\mathrm{d}\lambda^2(\boldsymbol{z}) && (84)\\ &> 0, && (85) \end{aligned}$$

where in (84) we applied the area formula Corollary H.3 upon noting that $g|_{\mathcal{D}}$ is one-to-one as an embedding and locally Lipschitz by real analyticity of g, and (85) is by Lemma H.4, $\lambda^2(\mathcal{D}) > 0$, and $Jg(\boldsymbol{z}) > 0$ for all $\boldsymbol{z} \in \mathcal{D}$. Since $g(\mathcal{D}) \subseteq \mathcal{U}$ and $\mathscr{H}^2(g(\mathcal{D})) > 0$, monotonicity of $\mathscr{H}^2$ yields $\mathscr{H}^2(\mathcal{U}) > 0$. Upon noting that $\mathscr{H}^{1+\ln 2/\ln 3}(\mathcal{U}) < \infty$ [38, Example 4.3], this results in a contradiction to Property i) in Lemma H.3.

Finally, to establish iii), towards a contradiction, suppose that $\mathbf{x}$ is 1-rectifiable. Then, Lemma III.4 implies that $\mathbf{x}$ admits a countably 1-rectifiable support set. As every countably 1-rectifiable set is the countable union of 1-rectifiable sets, the union bound implies that there must exist a 1-rectifiable set $\mathcal{V}$ with $\mathrm{P}[\mathbf{x} \in \mathcal{V}] > 0$. By Definition III.1, there must therefore exist a compact set $\mathcal{K} \subseteq \mathbb{R}$ and a Lipschitz mapping $f\colon \mathcal{K} \to \mathbb{R}^2$ such that $\mathcal{V} = f(\mathcal{K})$. It follows that

$$\begin{aligned} 0 &< \mu_{\mathbf{x}}(\mathcal{V}) && (86)\\ &= \mu_{\mathbf{x}}(f(\mathcal{K})) && (87)\\ &= \int_{\mathcal{C}} \lambda^1\big(\{t \in [0,1] : (c\ t)^{\mathsf{T}} \in f(\mathcal{K})\}\big)\,\mathrm{d}\pi(c), && (88)\\ &= \int_{\mathcal{C}} \lambda^1\big(\{t \in [0,1] : (c\ t)^{\mathsf{T}} \in f(\mathcal{A}_c)\}\big)\,\mathrm{d}\pi(c), && (89) \end{aligned}$$

where in (88) we applied Corollary H.1 (with the finite measure spaces $(\mathbb{R}, \mathscr{B}(\mathbb{R}), \pi)$ and $(\mathbb{R}, \mathscr{B}(\mathbb{R}), \lambda^1|_{[0,1]})$ and in (89) we set, for every $c \in \mathcal{C}$,

$$\mathcal{A}_c = f^{-1}\big(\{(c\ t)^{\mathsf{T}} : t \in [0,1]\}\big) \subseteq \mathcal{K}. \tag{90}$$

Note that the sets $\mathcal{A}_c \subseteq \mathcal{K}$ are pairwise disjoint as inverse images of pairwise disjoint sets. Now, Lemma H.4 together with (86)–(89) implies that there must exist a set $\mathcal{F} \subseteq \mathcal{C}$ with $\pi(\mathcal{F}) > 0$ such that

$$\lambda^1\big(\{t \in [0,1] : (c\ t)^{\mathsf{T}} \in f(\mathcal{A}_c)\}\big) > 0 \quad \text{for all } c \in \mathcal{F}. \tag{91}$$

Since $\pi = \mathscr{H}^{\ln 2/\ln 3}|_{\mathcal{C}} / \mathscr{H}^{\ln 2/\ln 3}(\mathcal{C})$ and $0 < \pi(\mathcal{F}) \leq 1$, the definition of Hausdorff dimension (see Definition H.4) implies $\dim_{\mathrm{H}}(\mathcal{F}) = \ln 2/\ln 3$. As every countable set has Hausdorff dimension zero [38, p. 48], we conclude that $\mathcal{F}$ must be uncountable. Moreover,

$$\begin{aligned} \lambda^1(\mathcal{A}_c) &= \mathscr{H}^1(\mathcal{A}_c) && (92)\\ &\geq \frac{1}{L}\mathscr{H}^1(f(\mathcal{A}_c)) && (93)\\ &\geq \frac{1}{L}\mathscr{H}^1\big(\{t \in [0,1] : (c\ t)^{\mathsf{T}} \in f(\mathcal{A}_c)\}\big) && (94)\\ &= \frac{1}{L}\lambda^1\big(\{t \in [0,1] : (c\ t)^{\mathsf{T}} \in f(\mathcal{A}_c)\}\big) && (95)\\ &> 0 \quad \text{for all } c \in \mathcal{F}, && (96) \end{aligned}$$

where (92) and (95) follow from Property iii) in Lemma H.3, (93) is by Property ii) in Lemma H.3 with L the Lipschitz constant of f, (94) is again by Property ii) in Lemma H.3 with the Lipschitz constant of the projection $\boldsymbol{e}_2^{\mathsf{T}}\colon \mathbb{R}^2 \to \mathbb{R}, (c\ t)^{\mathsf{T}} \mapsto t$ equal to one, and in (96) we used (91). As the sets $\mathcal{A}_c$ are pairwise disjoint subsets of positive Lebesgue measure of the compact set $\mathcal{K}$, it follows that

$$\sup_{\mathcal{E} \subseteq \mathcal{F} : |\mathcal{E}| < \infty} \sum_{c \in \mathcal{E}} \lambda^1(\mathcal{A}_c) \leq \lambda^1(\mathcal{K}) < \infty, \tag{97}$$

which, by Lemma H.11, contradicts the uncountability of $\mathcal{F}$. Therefore, $\mathbf{x}$ cannot be 1-rectifiable.

Our strong converse for analytic random vectors will be based on the following result.

Theorem IV.1. Let $\mathcal{A} \in \mathscr{B}(\mathbb{R}^s)$ be of positive Lebesgue measure, $h\colon \mathbb{R}^s \to \mathbb{R}^m$, with $s \leq m$, real analytic of s-dimensional Jacobian $Jh \not\equiv 0$, and $f\colon \mathbb{R}^m \to \mathbb{R}^n$ real analytic. If f is one-to-one on $h(\mathcal{A})$, then $n \geq s$.

Proof. See Section VII. □

With the help of Theorem IV.1, we can now prove the strong converse for s-analytic random vectors.

Corollary IV.2. For $\mathbf{x} \in \mathbb{R}^m$ s-analytic, $n \geq s$ is necessary for the existence of a measurement matrix $\boldsymbol{A} \in \mathbb{R}^{n \times m}$ and a Borel measurable mapping $g\colon \mathbb{R}^{n \times m} \times \mathbb{R}^n \to \mathbb{R}^m$ such that $\mathrm{P}[g(\boldsymbol{A}, \boldsymbol{A}\mathbf{x}) \neq \mathbf{x}] < 1$.

Proof. Suppose, to the contrary, that there exist a measurement matrix $\boldsymbol{A} \in \mathbb{R}^{n \times m}$ and a Borel measurable mapping $g\colon \mathbb{R}^{n \times m} \times \mathbb{R}^n \to \mathbb{R}^m$ satisfying $\mathrm{P}[g(\boldsymbol{A}, \boldsymbol{A}\mathbf{x}) \neq \mathbf{x}] < 1$ for $n < s$. Then, by Lemma IV.1, there must exist a $\mathcal{U} \in \mathscr{B}(\mathbb{R}^m)$ with $\mathrm{P}[\mathbf{x} \in \mathcal{U}] > 0$ such that $\boldsymbol{A}$ is one-to-one on $\mathcal{U}$. As $\mathrm{P}[\mathbf{x} \in \mathcal{U}] > 0$, the s-analyticity of $\mu_{\mathbf{x}}$ implies the existence of a set $\mathcal{A} \in \mathscr{B}(\mathbb{R}^s)$ of positive Lebesgue measure along with a real analytic mapping $h\colon \mathbb{R}^s \to \mathbb{R}^m$ of s-dimensional Jacobian $Jh \not\equiv 0$ such that $h(\mathcal{A}) \subseteq \mathcal{U}$. As $\boldsymbol{A}$ is one-to-one on $h(\mathcal{A})$ and linear mappings are trivially real-analytic, Theorem IV.1 implies that we must have $n \geq s$, which contradicts $n < s$. □

We next show that the s-rectifiable random vectors considered in Examples III.1–III.3 are all s-analytic with the analyticity parameter equal to the corresponding rectifiability parameter. We need the following result, which states that real analytic immersions preserve analyticity in the following sense.

Lemma IV.4. If $\mathbf{x} \in \mathbb{R}^m$ is s-analytic and $f\colon \mathbb{R}^m \to \mathbb{R}^k$, with $m \leq k$, is a real analytic immersion, then $f(\mathbf{x})$ is s-analytic.

Proof. See Appendix F. □

Example IV.3. We show that $\mathbf{x}$ in Example III.1 is s-analytic. To this end, we consider an arbitrary but fixed $\mathcal{U} \in \mathscr{B}(\mathbb{R}^m)$ with $\mu_{\mathbf{x}}(\mathcal{U}) > 0$ and establish the existence of a set $\mathcal{A} \in \mathscr{B}(\mathbb{R}^s)$ of positive Lebesgue measure and a real analytic mapping $h\colon \mathbb{R}^s \to \mathbb{R}^m$ of s-dimensional Jacobian $Jh \not\equiv 0$ such that $h(\mathcal{A}) \subseteq \mathcal{U}$. Since

$$0 < \mu_{\mathbf{x}}(\mathcal{U}) \tag{98}$$
$$= \mathrm{P}[(\boldsymbol{e}_{\mathsf{k}_1} \dots \boldsymbol{e}_{\mathsf{k}_s})\mathbf{z} \in \mathcal{U}] \tag{99}$$
$$\leq \sum_{1 \leq i_1 < \dots < i_s \leq m} \mathrm{P}[(\boldsymbol{e}_{i_1} \dots \boldsymbol{e}_{i_s})\mathbf{z} \in \mathcal{U}], \tag{100}$$

there must exist a set of indices $\{i_1, \dots, i_s\} \subseteq \{1, \dots, m\}$ with $i_1 < \dots < i_s$ such that

$$\mathrm{P}[u(\mathbf{z}) \in \mathcal{U}] > 0, \tag{101}$$

where $u\colon \mathbb{R}^s \to \mathbb{R}^m, \boldsymbol{z} \mapsto (\boldsymbol{e}_{i_1} \dots \boldsymbol{e}_{i_s})\boldsymbol{z}$. As $\mu_{\mathbf{z}} \ll \lambda^s$ by assumption, $\mathbf{z}$ is s-analytic thanks to Lemma IV.5 below. The mapping u is linear and, therefore, trivially real analytic. Furthermore,

$$Ju(\boldsymbol{z}) = \sqrt{\det\left((\boldsymbol{e}_{i_1} \dots \boldsymbol{e}_{i_s})^{\mathsf{T}}(\boldsymbol{e}_{i_1} \dots \boldsymbol{e}_{i_s})\right)} \tag{102}$$
$$= \sqrt{\det \boldsymbol{I}_s} \tag{103}$$
$$= 1 \quad \text{for all } \boldsymbol{z} \in \mathbb{R}^s, \tag{104}$$

where (102) follows from $Du(\boldsymbol{z}) = (\boldsymbol{e}_{i_1} \dots \boldsymbol{e}_{i_s})$ for all $\boldsymbol{z} \in \mathbb{R}^s$, which proves that u is an immersion. We can therefore employ Lemma IV.4 and conclude that $u(\mathbf{z})$ is s-analytic. Hence, Definition IV.2 together with (101) implies that there must exist a set $\mathcal{A} \in \mathscr{B}(\mathbb{R}^s)$ of positive Lebesgue measure and a real analytic mapping $h\colon \mathbb{R}^s \to \mathbb{R}^m$ of s-dimensional Jacobian $Jh \not\equiv 0$ such that $h(\mathcal{A}) \subseteq \mathcal{U}$.

Lemma IV.5. If $\mathbf{x} \in \mathbb{R}^m$ with $\mu_{\mathbf{x}} \ll \lambda^m$, then $\mathbf{x}$ is m-analytic.

Proof. We have to show that, for each $\mathcal{U} \in \mathscr{B}(\mathbb{R}^m)$ with $\mu_{\mathbf{x}}(\mathcal{U}) > 0$, we can find a set $\mathcal{A} \in \mathscr{B}(\mathbb{R}^m)$ of positive Lebesgue measure and a real analytic mapping $h\colon \mathbb{R}^m \to \mathbb{R}^m$ of m-dimensional Jacobian $Jh \not\equiv 0$ such that $h(\mathcal{A}) \subseteq \mathcal{U}$. For given such $\mathcal{U} \in \mathscr{B}(\mathbb{R}^m)$, simply take $\mathcal{A} = \mathcal{U}$ and h the identity mapping on $\mathbb{R}^m$. □

Example IV.4. We show that $\mathbf{x} = \mathbf{a} \otimes \mathbf{b} \in \mathbb{R}^{kl}$ as in Example III.2 is $(r + t - 1)$-analytic. To this end, let $\mathcal{U} \in \mathscr{B}(\mathbb{R}^{kl})$ with $\mu_{\mathbf{x}}(\mathcal{U}) > 0$ be arbitrary but fixed. We have to establish that there exist a set $\mathcal{A} \in \mathscr{B}(\mathbb{R}^{r+t-1})$ of positive Lebesgue measure and a real analytic mapping $h\colon \mathbb{R}^{r+t-1} \to \mathbb{R}^{kl}$ of $(r+t-1)$-dimensional Jacobian $Jh \not\equiv 0$ such that $h(\mathcal{A}) \subseteq \mathcal{U}$. Since

$$0 < \mu_{\mathbf{x}}(\mathcal{U}) \tag{105}$$
$$= \mathrm{P}[((\boldsymbol{e}_{\mathsf{p}_1} \dots \boldsymbol{e}_{\mathsf{p}_r})\mathbf{u}) \otimes ((\boldsymbol{e}_{\mathsf{q}_1} \dots \boldsymbol{e}_{\mathsf{q}_t})\mathbf{v}) \in \mathcal{U}] \tag{106}$$
$$= \mathrm{P}[((\boldsymbol{e}_{\mathsf{p}_1} \dots \boldsymbol{e}_{\mathsf{p}_r}) \otimes (\boldsymbol{e}_{\mathsf{q}_1} \dots \boldsymbol{e}_{\mathsf{q}_t}))(\mathbf{u} \otimes \mathbf{v}) \in \mathcal{U}] \tag{107}$$
$$\leq \sum_{\substack{1 \leq i_1 < \dots < i_r \leq k \\ 1 \leq j_1 < \dots < j_t \leq l}} \mathrm{P}[((\boldsymbol{e}_{i_1} \dots \boldsymbol{e}_{i_r}) \otimes (\boldsymbol{e}_{j_1} \dots \boldsymbol{e}_{j_t}))(\mathbf{u} \otimes \mathbf{v}) \in \mathcal{U}], \tag{108}$$

where (107) relies on [35, Lemma 4.2.10], there must exist a set of indices $\{i_1, \dots, i_r\} \subseteq \{1, \dots, m\}$ with $i_1 < \dots < i_r$ and a set of indices $\{j_1, \dots, j_t\} \subseteq \{1, \dots, m\}$ with $j_1 < \dots < j_t$ such that

$$\mathrm{P}[v(\mathbf{u} \otimes \mathbf{v}) \in \mathcal{U}] > 0, \tag{109}$$

where

$$v\colon \mathbb{R}^{rt} \to \mathbb{R}^{kl} \tag{110}$$
$$\boldsymbol{w} \mapsto ((\boldsymbol{e}_{i_1} \dots \boldsymbol{e}_{i_r}) \otimes (\boldsymbol{e}_{j_1} \dots \boldsymbol{e}_{j_t}))\boldsymbol{w}. \tag{111}$$

Since $\mu_{\mathbf{u}} \times \mu_{\mathbf{v}} \ll \lambda^{r+t}$ by assumption, it follows from Lemma IV.6 below that $\mathbf{u} \otimes \mathbf{v}$ is $(r + t - 1)$-analytic. The mapping v is linear and, therefore, trivially real analytic. Furthermore,

$$Jv(\boldsymbol{w}) = \sqrt{\det\left((\boldsymbol{E}_1 \otimes \boldsymbol{E}_2)^{\mathsf{T}}(\boldsymbol{E}_1 \otimes \boldsymbol{E}_2)\right)} \tag{112}$$
$$= \sqrt{\det\left((\boldsymbol{E}_1^{\mathsf{T}}\boldsymbol{E}_1) \otimes (\boldsymbol{E}_2^{\mathsf{T}}\boldsymbol{E}_2)\right)} \tag{113}$$
$$= \sqrt{\det(\boldsymbol{I}_r \otimes \boldsymbol{I}_t)} \tag{114}$$
$$= 1 \quad \text{for all } \boldsymbol{w} \in \mathbb{R}^{rt}, \tag{115}$$

where (112) follows from $Dv(\boldsymbol{w}) = \boldsymbol{E}_1 \otimes \boldsymbol{E}_2$ for all $\boldsymbol{w} \in \mathbb{R}^{rt}$ with $\boldsymbol{E}_1 = (\boldsymbol{e}_{i_1} \dots \boldsymbol{e}_{i_r})$ and $\boldsymbol{E}_2 = (\boldsymbol{e}_{j_1} \dots \boldsymbol{e}_{j_t})$, and (113) relies on [35, Equation (4.2.4)] and [35, Lemma 4.2.10], which proves that v is an immersion. We can therefore employ Lemma IV.4 and conclude that $v(\mathbf{u} \otimes \mathbf{v})$ is $(r+t-1)$-analytic. Hence, Definition IV.2 together with (109) implies that there must exist a set $\mathcal{A} \in \mathscr{B}(\mathbb{R}^{r+t-1})$ of positive Lebesgue measure and a real analytic mapping $h\colon \mathbb{R}^{r+t-1} \to \mathbb{R}^{kl}$ of $(r + t - 1)$-dimensional Jacobian $Jh \not\equiv 0$ such that $h(\mathcal{A}) \subseteq \mathcal{U}$.

Lemma IV.6. If $\mathbf{a} \in \mathbb{R}^k$ and $\mathbf{b} \in \mathbb{R}^l$ with $\mu_{\mathbf{a}} \times \mu_{\mathbf{b}} \ll \lambda^{k+l}$, then $\mathbf{a} \otimes \mathbf{b}$ is $(k + l - 1)$-analytic.

Proof. See Appendix G. □

Example IV.5. Let $\mathbf{x}$, z, and h be as in Example III.3. We first note that sin and cos are real analytic. In fact, it follows from the ratio test [42, Theorem 3.34] that the power series

$$\sin(z) = \sum_{n=0}^{\infty} \frac{(-1)^n z^{2n+1}}{(2n+1)!} \tag{116}$$

$$\cos(z) = \sum_{n=0}^{\infty} \frac{(-1)^n z^{2n}}{(2n)!} \tag{117}$$

are absolutely convergent for all $z \in \mathbb{R}$. Thus, sin and cos can both be represented by convergent power series at 0 with infinite convergence radius. Lemma K.1 therefore implies that sin and cos are both real analytic. As each component of h is real analytic, so is h. Furthermore, $Jh(z) = \sqrt{\sin^2(z) + \cos^2(z)} = 1$ for all $z \in \mathbb{R}$, which implies that h is a real analytic immersion. Since z is 1-analytic by Lemma IV.5 and $\mathsf{x} = h(\mathsf{z})$, Lemma IV.4 implies that x is 1-analytic.

V. Proof of Theorem II.1 (Achievability)

Suppose that $K(\mathsf{x}) < n$. It then follows from (12) that x must admit a support set $\mathcal{U} \subseteq \mathbb{R}^m$ with $\underline{\dim}_{\mathrm{MB}}(\mathcal{U}) < n$. We first construct a new support set $\mathcal{V} \subseteq \mathbb{R}^m$ for x as a countable union of compact sets satisfying $\underline{\dim}_{\mathrm{MB}}(\mathcal{V}) < n$. Based on this support set $\mathcal{V}$ we then prove the existence of a measurable decoder g satisfying $\mathrm{P}[g(\boldsymbol{A}, \boldsymbol{A}\mathsf{x}) \neq \mathsf{x}] = 0$. The construction of $\mathcal{V}$ starts by noting that, thanks to Property i) in Lemma H.15, $\underline{\dim}_{\mathrm{MB}}(\mathcal{U}) < n$ implies the existence of a covering $\{\mathcal{U}_i\}_{i\in\mathbb{N}}$ of $\mathcal{U}$ by nonempty compact sets $\mathcal{U}_i$ satisfying

$$\sup_{i\in\mathbb{N}} \underline{\dim}_{\mathrm{B}}(\mathcal{U}_i) < n. \tag{118}$$

For this covering, we set

$$\mathcal{V} = \bigcup_{i\in\mathbb{N}} \mathcal{U}_i, \tag{119}$$

and note that

$$\underline{\dim}_{\mathrm{MB}}(\mathcal{V}) = \inf\left\{ \sup_{i\in\mathbb{N}} \underline{\dim}_{\mathrm{B}}(\mathcal{V}_i) : \mathcal{V} \subseteq \bigcup_{i\in\mathbb{N}} \mathcal{V}_i \right\} \tag{120}$$

$$\leq \sup_{i\in\mathbb{N}} \underline{\dim}_{\mathrm{B}}(\mathcal{U}_i) \tag{121}$$

$$< n, \tag{122}$$

where (120) follows from Property i) in Lemma H.15 with the infimum taken over all coverings $\{\mathcal{V}_i\}_{i\in\mathbb{N}}$ of $\mathcal{V}$ by nonempty compact sets $\mathcal{V}_i$, (121) is by (119), and in (122) we used (118). Since $\dim_{\mathrm{MB}}(\mathbb{R}^m) = m$ by Property ii) of Lemma H.15, and $\underline{\dim}_{\mathrm{MB}}(\mathcal{V}) < n \leq m$, we must have $\mathcal{V} \subsetneq \mathbb{R}^m$. Now, $\mathcal{V}$ is a support set because it contains the support set $\mathcal{U}$ as a subset. Furthermore, since $\mathrm{P}[\mathsf{x} \in \mathcal{V}] = 1$ and $\mathcal{V} \subsetneq \mathbb{R}^m$, there must exist an $\boldsymbol{e} \in \mathbb{R}^m \setminus \mathcal{V}$ such that $\mathrm{P}[\mathsf{x} = \boldsymbol{e}] = 0$. This $\boldsymbol{e}$ will be used to declare a decoding error. We will show in Section V-A that there exists a Borel measurable mapping $g\colon \mathbb{R}^{n\times m} \times \mathbb{R}^n \to \mathbb{R}^m$ satisfying

$$g(\boldsymbol{A}, \boldsymbol{y}) \begin{cases} \in \{\boldsymbol{v} \in \mathcal{V} : \boldsymbol{A}\boldsymbol{v} = \boldsymbol{y}\} & \text{if } \exists \boldsymbol{v} \in \mathcal{V} : \boldsymbol{A}\boldsymbol{v} = \boldsymbol{y} \\ = \boldsymbol{e} & \text{else.} \end{cases} \tag{123}$$

The mapping g is guaranteed to deliver a $\boldsymbol{v} \in \mathcal{V}$ that is consistent with $(\boldsymbol{A}, \boldsymbol{y})$ (in the sense of $\boldsymbol{A}\boldsymbol{v} = \boldsymbol{y}$) if at least one such consistent $\boldsymbol{v} \in \mathcal{V}$ exists, otherwise an error is declared by delivering the "error symbol" $\boldsymbol{e}$. Next, for each $\boldsymbol{A} \in \mathbb{R}^{n\times m}$, let $p_{\mathrm{e}}(\boldsymbol{A})$ denote the probability of error defined as

$$p_{\mathrm{e}}(\boldsymbol{A}) = \mathrm{P}[g(\boldsymbol{A}, \boldsymbol{A}\mathsf{x}) \neq \mathsf{x}]. \tag{124}$$

It remains to show that $p_{\mathrm{e}}(\boldsymbol{A}) = 0$ for $\lambda^{n\times m}$-a.a. $\boldsymbol{A}$. Now,

$$p_{\mathrm{e}}(\boldsymbol{A}) \tag{125}$$

$$= \mathrm{P}[g(\boldsymbol{A}, \boldsymbol{A}\mathsf{x}) \neq \mathsf{x}, \mathsf{x} \in \mathcal{V}] + \mathrm{P}[g(\boldsymbol{A}, \boldsymbol{A}\mathsf{x}) \neq \mathsf{x}, \mathsf{x} \notin \mathcal{V}] \tag{126}$$

$$= \mathrm{P}[g(\boldsymbol{A}, \boldsymbol{A}\mathsf{x}) \neq \mathsf{x}, \mathsf{x} \in \mathcal{V}] \tag{127}$$

$$= \mathrm{P}[(\boldsymbol{A}, \mathsf{x}) \in \mathcal{A}] \quad \text{for all } \boldsymbol{A} \in \mathbb{R}^{n\times m}, \tag{128}$$

where (127) follows from $\mathrm{P}[\mathsf{x} \in \mathcal{V}] = 1$ and in (128) we set

$$\mathcal{A} = \{(\boldsymbol{A}, \boldsymbol{x}) \in \mathbb{R}^{n\times m} \times \mathcal{V} : g(\boldsymbol{A}, \boldsymbol{A}\boldsymbol{x}) \neq \boldsymbol{x}\}. \tag{129}$$

Since $\mathcal{A} \in \mathscr{B}(\mathbb{R}^{n\times m}) \otimes \mathscr{B}(\mathbb{R}^m)$ by Lemma V.1 below (with $\mathcal{X} = \mathbb{R}^m$, $\mathcal{Y} = \mathbb{R}^{n\times m}$, $f(\boldsymbol{x}, \boldsymbol{A}) = g(\boldsymbol{A}, \boldsymbol{A}\boldsymbol{x})$, and $\mathcal{V} \in \mathscr{B}(\mathbb{R}^m)$ as a countable union of compact sets), we can apply Corollary H.1 (with the σ-finite measure spaces $(\mathbb{R}^m, \mathscr{B}(\mathbb{R}^m), \mu_{\mathsf{x}})$ and $(\mathbb{R}^{n\times m}, \mathscr{B}(\mathbb{R}^{n\times m}), \lambda^{n\times m})$) to $\mathcal{A}$ and get

$$\int_{\mathbb{R}^{n\times m}} p_{\mathrm{e}}(\boldsymbol{A}) \, \mathrm{d}\lambda^{n\times m}(\boldsymbol{A}) \tag{130}$$

$$= \int_{\mathbb{R}^m} \lambda^{n\times m}(\{\boldsymbol{A} : (\boldsymbol{A}, \boldsymbol{x}) \in \mathcal{A}\}) \, \mathrm{d}\mu_{\mathsf{x}}(\boldsymbol{x}). \tag{131}$$

Next, note that for $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$ with $\boldsymbol{x} \in \mathcal{V}$, the vector $g(\boldsymbol{A}, \boldsymbol{y})$ can differ from $\boldsymbol{x}$ only if there is a $\boldsymbol{v} \in \mathcal{V}\backslash\{\boldsymbol{x}\}$ that is consistent with $\boldsymbol{y}$, i.e., if $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{v}$ for some $\boldsymbol{v} \in \mathcal{V}\backslash\{\boldsymbol{x}\}$. Thus,

$$\mathcal{A} \subseteq \{(\boldsymbol{A}, \boldsymbol{x}) \in \mathbb{R}^{n\times m} \times \mathcal{V} : \ker(\boldsymbol{A}) \cap \mathcal{V}_{\boldsymbol{x}} \neq \{\boldsymbol{0}\}\}, \tag{132}$$

where, for each $\boldsymbol{x} \in \mathcal{V}$, we set

$$\mathcal{V}_{\boldsymbol{x}} = \{\boldsymbol{v} - \boldsymbol{x} : \boldsymbol{v} \in \mathcal{V}\}. \tag{133}$$

As (132) yields

$$\{\boldsymbol{A} \in \mathbb{R}^{n\times m} : (\boldsymbol{A}, \boldsymbol{x}) \in \mathcal{A}\} \tag{134}$$

$$\subseteq \big\{\boldsymbol{A} \in \mathbb{R}^{n\times m} : \ker(\boldsymbol{A}) \cap \mathcal{V}_{\boldsymbol{x}} \neq \{\boldsymbol{0}\}\big\}, \tag{135}$$

monotonicity of $\lambda^{n\times m}$ implies

$$\lambda^{n\times m}(\{\boldsymbol{A} \in \mathbb{R}^{n\times m} : (\boldsymbol{A}, \boldsymbol{x}) \in \mathcal{A}\}) \tag{136}$$

$$\leq \lambda^{n\times m}\big(\big\{\boldsymbol{A} \in \mathbb{R}^{n\times m} : \ker(\boldsymbol{A}) \cap \mathcal{V}_{\boldsymbol{x}} \neq \{\boldsymbol{0}\}\big\}\big) \tag{137}$$

for all $\boldsymbol{x} \in \mathcal{V}$. The null-space property Proposition II.1, with $\mathcal{U} = \mathcal{V}_{\boldsymbol{x}}$ and $\underline{\dim}_{\mathrm{MB}}(\mathcal{V}_{\boldsymbol{x}}) = \underline{\dim}_{\mathrm{MB}}(\mathcal{V}) < n$ ((lower) modified Minkowski dimension is invariant under translation, as seen by translating covering balls accordingly) now implies that (137) equals zero for all $\boldsymbol{x} \in \mathcal{V}$. Therefore, (136) must equal zero as well for all $\boldsymbol{x} \in \mathcal{V}$. We conclude that (131) must equal zero as the integrand is identically zero (recall that $\mathcal{V}$ is a support set of x), which, by (130)-(131) and Lemma H.4, implies that we must have $p_{\mathrm{e}}(\boldsymbol{A}) = 0$ for $\lambda^{n\times m}$-a.a. $\boldsymbol{A}$, thereby completing the proof. □

A. Existence of Borel Measurable g

Recall that i) $\mathcal{V} = \bigcup_{i\in\mathbb{N}} \mathcal{U}_i \subsetneq \mathbb{R}^m$, where $\mathcal{U}_i \subseteq \mathbb{R}^m$ is nonempty and compact for all $i \in \mathbb{N}$ and ii) the error symbol $\boldsymbol{e} \in \mathbb{R}^m \setminus \mathcal{V}$. We have to show that there exists a Borel measurable mapping $g\colon \mathbb{R}^{n\times m} \times \mathbb{R}^n \to \mathbb{R}^m$ such that

$$g(\boldsymbol{A}, \boldsymbol{y}) \begin{cases} \in \{\boldsymbol{v} \in \mathcal{V} : \boldsymbol{A}\boldsymbol{v} = \boldsymbol{y}\} & \text{if } \exists \boldsymbol{v} \in \mathcal{V} : \boldsymbol{A}\boldsymbol{v} = \boldsymbol{y} \\ = \boldsymbol{e} & \text{else.} \end{cases} \tag{138}$$

To this end, first consider the mapping

$$f\colon \mathbb{R}^{n\times m}\times\mathbb{R}^n\times\mathbb{R}^m \to \mathbb{R} \tag{139}$$
$$(\boldsymbol{A},\boldsymbol{y},\boldsymbol{v}) \mapsto \|\boldsymbol{y}-\boldsymbol{A}\boldsymbol{v}\|_2. \tag{140}$$

Since f is continuous, Lemma I.3 implies that f is a normal integrand (see Definition I.4) with respect to $\mathscr{B}(\mathbb{R}^{n\times m}\times\mathbb{R}^n)$. For each $i\in\mathbb{N}$, let

$$\mathcal{T}_i = \{(\boldsymbol{A},\boldsymbol{y})\in\mathbb{R}^{n\times m}\times\mathbb{R}^n : \exists \boldsymbol{u}\in\mathcal{U}_i \text{ with } f(\boldsymbol{A},\boldsymbol{y},\boldsymbol{u})\le 0\} \tag{141}$$
$$= \{(\boldsymbol{A},\boldsymbol{y})\in\mathbb{R}^{n\times m}\times\mathbb{R}^n : \exists \boldsymbol{u}\in\mathcal{U}_i \text{ with } \boldsymbol{A}\boldsymbol{u}=\boldsymbol{y}\}. \tag{142}$$

It now follows from Properties ii) and iii) of Lemma I.5 (with $\mathcal{T}=\mathbb{R}^{n\times m}\times\mathbb{R}^n$, $\alpha=0$, $\mathcal{K}=\mathcal{U}_i$, and f as in (139)–(140), which is a normal integrand with respect to $\mathscr{B}(\mathbb{R}^{n\times m}\times\mathbb{R}^n)$) that i) $\mathcal{T}_i\in\mathscr{B}(\mathbb{R}^{n\times m}\times\mathbb{R}^n)$ for all $i\in\mathbb{N}$ and ii) for every $i\in\mathbb{N}$, there exists a Borel measurable mapping

$$p_i\colon \mathcal{T}_i\to\mathbb{R}^m \tag{143}$$
$$(\boldsymbol{A},\boldsymbol{y})\mapsto p_i(\boldsymbol{A},\boldsymbol{y})\in\{\boldsymbol{u}\in\mathcal{U}_i : \boldsymbol{A}\boldsymbol{u}=\boldsymbol{y}\}. \tag{144}$$

For each $i\in\mathbb{N}$, the mapping p_i can be extended to a mapping $g_i\colon\mathbb{R}^{n\times m}\times\mathbb{R}^n\to\mathbb{R}^m$ by setting

$$g_i|_{\mathcal{T}_i} = p_i \tag{145}$$
$$g_i|_{(\mathbb{R}^{n\times m}\times\mathbb{R}^n)\setminus\mathcal{T}_i} = \boldsymbol{e}, \tag{146}$$

which is Borel measurable thanks to Lemma H.8 as p_i is Borel measurable and $\mathcal{T}_i\in\mathscr{B}(\mathbb{R}^{n\times m}\times\mathbb{R}^n)$. Based on this sequence $\{g_i\}_{i\in\mathbb{N}}$ of Borel measurable mappings g_i, we now construct a Borel measurable mapping satisfying (138). The idea underlying this construction is as follows. For a given pair $(\boldsymbol{A},\boldsymbol{y})$, we first use g_1 to try to find a consistent (in the sense of $\boldsymbol{y}=\boldsymbol{A}\boldsymbol{u}$) $\boldsymbol{u}\in\mathcal{U}_1$. If g_1 delivers the error symbol $\boldsymbol{e}$, we use g_2 to try to find a consistent $\boldsymbol{u}\in\mathcal{U}_2$. This procedure is continued until a g_i delivers a consistent $\boldsymbol{u}\in\mathcal{U}_i$. If no g_i yields a consistent $\boldsymbol{u}\in\mathcal{U}_i$, we deliver the error symbol $\boldsymbol{e}$ as the final decoder output. The formal construction is as follows. We set $G_1=g_1$ and, for every $i\in\mathbb{N}\setminus\{1\}$, we define the mapping $G_i\colon\mathbb{R}^{n\times m}\times\mathbb{R}^n\to\mathbb{R}^m$ iteratively by setting

$$G_i(\boldsymbol{A},\boldsymbol{y}) = \begin{cases} G_{i-1}(\boldsymbol{A},\boldsymbol{y}) & \text{if } G_{i-1}(\boldsymbol{A},\boldsymbol{y})\ne\boldsymbol{e} \\ g_i(\boldsymbol{A},\boldsymbol{y}) & \text{else.}\end{cases} \tag{147}$$

Then, G_1 $(=g_1)$ is Borel measurable, and, for each $i\in\mathbb{N}\setminus\{1\}$, the Borel-measurability of G_i follows from the Borel-measurability of G_{i-1} and g_i thanks to Lemma V.2 below. Note that by construction

$$G_i(\boldsymbol{A},\boldsymbol{y})\in\left\{\boldsymbol{v}\in\bigcup_{j=1}^{i}\mathcal{U}_j : \boldsymbol{A}\boldsymbol{v}=\boldsymbol{y}\right\} \tag{148}$$

if there exists a $\boldsymbol{v}\in\bigcup_{j=1}^{i}\mathcal{U}_j$ such that $\boldsymbol{A}\boldsymbol{v}=\boldsymbol{y}$ and $G_i(\boldsymbol{A},\boldsymbol{y})=\boldsymbol{e}$ else. Finally, we obtain $g\colon\mathbb{R}^{n\times m}\times\mathbb{R}^n\to\mathbb{R}^m$ according to

$$g(\boldsymbol{A},\boldsymbol{y}) = \lim_{i\to\infty} G_i(\boldsymbol{A},\boldsymbol{y}), \tag{149}$$

which satisfies (138) by construction. As the pointwise limit of a sequence of Borel measurable mappings, g is Borel measurable thanks to Corollary H.2. $\square$

Lemma V.1. Let $\mathcal{X}$ and $\mathcal{Y}$ be Euclidean spaces, consider a Borel measurable mapping

$$f\colon\mathcal{X}\times\mathcal{Y}\to\mathcal{X}, \tag{150}$$

and let $\mathcal{V}\in\mathscr{B}(\mathcal{X})$. Then,

$$\mathcal{A} = \{(x,y)\in\mathcal{V}\times\mathcal{Y} : f(x,y)\ne x\} \tag{151}$$
$$\in\mathscr{B}(\mathcal{X}\times\mathcal{Y}) = \mathscr{B}(\mathcal{X})\otimes\mathscr{B}(\mathcal{Y}). \tag{152}$$

Proof. We first note that $\mathscr{B}(\mathcal{X}\times\mathcal{X})=\mathscr{B}(\mathcal{X})\otimes\mathscr{B}(\mathcal{X})$ and $\mathscr{B}(\mathcal{X}\times\mathcal{Y})=\mathscr{B}(\mathcal{X})\otimes\mathscr{B}(\mathcal{Y})$, both thanks to Lemma H.5. Therefore, $\mathcal{V}\times\mathcal{X}\in\mathscr{B}(\mathcal{X}\times\mathcal{X})$. Now, consider the diagonal $\mathcal{D}=\{(x,x):x\in\mathcal{X}\}$ and note that $\mathcal{D}$ as the inverse image of $\{0\}$ under the Borel measurable mapping $g\colon\mathcal{X}\times\mathcal{X}\to\mathcal{X}$, $(u,v)\mapsto u-v$ is in $\mathscr{B}(\mathcal{X}\times\mathcal{X})$. Let $\mathcal{C}=(\mathcal{V}\times\mathcal{X})\cap((\mathcal{X}\times\mathcal{X})\setminus\mathcal{D})$. Since $\mathcal{D}\in\mathscr{B}(\mathcal{X}\times\mathcal{X})$, it follows that $\mathcal{C}\in\mathscr{B}(\mathcal{X}\times\mathcal{X})$. Define the mapping

$$F\colon\mathcal{X}\times\mathcal{Y}\to\mathcal{X}\times\mathcal{X} \tag{153}$$
$$(x,y)\mapsto(x,f(x,y)), \tag{154}$$

and note that it is Borel measurable thanks to Lemma H.7 (with $f_1\colon\mathcal{X}\times\mathcal{Y}\to\mathcal{X}$, $(x,y)\mapsto x$ and $f_2\colon\mathcal{X}\times\mathcal{Y}\to\mathcal{X}$, $(x,y)\mapsto f(x,y)$). Finally, $\mathcal{A}\in\mathscr{B}(\mathcal{X}\times\mathcal{Y})$ as $\mathcal{A}=F^{-1}(\mathcal{C})$. $\square$

Lemma V.2. Let $\mathcal{X}$ and $\mathcal{Y}$ be topological spaces and $y_0\in\mathcal{Y}$ and suppose that $f,g\colon\mathcal{X}\to\mathcal{Y}$ are both Borel measurable. Then, $h\colon\mathcal{X}\to\mathcal{Y}$,

$$h(x) = \begin{cases} f(x) & \text{if } f(x)\ne y_0 \\ g(x) & \text{else}\end{cases} \tag{155}$$

is Borel measurable.

Proof. We have to show that $h^{-1}(\mathcal{U})\in\mathscr{B}(\mathcal{X})$ for all $\mathcal{U}\in\mathscr{B}(\mathcal{Y})$. To this end, consider an arbitrary but fixed $\mathcal{U}\in\mathscr{B}(\mathcal{Y})$. Now, $\{y_0\}\in\mathscr{B}(\mathcal{Y})$ implies $\mathcal{U}\setminus\{y_0\}\in\mathscr{B}(\mathcal{Y})$. We write $h^{-1}(\mathcal{U})=\mathcal{A}\cup\mathcal{B}$ with

$$\mathcal{A} = \{x\in\mathcal{X} : h(x)\in\mathcal{U}, f(x)\ne y_0\} \tag{156}$$
$$\mathcal{B} = \{x\in\mathcal{X} : h(x)\in\mathcal{U}, f(x)= y_0\} \tag{157}$$

and show that $\mathcal{A}$ and $\mathcal{B}$ are both in $\mathscr{B}(\mathcal{X})$, which in turn implies $h^{-1}(\mathcal{U})\in\mathscr{B}(\mathcal{X})$. Since

$$\mathcal{A} = \{x\in\mathcal{X} : f(x)\in\mathcal{U}, f(x)\ne y_0\} \tag{158}$$
$$= f^{-1}(\mathcal{U}\setminus\{y_0\}), \tag{159}$$

$\mathcal{U}\setminus\{y_0\}\in\mathscr{B}(\mathcal{Y})$, and f is Borel measurable by assumption, it follows that $\mathcal{A}\in\mathscr{B}(\mathcal{X})$. Finally, as

$$\mathcal{B} = \{x\in\mathcal{X} : g(x)\in\mathcal{U}, f(x)= y_0\} \tag{160}$$
$$= f^{-1}(\{y_0\})\cap g^{-1}(\mathcal{U}), \tag{161}$$

$\{y_0\}\in\mathscr{B}(\mathcal{Y})$, $\mathcal{U}\in\mathscr{B}(\mathcal{Y})$, and f and g are both Borel measurable by assumption, it follows that $\mathcal{B}\in\mathscr{B}(\mathcal{X})$. Thus, $h^{-1}(\mathcal{U})=\mathcal{A}\cup\mathcal{B}\in\mathscr{B}(\mathcal{X})$. Since $\mathcal{U}$ was arbitrary, we conclude that h is Borel measurable. $\square$

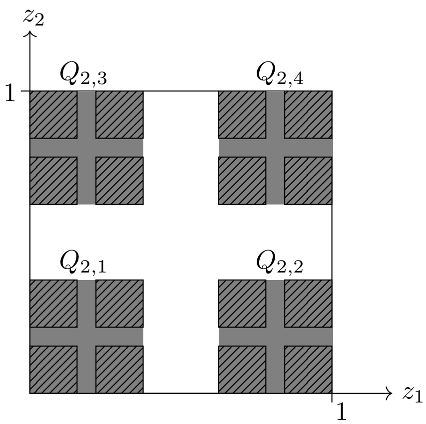

Figure 2. The set Q_2 consists of the four grey squares. The set Q_3 consists of the sixteen shaded black squares.

VI. Proof of Theorem III.2

Construction of $\mathcal{A}$. Consider the sequence $\{a_k\}_{k\in\mathbb{N}}$, where $a_k = 1/2 + 1/2^k$, and note that

$$\lim_{k\to\infty} a_k = \tfrac{1}{2} \tag{162}$$

$$a_k > a_{k+1} \quad \text{for all } k \in \mathbb{N}. \tag{163}$$

Let $\mathcal{Q}_1 = [0,1]^2$ be the unit square of side length one. We define

$$\mathcal{Q}_2 = \mathcal{Q}_{2,1} \cup \mathcal{Q}_{2,2} \cup \mathcal{Q}_{2,3} \cup \mathcal{Q}_{2,4}, \tag{164}$$

where every square $\mathcal{Q}_{2,i} \subseteq \mathcal{Q}_1$ has side length $a_2/2$ with $(0\ 0)^\mathsf{T} \in \mathcal{Q}_{2,1}$, $(1\ 0)^\mathsf{T} \in \mathcal{Q}_{2,2}$, $(0\ 1)^\mathsf{T} \in \mathcal{Q}_{2,3}$, and $(1\ 1)^\mathsf{T} \in \mathcal{Q}_{2,4}$. It follows from (163) that the squares $\mathcal{Q}_{2,1}, \dots, \mathcal{Q}_{2,4}$ are pairwise disjoint and $\mathcal{Q}_2 \subsetneq \mathcal{Q}_1$. To define $\mathcal{Q}_3$, we follow the same procedure and split up every set $\mathcal{Q}_{2,i}$ into the disjoint union of four squares with side length $a_3/4$. The sets $\mathcal{Q}_1$, $\mathcal{Q}_2$, and $\mathcal{Q}_3$ are depicted in Figure 2. We iterate this construction and obtain a sequence $\{\mathcal{Q}_k\}_{k\in\mathbb{N}}$, where $\mathcal{Q}_k$ is the disjoint union of 4^{k-1} squares $\mathcal{Q}_{k,i}, i = 1, \dots, 4^{k-1}$, of side length $a_k/2^{k-1}$ and

$$\mathcal{Q}_{k+1} \subsetneq \mathcal{Q}_k \quad \text{for all } k \in \mathbb{N}. \tag{165}$$

Next, we set

$$\mathcal{A} = \bigcap_{k\in\mathbb{N}} \mathcal{Q}_k, \tag{166}$$

which, as the intersection of closed sets, is closed. Since $\mathcal{A}$ is also bounded it must be compact by the Heine-Borel theorem [42, Theorem 2.41]. Finally,

$$\lambda^2(\mathcal{A}) = \lim_{k\to\infty} \lambda^2(\mathcal{Q}_k) \tag{167}$$

$$= \lim_{k\to\infty} \frac{4^{k-1} a_k^2}{(2^{k-1})^2} \tag{168}$$

$$= \lim_{k\to\infty} a_k^2 \tag{169}$$

$$= \tfrac{1}{4}, \tag{170}$$

where (167) follows from Property iii) of Lemma H.1.

Construction of κ. We now construct a C^∞-function $\kappa\colon \mathbb{R}^2 \to \mathbb{R}$ that is one-to-one on $\mathcal{A}$ as defined in (166). This will be accomplished by building compactly supported C^∞-functions $\varphi_{k,i}\colon \mathbb{R}^2 \to [0,1]$, $i = 1, \dots, 4^{k-1}$, $k \in \mathbb{N}$, such that

$$\varphi_{k,i}(\boldsymbol{z}) = \begin{cases} 1 & \text{if } \boldsymbol{z} \in \mathcal{Q}_{k,i}, \\ 0 & \text{if } \boldsymbol{z} \in \mathcal{Q}_{k,j},\ j \in \{1, \dots, 4^{k-1}\}\backslash\{i\}. \end{cases} \tag{171}$$

The construction of these functions is effected by Lemma VI.1 below with $\varphi_{k,i}(\boldsymbol{z}) = \psi_{\delta_{k,i}, a_{k,i}, \boldsymbol{w}_{k,i}}(\boldsymbol{z})$, where $\boldsymbol{w}_{k,i}$ denotes the center of $\mathcal{Q}_{k,i}$, $a_{k,i}$ equals half the side-length of $\mathcal{Q}_{k,i}$, and $\delta_{k,i}$ is chosen sufficiently small for (171) to hold (recall that the squares $\mathcal{Q}_{k,i}$ are closed and disjoint). Next, we define the C^∞-functions

$$\varphi_k = \sum_{i=1}^{4^{k-1}} \frac{i}{4^{k-1}} \varphi_{k,i} \tag{172}$$

and note that

$$\varphi_k(\boldsymbol{z}) = \frac{i}{4^{k-1}} \tag{173}$$

for all $\boldsymbol{z} \in \mathcal{Q}_{k,i}$, $i = 1, \dots, 4^{k-1}$, and $k \in \mathbb{N}$, and

$$|\varphi_k(\boldsymbol{z}) - \varphi_k(\boldsymbol{w})| = \frac{|i-j|}{4^{k-1}} \tag{174}$$

$$\geq \frac{1}{4^{k-1}} \tag{175}$$

for all $\boldsymbol{z} \in \mathcal{Q}_{k,i}$, $\boldsymbol{w} \in \mathcal{Q}_{k,j}$, $1 \leq i < j \leq 4^{k-1}$, and $k \in \mathbb{N}$. For $l \in \mathbb{N}$ and $a, b \in \mathbb{N}_0$, consider now the C^∞-function

$$s_l^{(a,b)}(\boldsymbol{z}) := \frac{\partial^{a+b}}{\partial z_1^a \partial z_2^b} \sum_{k=1}^{l} \frac{1}{8^{2^k}(M_k+1)} \varphi_k(\boldsymbol{z}) \tag{176}$$

$$= \sum_{k=1}^{l} \frac{1}{8^{2^k}(M_k+1)} \frac{\partial^{a+b}\varphi_k(\boldsymbol{z})}{\partial z_1^a \partial z_2^b} \tag{177}$$

with

$$M_k = \max_{1\leq i\leq k} \max_{1\leq j<k} d(i,j), \tag{178}$$

$$d(i,j) = \sup_{\boldsymbol{z}\in\mathbb{R}^2} \max\left\{ \left| \frac{\partial^j \varphi_i(\boldsymbol{z})}{\partial z_1^a \partial z_2^b} \right| : a, b \in \mathbb{N}_0,\ a+b=j \right\}. \tag{179}$$

We now show that this particular choice for the constants M_k guarantees, for each $a, b \in \mathbb{N}_0$, that the sequence $\{s_l^{(a,b)}\}_{l\in\mathbb{N}}$ of C^∞-functions converges uniformly and denote the corresponding limiting functions by $\kappa^{(a,b)}$. Corollary J.1 then implies

$$\frac{\partial \kappa^{(a,b)}(\boldsymbol{z})}{\partial z_1} = \kappa^{(a+1,b)}(\boldsymbol{z}) \tag{180}$$

$$\frac{\partial \kappa^{(a,b)}(\boldsymbol{z})}{\partial z_2} = \kappa^{(a,b+1)}(\boldsymbol{z}) \tag{181}$$

for all $a, b \in \mathbb{N}_0$ and $\boldsymbol{z} \in \mathbb{R}^2$, and $\kappa^{(0,0)}$ must therefore be C^∞. We set $\kappa = \kappa^{(0,0)}$. It remains to prove uniform convergence of the sequences $\{s_l^{(a,b)}\}_{l\in\mathbb{N}}$ of C^∞-functions for all $a, b \in \mathbb{N}_0$. To this end, let $a, b \in \mathbb{N}_0$ be arbitrary but fixed and note that

$$\frac{1}{8^{2^k}(M_k+1)} \left| \frac{\partial^{a+b}\varphi_k(\boldsymbol{z})}{\partial z_1^a \partial z_2^b} \right| \leq \frac{d(k, a+b)}{8^{2^k}(M_k+1)} \tag{182}$$

$$< \frac{1}{8^{2^k}} \tag{183}$$

for all $k > a+b$ and $\boldsymbol{z} \in \mathbb{R}^2$. Furthermore, by the sum formula for the geometric series,

$$\sum_{k=a+b+1}^{\infty} \frac{1}{8^{2^k}} < \sum_{k\in\mathbb{N}} \frac{1}{8^k} = \frac{1}{7}. \tag{184}$$

We can now conclude from (182) and (184) that the sequence $\{t_l^{(a,b)}\}_{l\in\mathbb{N}}$ of C^∞-functions

$$t_l^{(a,b)}(\boldsymbol{z}) := \sum_{k=a+b+1}^{a+b+l} \frac{1}{8^{2^k}(M_k+1)} \frac{\partial^{\,a+b}\varphi_k(\boldsymbol{z})}{\partial z_1^a \partial z_2^b} \tag{185}$$

satisfies the assumptions of Theorem J.1 and therefore converges uniformly to a function, which we denote by $\rho^{(a,b)}$. As

$$s_l^{(a,b)}(\boldsymbol{z}) = \sum_{k=1}^{a+b} \frac{1}{8^{2^k}(M_k+1)} \frac{\partial^{\,a+b}\varphi_k(\boldsymbol{z})}{\partial z_1^a \partial z_2^b} + t_{l-a-b}^{(a,b)}(\boldsymbol{z}) \tag{186}$$

for all $l > a+b$, we conclude that $\{s_l^{(a,b)}\}_{l\in\mathbb{N}}$ must converge uniformly to

$$\kappa^{(a,b)}(\boldsymbol{z}) := \sum_{k=1}^{a+b} \frac{1}{8^{2^k}(M_k+1)} \frac{\partial^{\,a+b}\varphi_k(\boldsymbol{z})}{\partial z_1^a \partial z_2^b} + \rho^{(a,b)}(\boldsymbol{z}). \tag{187}$$

Since a and b are arbitrary, this implies that $\{s_l^{(a,b)}\}_{l\in\mathbb{N}}$ converges uniformly for all $a, b \in \mathbb{N}_0$, thereby concluding the proof of κ being C^∞.

It remains to show that κ is one-to-one on $\mathcal{A}$. To this end, consider arbitrary but fixed $\boldsymbol{z}_0$ and $\boldsymbol{w}_0$ in $\mathcal{A}$ with $\boldsymbol{z}_0 \neq \boldsymbol{w}_0$. We have to show that $\kappa(\boldsymbol{z}_0) \neq \kappa(\boldsymbol{w}_0)$. Note that by construction of $\mathcal{A}$ (see (166)), there exists a $k_0 \in \mathbb{N}$ such that

i) for every $k \geq k_0$, there exist i_k and j_k in $\{1, \dots, 4^{k-1}\}$ with $i_k \neq j_k$ such that $\boldsymbol{z}_0 \in \mathcal{Q}_{k,i_k}$ and $\boldsymbol{w}_0 \in \mathcal{Q}_{k,j_k}$, and
ii) for every $k < k_0$, there exists an i_k such that $\boldsymbol{z}_0$ and $\boldsymbol{w}_0$ are both in $\mathcal{Q}_{k,i_k}$.

We therefore have

$$\begin{aligned}
&|\kappa(\boldsymbol{z}_0) - \kappa(\boldsymbol{w}_0)| && (188)\\
&= \left| \sum_{k\in\mathbb{N}} \frac{\varphi_k(\boldsymbol{z}_0) - \varphi_k(\boldsymbol{w}_0)}{8^{2^k}(M_k+1)} \right| && (189)\\
&= \left| \sum_{k=k_0}^{\infty} \frac{\varphi_k(\boldsymbol{z}_0) - \varphi_k(\boldsymbol{w}_0)}{8^{2^k}(M_k+1)} \right| && (190)\\
&\geq \frac{|\varphi_{k_0}(\boldsymbol{z}_0) - \varphi_{k_0}(\boldsymbol{w}_0)|}{8^{2^{k_0}}(M_{k_0}+1)} - \left| \sum_{k=k_0+1}^{\infty} \frac{\varphi_k(\boldsymbol{z}_0) - \varphi_k(\boldsymbol{w}_0)}{8^{2^k}(M_k+1)} \right| && (191)\\
&\geq \frac{|\varphi_{k_0}(\boldsymbol{z}_0) - \varphi_{k_0}(\boldsymbol{w}_0)|}{8^{2^{k_0}}(M_{k_0}+1)} - \sum_{k=k_0+1}^{\infty} \frac{|\varphi_k(\boldsymbol{z}_0) - \varphi_k(\boldsymbol{w}_0)|}{8^{2^k}(M_k+1)} && (192)\\
&\geq \frac{1}{M_{k_0}+1}\Bigg(\frac{|\varphi_{k_0}(\boldsymbol{z}_0) - \varphi_{k_0}(\boldsymbol{w}_0)|}{8^{2^{k_0}}} - \sum_{k=k_0+1}^{\infty} \frac{|\varphi_k(\boldsymbol{z}_0) - \varphi_k(\boldsymbol{w}_0)|}{8^{2^k}} \Bigg), && (193)
\end{aligned}$$

where (189) follows from the uniform convergence of $\{s_l^{(0,0)}\}_{l\in\mathbb{N}}$ to κ, in (190) we used (173) and ii) above, (191) is by the reverse triangle inequality, and (193) follows from $M_k \leq M_{k+1}$ for all $k \in \mathbb{N}$ (see (178)). Moreover, (174)–(175) imply $|\varphi_{k_0}(\boldsymbol{z}_0) - \varphi_{k_0}(\boldsymbol{w}_0)| \geq 1/4^{k_0-1}$ and (173) yields

$$\begin{aligned}
|\varphi_k(\boldsymbol{z}_0) - \varphi_k(\boldsymbol{w}_0)| &\leq |\varphi_k(\boldsymbol{z}_0)| + |\varphi_k(\boldsymbol{w}_0)| && (194)\\
&= \frac{i_k + j_k}{4^{k-1}} && (195)\\
&\leq 2 \quad \text{for all } k \in \mathbb{N}. && (196)
\end{aligned}$$

We can therefore further lower-bound $|\kappa(\boldsymbol{z}_0) - \kappa(\boldsymbol{w}_0)|$ according to

$$\begin{aligned}
&|\kappa(\boldsymbol{z}_0) - \kappa(\boldsymbol{w}_0)| && (197)\\
&\geq \frac{1}{M_{k_0}+1}\left(\frac{1}{4^{k_0-1}8^{2^{k_0}}} - \sum_{k=k_0+1}^{\infty} \frac{2}{8^{2^k}} \right) && (198)\\
&= \frac{1}{M_{k_0}+1}\left(\frac{1}{4^{k_0-1}8^{2^{k_0}}} - \sum_{k\in\mathbb{N}_0} \frac{2}{8^{2^{k+k_0+1}}} \right) && (199)\\
&\geq \frac{1}{M_{k_0}+1}\left(\frac{1}{4^{k_0-1}8^{2^{k_0}}} - \frac{2}{8^{2^{k_0+1}}} \sum_{k\in\mathbb{N}_0} \frac{1}{8^k} \right) && (200)\\
&> \frac{4}{(M_{k_0}+1)8^{2^{k_0}}} \left(\frac{1}{4^{k_0}} - \frac{1}{8^{2^{k_0}}} \right) && (201)\\
&> 0. && (202)
\end{aligned}$$

Since $\boldsymbol{z}_0$ and $\boldsymbol{w}_0$ are arbitrary, this establishes that κ is indeed one-to-one, thereby finishing the proof. ☐

Lemma VI.1. For $a > 0$, $\delta > 0$, and $\boldsymbol{w} = (w_1 \; w_2)^\mathsf{T} \in \mathbb{R}^2$, consider the mapping

$$\begin{aligned}
\psi_{\delta,a,\boldsymbol{w}} \colon \mathbb{R}^2 &\to [0,1] && (203)\\
\boldsymbol{z} &\mapsto \rho_{\delta,a}(z_1 - w_1)\rho_{\delta,a}(z_2 - w_2), && (204)
\end{aligned}$$

where

$$\begin{aligned}
\rho_{\delta,a} \colon \mathbb{R} &\to [0,1] && (205)\\
t &\mapsto \frac{f(a+\delta-|t|)}{f(a+\delta-|t|) + f(|t|-a)} && (206)
\end{aligned}$$

with $f(t) = e^{-1/t}\mathbb{1}_{\mathbb{R}_+}(t)$. Then, $\psi_{\delta,a,\boldsymbol{w}}$ is C^∞ and satisfies

$$\psi_{\delta,a,\boldsymbol{w}}(\boldsymbol{z}) = \begin{cases} 1 & \text{if } \max\{|z_1 - w_1|, |z_2 - w_2|\} \leq a, \\ 0 & \text{if } \min\{|z_1 - w_1|, |z_2 - w_2|\} \geq a + \delta. \end{cases} \tag{207}$$

Proof. It follows from [43, Lemma 2.22] with $H = \rho_{\delta,a}$, $r_1 = a$, and $r_2 = a + \delta$ that $\rho_{\delta,a}$ is C^∞ and satisfies

$$\rho_{\delta,a}(t) = \begin{cases} 1 & \text{if } |t| \in [0,a], \\ 0 & \text{if } |t| \in [a+\delta, \infty). \end{cases} \tag{208}$$

The claim now follows from $\psi_{\delta,a,\boldsymbol{w}} = \rho_{\delta,a}(z_1 - w_1)\rho_{\delta,a}(z_2 - w_2)$ and the properties of $\rho_{\delta,a}$. ☐

VII. Proof of Theorem IV.1 (Strong Converse)

Towards a contradiction, suppose that the statement is false. That is, we can find an $s \in \mathbb{N}$ such that there exist

i) an $\mathcal{A} \in \mathscr{B}(\mathbb{R}^s)$ with $\lambda^s(\mathcal{A}) > 0$,
ii) a real analytic mapping $h\colon \mathbb{R}^s \to \mathbb{R}^m$, with $s \leq m$, of s-dimensional Jacobian $Jh \not\equiv 0$, and
iii) an $n \in \mathbb{N}$ with $n < s$ and a real analytic mapping $f\colon \mathbb{R}^m \to \mathbb{R}^n$ that is one-to-one on $h(\mathcal{A})$.

Let s_0 be the smallest $s \in \mathbb{N}$ such that i)–iii) hold. The proof will be effected by showing that this implies validity of i)–iii) for $s_0 - 1$ and $n-1$, which contradicts the assumption that s_0 is the smallest natural number for i)–iii) to hold. The reader might wonder what happens to this argument in the case where $n = 1$. In fact we establish below that, if i)–iii) is satisfied, then necessarily $n \geq 2$, see the claims a)–c) right after (223).

Let $\mathcal{A}$, h, n, and f satisfy i)–iii) for s_0. We start by noting that

$$m \geq s_0 > n \geq 1 \tag{209}$$

by ii) and iii). Next, we write

$$f(\boldsymbol{x}) = (f_1(\boldsymbol{x}) \dots f_n(\boldsymbol{x}))^{\mathsf{T}} \tag{210}$$

and set $\psi_i = f_i \circ h$, $i = 1, \dots, n$, and $\psi = f \circ h$, which, by Corollary K.2, are all real analytic as compositions of real analytic mappings. We now show that there must exist an $i_0 \in \{1, \dots, n\}$ and a set $\mathcal{A}_{i_0} \subseteq \mathcal{A}$ such that

$$\lambda^{s_0}(\mathcal{A}_{i_0}) > 0, \tag{211}$$

$J\psi_{i_0}(\boldsymbol{z}) > 0$, and $Jh(\boldsymbol{z}) > 0$ for all $\boldsymbol{z} \in \mathcal{A}_{i_0}$. To this end, we first decompose

$$\mathcal{A} = \mathcal{A}_0 \cup \bigcup_{i=1}^{n} \mathcal{A}_i, \tag{212}$$

where

$$\mathcal{A}_i = \{\boldsymbol{z} \in \mathcal{A} : D\psi_i(\boldsymbol{z}) \neq \boldsymbol{0}, Jh(\boldsymbol{z}) > 0\} \tag{213}$$
$$= \{\boldsymbol{z} \in \mathcal{A} : J\psi_i(\boldsymbol{z}) > 0, Jh(\boldsymbol{z}) > 0\} \tag{214}$$

for $i = 1, \dots, n$ and

$$\mathcal{A}_0 = \{\boldsymbol{z} \in \mathcal{A} : D\psi(\boldsymbol{z}) = \boldsymbol{0}\} \cup \{\boldsymbol{z} \in \mathcal{A} : Jh(\boldsymbol{z}) = 0\}. \tag{215}$$

By Lemma VII.1 below (with $s = s_0$), $D\psi(\boldsymbol{z}) \neq \boldsymbol{0}$ for λ^{s_0}-a.a. $\boldsymbol{z} \in \mathcal{A}$. Furthermore, $Jh(\boldsymbol{z}) \neq 0$ for λ^{s_0}-a.a. $\boldsymbol{z} \in \mathcal{A}$ because of $Jh \not\equiv 0$ and Lemma K.7. Thus, $\lambda^{s_0}(\mathcal{A}_0) = 0$ by the countable subadditivity of Lebesgue measure. Since $\lambda^{s_0}(\mathcal{A}) > 0$ by assumption, and $\lambda^{s_0}(\mathcal{A}_0) = 0$, it follows, again by the countable subadditivity of Lebesgue measure, that there must exist an i_0 such that (211) holds.

Now, for each $y \in \mathbb{R}$, let

$$\mathcal{M}_y = \psi_{i_0}^{-1}(\{y\}). \tag{216}$$

We show in Section VII-A below that there exist a $y_0 \in \mathbb{R}$ and a $\boldsymbol{z}_0 \in \mathcal{A}_{i_0} \cap \mathcal{M}_{y_0}$ such that

$$\mathscr{H}^{s_0-1}\big(\mathcal{B}_{s_0}(\boldsymbol{z}_0, r) \cap \mathcal{A}_{i_0} \cap \mathcal{M}_{y_0}\big) > 0 \quad \text{for all } r > 0, \tag{217}$$
$$J\psi_{i_0}(\boldsymbol{z}) > 0 \quad \text{for all } \boldsymbol{z} \in \mathcal{M}_{y_0}, \tag{218}$$
$$Jh(\boldsymbol{z}_0) > 0. \tag{219}$$

It now follows from (218), real analyticity of ψ_{i_0}, $\mathcal{M}_{y_0} \neq \emptyset$, and Lemma K.11 that $\mathcal{M}_{y_0}$ is a $(s_0 - 1)$-dimensional real analytic submanifold of $\mathbb{R}^{s_0}$. Therefore, by Lemma K.9, there exist a real analytic embedding $\zeta\colon \mathbb{R}^{s_0-1} \to \mathbb{R}^{s_0}$ and an $\eta > 0$ such that

$$\zeta(\boldsymbol{0}) = \boldsymbol{z}_0, \tag{220}$$
$$\mathcal{B}_{s_0}(\boldsymbol{z}_0, \eta) \cap \mathcal{M}_{y_0} \subseteq \zeta\big(\mathbb{R}^{s_0-1}\big), \tag{221}$$

and $\zeta\big(\mathbb{R}^{s_0-1}\big)$ is relatively open in $\mathcal{M}_{y_0}$, i.e., there exists an open set $\mathcal{V} \subseteq \mathbb{R}^{s_0}$ with $\zeta\big(\mathbb{R}^{s_0-1}\big) = \mathcal{V} \cap \mathcal{M}_{y_0}$. Combining (217) and (221) yields

$$\mathscr{H}^{s_0-1}\big(\zeta\big(\mathbb{R}^{s_0-1}\big) \cap \mathcal{A}_{i_0}\big) > 0. \tag{222}$$

Next, let

$$\mathcal{C}_{i_0} = \zeta^{-1}\big(\zeta\big(\mathbb{R}^{s_0-1}\big) \cap \mathcal{A}_{i_0}\big). \tag{223}$$

We now show that

a) $\mathcal{C}_{i_0} \in \mathscr{B}(\mathbb{R}^{s_0-1})$ with $\lambda^{s_0-1}(\mathcal{C}_{i_0}) > 0$,
b) $\tilde{h} = h \circ \zeta\colon \mathbb{R}^{s_0-1} \to \mathbb{R}^m$ is real analytic and of (s_0-1)-dimensional Jacobian $J\tilde{h} \not\equiv 0$, and
c) $s_0 - 1 > n - 1 > 0$ and the real analytic mapping

$$\tilde{f}\colon \mathbb{R}^m \to \mathbb{R}^{n-1} \tag{224}$$
$$\boldsymbol{x} \mapsto (f_1(\boldsymbol{x}) \dots f_{i_0-1}(\boldsymbol{x})\, f_{i_0+1}(\boldsymbol{x}) \dots f_n(\boldsymbol{x}))^{\mathsf{T}} \tag{225}$$

is one-to-one on $\tilde{h}(\mathcal{C}_{i_0})$,

which finally yields the desired contradiction to the statement of s_0 being the smallest natural number such that i)–iii) at the beginning of the proof are satisfied.

Proof of a). We first establish that $\mathcal{C}_{i_0} \in \mathscr{B}(\mathbb{R}^{s_0-1})$. Since $\zeta\big(\mathbb{R}^{s_0-1}\big)$ is relatively open in $\mathcal{M}_{y_0}$ and, therefore, a Borel set in $\mathbb{R}^{s_0}$, it follows from (223) that $\mathcal{C}_{i_0}$ is the inverse image of a finite intersection of Borel sets under the real analytic embedding ζ and, therefore, also a Borel set. Next, we show that $\lambda^{s_0-1}(\mathcal{C}_{i_0}) > 0$. Since

$$\int_{\mathcal{C}_{i_0}} J\zeta(\boldsymbol{w})\, \mathrm{d}\lambda^{s_0-1}(\boldsymbol{w}) = \mathscr{H}^{s_0-1}(\zeta(\mathcal{C}_{i_0})) \tag{226}$$
$$= \mathscr{H}^{s_0-1}\big(\zeta(\mathbb{R}^{s_0-1}) \cap \mathcal{A}_{i_0}\big) \tag{227}$$
$$> 0, \tag{228}$$

where (226) follows from the area formula Corollary H.3 upon noting that ζ is one-to-one as an embedding and locally Lipschitz by real analyticity, (227) is by (223), and in (228) we applied (222). Using Lemma H.4, we conclude from (226)–(228) that $\lambda^{s_0-1}(\mathcal{C}_{i_0}) > 0$.

Proof of b). By Corollary K.2, $\tilde{h}$ is real analytic as the composition of real analytic mappings. It remains to show that $J\tilde{h} \not\equiv 0$. To this end, we establish $J\tilde{h}(\mathbf{0}) > 0$. First note that the chain rule yields

$$D\tilde{h}(\mathbf{0}) = (Dh)(\zeta(\mathbf{0}))D\zeta(\mathbf{0}) \tag{229}$$
$$= Dh(\boldsymbol{z}_0)D\zeta(\mathbf{0}), \tag{230}$$

where the second equality is by (220). Since $\zeta\colon \mathbb{R}^{s_0-1} \to \mathbb{R}^{s_0}$ is an embedding, it follows that

$$\operatorname{rank}(D\zeta(\mathbf{0})) = s_0 - 1. \tag{231}$$

Moreover, as $h\colon \mathbb{R}^{s_0} \to \mathbb{R}^m$ with $m \geq s_0$, (219) implies

$$\operatorname{rank}(Dh(\boldsymbol{z}_0)) = s_0. \tag{232}$$

Applying Lemma K.12 to $D\tilde{h}(\mathbf{0}) = Dh(\boldsymbol{z}_0)D\zeta(\mathbf{0})$ therefore yields $\operatorname{rank}(D\tilde{h}(\mathbf{0})) \geq s_0 - 1$, which in turn implies $J\tilde{h}(\mathbf{0}) > 0$ because $D\tilde{h}(\mathbf{0}) \in \mathbb{R}^{m\times(s_0-1)}$.

Proof of c). $s_0 - 1 > n - 1$ simply follows from $s_0 > n$ (see (209)). To prove that $\tilde{f}$ is one-to-one on $\tilde{h}(\mathcal{C}_{i_0})$, we first show that f_{i_0} is constant on $\tilde{h}(\mathcal{C}_{i_0})$. In fact, since $\zeta(\mathcal{C}_{i_0}) \subseteq \mathcal{M}_{y_0}$ and $\mathcal{M}_{y_0} = (f_{i_0} \circ h)^{-1}(\{y_0\})$, it follows that

$$f_{i_0}(\tilde{h}(\boldsymbol{w})) = y_0 \quad \text{for all } \boldsymbol{w} \in \mathcal{C}_{i_0}. \tag{233}$$

As f is one-to-one on $h(\mathcal{A})$ and f_{i_0} is constant on $\tilde{h}(\mathcal{C}_{i_0})$ with $\tilde{h}(\mathcal{C}_{i_0}) = (h \circ \zeta)(\mathcal{C}_{i_0}) \subseteq h(\mathcal{A})$ by (223), we conclude that $\tilde{f}$ must also be one-to-one on $\tilde{h}(\mathcal{C}_{i_0})$. It remains to show that we must have $n > 1$, which obviously implies $n-1 > 0$. Suppose, to the contrary, that $n = 1$. Since $\tilde{h}$ is real analytic with $J\tilde{h} \not\equiv 0$ and $\lambda^{s_0-1}(\mathcal{C}_{i_0}) > 0$, it follows from Property ii) of Lemma K.8 that $\tilde{h}$ is one-to-one on a subset of $\mathcal{C}_{i_0}$ of positive Lebesgue measure. Thus, the set $\tilde{h}(\mathcal{C}_{i_0})$ is uncountable. Now, since by (233) f_{i_0} is constant on $\tilde{h}(\mathcal{C}_{i_0})$, and a constant function cannot be one-to-one on a set of cardinality larger than one, the uncountability of $\tilde{h}(\mathcal{C}_{i_0})$ implies that f_{i_0} cannot be one-to-one on $\tilde{h}(\mathcal{C}_{i_0})$. But $f = f_{i_0}$ for $n = 1$ (with $i_0 = 1$) and f is one-to-one on $\tilde{h}(\mathcal{C}_{i_0}) \subseteq h(\mathcal{A})$ by assumption, which results in a contradiction. Therefore, we necessarily have $n > 1$ and we can conclude that a)–c) must hold, which finalizes the proof of the theorem.

A. *Proof of* (217)–(219)

We start by noting that

$$\int_{\mathbb{R}} \mathscr{H}^{s_0-1}(\mathcal{A}_{i_0} \cap \mathcal{M}_y)\,\mathrm{d}\lambda^1(y) \tag{234}$$
$$= \int_{\mathcal{A}_{i_0}} J\psi_{i_0}(\boldsymbol{z})\,\mathrm{d}\lambda^{s_0}(\boldsymbol{z}) \tag{235}$$
$$= \int_{\mathbb{R}^{s_0}} \mathbb{1}_{\mathcal{A}_{i_0}}(\boldsymbol{z}) J\psi_{i_0}(\boldsymbol{z})\,\mathrm{d}\lambda^{s_0}(\boldsymbol{z}) \tag{236}$$
$$> 0, \tag{237}$$

where (235) is by the coarea formula Corollary H.4 upon noting that ψ_{i_0} is locally Lipschitz by real analyticity, and (237) follows from $\mathbb{1}_{\mathcal{A}_{i_0}}(\boldsymbol{z}) J\psi_{i_0}(\boldsymbol{z}) > 0$ for all $\boldsymbol{z} \in \mathcal{A}_{i_0}$, $\lambda^{s_0}(\mathcal{A}_{i_0}) > 0$ (see (211)), and Lemma H.4. Also by Lemma H.4 and (234)–(237), there must exist a set $\mathcal{D} \subseteq \mathbb{R}$ with $\lambda^1(\mathcal{D}) > 0$ such that

$$\mathscr{H}^{s_0-1}(\mathcal{A}_{i_0} \cap \mathcal{M}_y) > 0 \quad \text{for all } y \in \mathcal{D}. \tag{238}$$

Furthermore,

$$\lambda^1\big(\{y \in \mathbb{R} : \exists \boldsymbol{z} \in \mathcal{M}_y \text{ with } J\psi_{i_0}(\boldsymbol{z}) = 0\}\big) \tag{239}$$
$$= \lambda^1\big(\psi_{i_0}\big(\{\boldsymbol{z} \in \mathbb{R}^{s_0} : J\psi_{i_0}(\boldsymbol{z}) = 0\}\big)\big) \tag{240}$$
$$= 0, \tag{241}$$

where (240) is by (216) and (241) follows from Property ii) of Theorem H.2 with ψ_{i_0} being C^∞ (recall that $\psi_{i_0} = f_{i_0} \circ h$ is real analytic as a composition of real analytic mappings and, therefore, C^∞ by Lemma K.3). Now, (239)–(241) together with $\lambda^1(\mathcal{D}) > 0$ implies the existence of a $y_0 \in \mathcal{D}$ such that (218) holds. For this y_0, we must have $\mathscr{H}^{s_0-1}(\mathcal{A}_{i_0} \cap \mathcal{M}_{y_0}) > 0$ by (238). Therefore, Lemma H.9 implies that there must exist a $\boldsymbol{z}_0 \in \mathcal{A}_{i_0} \cap \mathcal{M}_{y_0}$ such that (217) holds. As this $\boldsymbol{z}_0 \in \mathcal{A}_{i_0} \cap \mathcal{M}_{y_0} \subseteq \mathcal{A}_{i_0}$, (219) finally follows from (214). ☐

Lemma VII.1. Let $\mathcal{A} \in \mathscr{B}(\mathbb{R}^s)$ be of positive Lebesgue measure, $h\colon \mathbb{R}^s \to \mathbb{R}^m$, with $s \leq m$, real analytic of s-dimensional Jacobian $Jh \not\equiv 0$, and $f\colon \mathbb{R}^m \to \mathbb{R}^n$ real analytic. If f is one-to-one on $h(\mathcal{A})$, then $D(f \circ h)(\boldsymbol{z}) \neq \mathbf{0}$ for λ^s-a.a. $\boldsymbol{z} \in \mathcal{A}$.

Proof. Towards a contradiction, suppose that f is one-to-one on $h(\mathcal{A})$ and $D(f \circ h) \equiv \mathbf{0}$ on a set $\mathcal{B} \subseteq \mathcal{A}$ with $\lambda^s(\mathcal{B}) > 0$. Since $f \circ h$ is real analytic by Corollary K.2, $D(f \circ h)$ is real analytic as a consequence of Lemma K.3. It therefore follows from Lemma K.5 that $D(f \circ h)$ is identically zero. Hence, $f \circ h$ must be constant. In particular, $f \circ h$ must be constant on $\mathcal{B}$ and, therefore, f must be constant on $h(\mathcal{B})$. Since f is one-to-one and constant on $h(\mathcal{B})$, it follows that h must be constant on $\mathcal{B}$. Thus, $Jh(\boldsymbol{z}) = 0$ for all $\boldsymbol{z} \in \mathcal{B}$, and Lemma K.5 implies $Jh \equiv 0$. This is a contradiction to $Jh \not\equiv 0$. ☐

Appendix A
Proof of Lemma III.1

Proof of i). Suppose that $\mathcal{U}$ is s-rectifiable. Then, $\mathcal{U} = \varphi(\mathcal{A})$ with $\mathcal{A} \subseteq \mathbb{R}^s$ compact and φ Lipschitz. For $t \in \mathbb{N}$ with $t > s$, take

$$\mathcal{B} = \left\{ \begin{pmatrix} \boldsymbol{z} \\ \mathbf{0} \end{pmatrix} : \boldsymbol{z} \in \mathcal{A} \right\} \subseteq \mathbb{R}^t \tag{242}$$

and define

$$\psi\colon \mathcal{B} \to \mathbb{R}^m \tag{243}$$
$$\begin{pmatrix} \boldsymbol{z} \\ \mathbf{0} \end{pmatrix} \mapsto \varphi(\boldsymbol{z}). \tag{244}$$

Then, $\mathcal{U} = \psi(\mathcal{B})$ is t-rectifiable as a consequence of $\mathcal{B}$ compact and ψ Lipschitz.

Proof of ii). First note that we can cover $\mathbb{R}^s$ according to

$$\mathbb{R}^s = \bigcup_{k\in\mathbb{N}} \mathcal{A}_k, \tag{245}$$

with $\mathcal{A}_k \subseteq \mathbb{R}^s$ compact for all $k \in \mathbb{N}$ (take, for example, $\mathcal{A}_k = \overline{\mathcal{B}_s(\mathbf{0}, k)}$). Setting, for every $i, k \in \mathbb{N}$, $\varphi_{i,k} = \varphi_i|_{\mathcal{A}_k}$ (locally Lipschitz mappings are Lipschitz on compact sets by Lemma H.12), we can write

$$\mathcal{V} = \bigcup_{i,k\in\mathbb{N}} \varphi_{i,k}(\mathcal{A}_k), \tag{246}$$

which implies that $\mathcal{V}$ is countably s-rectifiable.

Proof of iii). Suppose that $\mathcal{U}$ is countably s-rectifiable and $\mathcal{V}$ is countably t-rectifiable. Then, we can write

$$\mathcal{U} = \bigcup_{i\in\mathbb{N}} \varphi_i(\mathcal{A}_i) \tag{247}$$

$$\mathcal{V} = \bigcup_{j\in\mathbb{N}} \psi_j(\mathcal{B}_j), \tag{248}$$

where the $\mathcal{A}_i \subseteq \mathbb{R}^s$ and the $\mathcal{B}_j \subseteq \mathbb{R}^t$ are compact and the $\varphi_i\colon \mathbb{R}^s \to \mathbb{R}^m$ and the $\psi_j\colon \mathbb{R}^t \to \mathbb{R}^n$ are Lipschitz. Thus,

$$\mathcal{W} = \bigcup_{i,j\in\mathbb{N}} \theta_{i,j}(\mathcal{C}_{i,j}), \tag{249}$$

where we defined

$$\theta_{i,j}\colon \mathbb{R}^{s+t} \to \mathbb{R}^{m+n} \tag{250}$$

$$(\boldsymbol{a}\ \boldsymbol{b})^{\mathsf{T}} \mapsto (\varphi_i(\boldsymbol{a})\ \psi_j(\boldsymbol{b}))^{\mathsf{T}} \tag{251}$$

and set

$$\mathcal{C}_{i,j} = \{(\boldsymbol{a}\ \boldsymbol{b})^{\mathsf{T}} : \boldsymbol{a} \in \mathcal{A}_i, \boldsymbol{b} \in \mathcal{B}_j\}. \tag{252}$$

Now, the $\theta_{i,j}$ are Lipschitz because the φ_i and the ψ_j are Lipschitz for all $i, j \in \mathbb{N}$, and the $\mathcal{C}_{i,j}$ are compact by Tychonoff's theorem [44, Theorem 4.42] thanks to $\mathcal{A}_i$ and $\mathcal{B}_j$ compact for all $i, j \in \mathbb{N}$. Therefore, $\mathcal{W}$ is countably $(s+t)$-rectifiable.

Proof of iv). Let $\mathcal{M}$ be an s-dimensional C^1-submanifold of $\mathbb{R}^m$. By [36, Definition 5.3.1], we can write

$$\mathcal{M} = \bigcup_{\boldsymbol{x}\in\mathcal{M}} \varphi_{\boldsymbol{x}}(\mathcal{U}_{\boldsymbol{x}}), \tag{253}$$

where, for every $\boldsymbol{x} \in \mathcal{M}$, $\mathcal{U}_{\boldsymbol{x}} \subseteq \mathbb{R}^s$ is open, and $\varphi_{\boldsymbol{x}}\colon \mathcal{U}_{\boldsymbol{x}} \to \mathbb{R}^m$ is C^1 and satisfies $\boldsymbol{x} \in \varphi_{\boldsymbol{x}}(\mathcal{U}_{\boldsymbol{x}})$ and $\varphi_{\boldsymbol{x}}(\mathcal{U}_{\boldsymbol{x}}) = \mathcal{V}_{\boldsymbol{x}} \cap \mathcal{M}$ with $\mathcal{V}_{\boldsymbol{x}} \subseteq \mathbb{R}^m$ open. Now, there must exist a countable set $\{\boldsymbol{x}_i : i \in \mathbb{N}\} \subseteq \mathcal{M}$ such that

$$\mathcal{M} = \bigcup_{i\in\mathbb{N}} \mathcal{V}_{\boldsymbol{x}_i} \cap \mathcal{M}. \tag{254}$$

For if such a countable set does not exist, the open set

$$\mathcal{V} = \bigcup_{\boldsymbol{x}\in\mathcal{M}} \mathcal{V}_{\boldsymbol{x}} \tag{255}$$

would not admit a countable subcover, which would contradict that $\mathcal{V}$ as an open set in $\mathbb{R}^m$ is Lindelöf [45, Definition 5.6.19, Proposition 5.6.22]. With the countable set $\{\boldsymbol{x}_i : i \in \mathbb{N}\} \subseteq \mathcal{M}$ we can now write

$$\mathcal{M} = \bigcup_{i\in\mathbb{N}} \varphi_i(\mathcal{U}_i), \tag{256}$$

where we set $\varphi_i = \varphi_{\boldsymbol{x}_i}$ and $\mathcal{U}_i = \mathcal{U}_{\boldsymbol{x}_i}$. Now fix $i \in \mathbb{N}$ arbitrarily. As $\mathcal{U}_i$ is open, for every $\boldsymbol{u}_i \in \mathcal{U}_i$, there exists an $r_{\boldsymbol{u}_i} > 0$ such that

$$\mathcal{B}^s(\boldsymbol{u}_i, r_{\boldsymbol{u}_i}) \subseteq \mathcal{U}_i. \tag{257}$$

We can thus write

$$\mathcal{U}_i = \bigcup_{\boldsymbol{u}_i\in\mathcal{U}_i} \mathcal{B}^s\Big(\boldsymbol{u}_i, \frac{r_{\boldsymbol{u}_i}}{2}\Big). \tag{258}$$

Since $\mathcal{U}_i$ as an open set in $\mathbb{R}^m$ is Lindelöf [45, Definition 5.6.19, Proposition 5.6.22], there exists a countable set $\{\boldsymbol{u}_{i,j} : j \in \mathbb{N}\} \subseteq \mathcal{U}_i$ such that

$$\mathcal{U}_i = \bigcup_{j\in\mathbb{N}} \mathcal{B}^s\Big(\boldsymbol{u}_{i,j}, \frac{r_{i,j}}{2}\Big), \tag{259}$$

where we set $r_{i,j} = r_{\boldsymbol{u}_{i,j}}$ for all $j \in \mathbb{N}$. Using $\mathcal{B}^s(\boldsymbol{u}_{i,j}, r_{i,j}) \subseteq \mathcal{U}_i$ for all $j \in \mathbb{N}$ (see (257)), it follows that

$$\mathcal{U}_i = \bigcup_{j\in\mathbb{N}} \overline{\mathcal{B}^s\Big(\boldsymbol{u}_{i,j}, \frac{r_{i,j}}{2}\Big)}. \tag{260}$$

Since $i \in \mathbb{N}$ was arbitrary, using (256) and (260), we get

$$\mathcal{M} = \bigcup_{i,j\in\mathbb{N}} \varphi_i\Big(\overline{\mathcal{B}^s\Big(\boldsymbol{u}_{i,j}, \frac{r_{i,j}}{2}\Big)}\Big). \tag{261}$$

Finally, all the φ_i as C^1 mappings are locally Lipschitz and, therefore, Lipschitz on the compact sets $\overline{\mathcal{B}^s\big(\boldsymbol{u}_{i,j}, \frac{r_{i,j}}{2}\big)}$ for all $j \in \mathbb{N}$, which establishes countable rectifiability of $\mathcal{M}$.

Proof of v). Since every s_i-rectifiable set, with $s_i \leq s$, is s-rectifiable by i), it follows that $\mathcal{A}$ is a countable union of s-rectifiable sets and, therefore, countably s-rectifiable. □

APPENDIX B
PROOF OF LEMMA III.3

Proof of i). Suppose that $\mathcal{U}$ is s-rectifiable. Then, there exist a compact set $\mathcal{A} \subseteq \mathbb{R}^s$ and a Lipschitz mapping $\varphi\colon \mathcal{A} \to \mathbb{R}^m$ such that $\mathcal{U} = \varphi(\mathcal{A})$. As $f \circ \varphi\colon \mathcal{A} \to \mathbb{R}^n$ is Lipschitz owing to Lemmata H.12 and H.14, $f(\mathcal{U}) = (f \circ \varphi)(\mathcal{A})$ is s-rectifiable.

Proof of ii). Suppose that $\mathcal{U}$ is countably s-rectifiable. We can write $\mathcal{U} = \bigcup_{i\in\mathbb{N}} \mathcal{U}_i$, where $\mathcal{U}_i$ is s-rectifiable for all $i \in \mathbb{N}$. As $f(\mathcal{U}) = \bigcup_{i\in\mathbb{N}} f(\mathcal{U}_i)$ and $f(\mathcal{U}_i)$ is s-rectifiable for all $i \in \mathbb{N}$ by i), it follows that $f(\mathcal{U})$ is countably s-rectifiable.

Proof of iii). Suppose that $\mathcal{U} \in \mathscr{B}(\mathbb{R}^m)$ is countably $(\mathscr{H}^s, s)$-rectifiable. We have to show that $f(\mathcal{U}) = \mathcal{A} \cup \mathcal{B}$, where $\mathcal{A} \in \mathscr{B}(\mathbb{R}^n)$ is countably $(\mathscr{H}^s, s)$-rectifiable and $\mathscr{H}^s(\mathcal{B}) = 0$. As $\mathcal{U} \in \mathscr{B}(\mathbb{R}^m)$ is countably $(\mathscr{H}^s, s)$-rectifiable, [46, Lemma 15.5] implies that $\mathscr{H}^s|_{\mathcal{U}}$ is σ-finite. We can therefore employ Lemma H.10 to decompose

$$\mathcal{U} = \mathcal{V}_0 \cup \bigcup_{j\in\mathbb{N}} \mathcal{V}_j, \tag{262}$$

where $\mathscr{H}^s(\mathcal{V}_0) = \mathscr{H}^s|_{\mathcal{U}}(\mathcal{V}_0) = 0$ and $\mathcal{V}_j \subseteq \mathbb{R}^m$ is compact for all $j \in \mathbb{N}$. This decomposition allows us to write $f(\mathcal{U}) = \mathcal{A} \cup \mathcal{B}$, where

$$\mathcal{A} = \bigcup_{j\in\mathbb{N}} f(\mathcal{V}_j) \tag{263}$$

$$\mathcal{B} = f(\mathcal{V}_0). \tag{264}$$

Now, thanks to Lemma H.13, the $f(\mathcal{V}_j)$ are compact. Therefore, $\mathcal{A}$ as a countable union of compact sets is Borel. Furthermore, $\mathscr{H}^s(\mathcal{B}) = 0$ because of $\mathscr{H}^s(\mathcal{V}_0) = 0$ and Lemma B.1 below. It remains to show that $\mathcal{A}$ is countably $(\mathscr{H}^s, s)$-rectifiable. Since $\mathcal{U}$ is countably $(\mathscr{H}^s, s)$-rectifiable, there exists a countably s-rectifiable set $\mathcal{U}_1$ such that $\mathscr{H}^s(\mathcal{U}\setminus$

$\mathcal{U}_1) = 0$. Furthermore, as $\mathcal{U}_1$ is countably s-rectifiable, $f(\mathcal{U}_1)$ is countably s-rectifiable by ii). Finally,

$$\mathscr{H}^s(\mathcal{A}\backslash f(\mathcal{U}_1)) \le \mathscr{H}^s(f(\mathcal{U})\backslash f(\mathcal{U}_1)) \tag{265}$$
$$\le \mathscr{H}^s(f(\mathcal{U}\backslash\mathcal{U}_1)) \tag{266}$$
$$= 0, \tag{267}$$

where (265) follows from the monotonicity of $\mathscr{H}^s$ and $f(\mathcal{U}) = \mathcal{A} \cup \mathcal{B}$, (266) is by monotonicity of $\mathscr{H}^s$ and $f(\mathcal{U})\backslash f(\mathcal{U}_1) \subseteq f(\mathcal{U}\backslash\mathcal{U}_1)$, and (267) follows from Lemma B.1 below. This proves that $\mathcal{A}$ is countably $(\mathscr{H}^s, s)$-rectifiable, thereby concluding the proof. □

Lemma B.1. If $\mathcal{U} \subseteq \mathbb{R}^m$ with $\mathscr{H}^s(\mathcal{U}) = 0$ and $f\colon \mathbb{R}^m \to \mathbb{R}^n$ is locally Lipschitz, then $\mathscr{H}^s(f(\mathcal{U})) = 0$.

Proof. We employ the following chain of arguments:

$$\mathscr{H}^s(f(\mathcal{U})) = \mathscr{H}^s\Big(f\Big(\mathcal{U} \cap \bigcup_{l\in\mathbb{N}} \mathcal{B}_m(\mathbf{0}, l)\Big)\Big) \tag{268}$$
$$= \mathscr{H}^s\Big(\bigcup_{l\in\mathbb{N}} f(\mathcal{B}_m(\mathbf{0}, l) \cap \mathcal{U})\Big) \tag{269}$$
$$\le \sum_{l\in\mathbb{N}} \mathscr{H}^s(f(\mathcal{B}_m(\mathbf{0}, l) \cap \mathcal{U})) \tag{270}$$
$$\le \sum_{l\in\mathbb{N}} L_l^s \mathscr{H}^s(\mathcal{B}_m(\mathbf{0}, l) \cap \mathcal{U}) \tag{271}$$
$$\le \sum_{l\in\mathbb{N}} L_l^s \mathscr{H}^s(\mathcal{U}) \tag{272}$$
$$= 0, \tag{273}$$

where (270) follows from the countable subadditivity of $\mathscr{H}^s$, in (271) we applied Property ii) of Lemma H.3, where, for every $l \in \mathbb{N}$, L_l denotes the Lipschitz constant of $f|_{\overline{\mathcal{B}_m(\mathbf{0},l)}}$, and in (272) we used the monotonicity of $\mathscr{H}^s$. □

Appendix C
Proof of Lemma III.5

Towards a contradiction, suppose that there exist $r, t \in \mathbb{N}$ with $r \neq t$ such that x is r-rectifiable and t-rectifiable. We can assume, w.l.o.g., that $r < t$. Now, Lemma III.4 implies the existence of

i) a countably r-rectifiable set $\mathcal{U}$ satisfying $\mathrm{P}[\mathsf{x} \in \mathcal{U}] = 1$ and
ii) a countably t-rectifiable set $\mathcal{V}$ satisfying $\mathrm{P}[\mathsf{x} \in \mathcal{V}] = 1$.

With Definition III.1, we can conclude that there exist

i) compact sets $\mathcal{A}_i \subseteq \mathbb{R}^r$ and Lipschitz mappings $\varphi_i\colon \mathcal{A}_i \to \mathbb{R}^m$, $i \in \mathbb{N}$, such that
$$\mathcal{U} = \bigcup_{i\in\mathbb{N}} \varphi_i(\mathcal{A}_i), \tag{274}$$
and
ii) compact sets $\mathcal{B}_j \subseteq \mathbb{R}^t$ and Lipschitz mappings $\psi_j\colon \mathcal{B}_j \to \mathbb{R}^m$, $j \in \mathbb{N}$, such that
$$\mathcal{V} = \bigcup_{j\in\mathbb{N}} \psi_j(\mathcal{B}_j). \tag{275}$$

The union bound now yields

$$\mathrm{P}[\mathsf{x} \in \mathcal{U} \cap \mathcal{V}] = \mathrm{P}\Big[\mathsf{x} \in \bigcup_{i\in\mathbb{N}} \bigcup_{j\in\mathbb{N}} \varphi_i(\mathcal{A}_i) \cap \psi_j(\mathcal{B}_j)\Big] \tag{276}$$
$$\le \sum_{i\in\mathbb{N}} \sum_{j\in\mathbb{N}} \mathrm{P}[\mathsf{x} \in \varphi_i(\mathcal{A}_i) \cap \psi_j(\mathcal{B}_j)]. \tag{277}$$

Since $\mathrm{P}[\mathsf{x} \in \mathcal{U} \cap \mathcal{V}] = 1$, which follows from $\mathrm{P}[\mathsf{x} \in \mathcal{U}] = 1$ and $\mathrm{P}[\mathsf{x} \in \mathcal{V}] = 1$, (276)–(277) guarantee the existence of an i_0 and a j_0, both in $\mathbb{N}$, such that $\mu_{\mathsf{x}}(\varphi_{i_0}(\mathcal{A}_{i_0}) \cap \psi_{j_0}(\mathcal{B}_{j_0})) > 0$. As $\mu_{\mathsf{x}} \ll \mathscr{H}^t|_{\mathcal{V}}$ by the t-rectifiability of x and $\mathscr{H}^t|_{\mathcal{V}} \ll \mathscr{H}^t$ by the monotonicity of $\mathscr{H}^t$, it follows that $\mathscr{H}^t(\varphi_{i_0}(\mathcal{A}_{i_0}) \cap \psi_{j_0}(\mathcal{B}_{j_0})) > 0$. Property i) of Lemma H.3 therefore implies (recall that $r < t$ by assumption) $\mathscr{H}^r(\varphi_{i_0}(\mathcal{A}_{i_0}) \cap \psi_{j_0}(\mathcal{B}_{j_0})) = \infty$. The monotonicity of $\mathscr{H}^r$ then yields

$$\mathscr{H}^r(\varphi_{i_0}(\mathcal{A}_{i_0})) \ge \mathscr{H}^r(\varphi_{i_0}(\mathcal{A}_{i_0}) \cap \psi_{j_0}(\mathcal{B}_{j_0})) \tag{278}$$
$$= \infty. \tag{279}$$

But we also have

$$\mathscr{H}^r(\varphi_{i_0}(\mathcal{A}_{i_0})) \le L^r \mathscr{H}^r(\mathcal{A}_{i_0}) \tag{280}$$
$$= L^r \lambda^r(\mathcal{A}_{i_0}) \tag{281}$$
$$< \infty, \tag{282}$$

where (280) follows from Property ii) of Lemma H.3 with L the Lipschitz constant of φ_{i_0}, (281) is by Property iii) of Lemma H.3 and the fact that $\mathcal{A}_{i_0}$ as a compact set is in $\mathscr{B}(\mathbb{R}^r)$, and finally (282) is a consequence, again, of $\mathcal{A}_{i_0}$ being compact. This contradicts (279) and thereby concludes the proof. □

Appendix D
Proof of Lemma IV.1

Suppose that there exists a Borel measurable mapping $g\colon \mathbb{R}^{n\times m} \times \mathbb{R}^n \to \mathbb{R}^m$ such that $\mathrm{P}[g(\boldsymbol{A}, \boldsymbol{A}\mathsf{x}) \neq \mathsf{x}] < 1$ and set $\mathcal{U} = \{\boldsymbol{x} \in \mathbb{R}^m : g(\boldsymbol{A}, \boldsymbol{A}\boldsymbol{x}) = \boldsymbol{x}\}$. We first show that $\mathcal{U} \in \mathscr{B}(\mathbb{R}^m)$. To this end, consider the mapping $h_{\boldsymbol{A}}\colon \mathbb{R}^m \to \mathbb{R}^m$, $\boldsymbol{x} \mapsto g(\boldsymbol{A}, \boldsymbol{A}\boldsymbol{x})$, which as the composition of two Borel measurable mappings, namely, $\boldsymbol{x} \mapsto (\boldsymbol{A}, \boldsymbol{A}\boldsymbol{x})$ and g, is Borel measurable. Application of Lemma H.7 (with $\mathcal{X} = \mathcal{Y} = \mathbb{R}^m$, f_1 the identity mapping on $\mathbb{R}^m$, and $f_2 = h_{\boldsymbol{A}}$) therefore establishes that

$$F\colon \mathbb{R}^m \to \mathbb{R}^m \times \mathbb{R}^m \tag{283}$$
$$\boldsymbol{x} \mapsto (\boldsymbol{x}, h_{\boldsymbol{A}}(\boldsymbol{x})) \tag{284}$$

is Borel measurable. Now, consider the diagonal $\mathcal{D} = \{(\boldsymbol{x}, \boldsymbol{x}) : \boldsymbol{x} \in \mathbb{R}^m\}$ and note that $\mathcal{D}$ as the inverse image of $\{\mathbf{0}\}$ under the Borel measurable mapping $d\colon \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^m$, $(\boldsymbol{u}, \boldsymbol{v}) \mapsto \boldsymbol{u} - \boldsymbol{v}$ is in $\mathscr{B}(\mathbb{R}^m \times \mathbb{R}^m)$. Since $\mathcal{U} = F^{-1}(\mathcal{D})$, we conclude that $\mathcal{U} \in \mathscr{B}(\mathbb{R}^m)$. Now, $\mathrm{P}[\mathsf{x} \in \mathcal{U}] > 0$ as $\mathrm{P}[\mathsf{x} \notin \mathcal{U}] = \mathrm{P}[g(\boldsymbol{A}, \boldsymbol{A}\mathsf{x}) \neq \mathsf{x}] < 1$ by assumption. It remains to show that $\boldsymbol{A}$ is one-to-one on $\mathcal{U}$. To this end, consider arbitrary but fixed $\boldsymbol{u}, \boldsymbol{v} \in \mathcal{U}$ and suppose that $\boldsymbol{A}\boldsymbol{u} = \boldsymbol{A}\boldsymbol{v}$. This implies $g(\boldsymbol{A}, \boldsymbol{A}\boldsymbol{u}) = g(\boldsymbol{A}, \boldsymbol{A}\boldsymbol{v})$. As $\boldsymbol{u}, \boldsymbol{v} \in \mathcal{U}$ by assumption, this is only possible if $\boldsymbol{u} = \boldsymbol{v}$. Thus, $\boldsymbol{A}$ is one-to-one on $\mathcal{U}$, which concludes the proof. □

Appendix E
Proof of Lemma IV.3

Proof of i). The proof is by induction. Suppose that μ is s-analytic for $s \in \mathbb{N}\backslash\{1\}$. We have to show that this implies $(s-1)$-analyticity of μ. Consider $\mathcal{C} \in \mathscr{B}(\mathbb{R}^m)$ with $\mu(\mathcal{C}) > 0$. Then, by Definition IV.1, there exist a set $\mathcal{A} \in \mathscr{B}(\mathbb{R}^s)$ of positive Lebesgue measure and a real analytic mapping $h\colon \mathbb{R}^s \to \mathbb{R}^m$ of s-dimensional Jacobian $Jh \not\equiv 0$ such that $h(\mathcal{A}) \subseteq \mathcal{C}$. By Property ii) of Lemma K.8, we can assume, w.l.o.g., that $h|_\mathcal{A}$ is an embedding. For each $z \in \mathbb{R}$, let

$$\mathcal{A}_z = \{\boldsymbol{v} \in \mathbb{R}^{s-1} : (\boldsymbol{v}\ z)^\mathsf{T} \in \mathcal{A}\}. \tag{285}$$

We now use Fubini's Theorem to show that there exists a $z_0 \in \mathbb{R}$ such that $\lambda^{s-1}(\mathcal{A}_{z_0}) > 0$. Concretely,

$$\begin{aligned} 0 &< \int_{\mathbb{R}^s} \mathbb{1}_\mathcal{A}(\boldsymbol{z}) Jh(\boldsymbol{z})\, \mathrm{d}\lambda^s(\boldsymbol{z}) && (286)\\ &= \int_\mathbb{R} \left(\int_{\mathbb{R}^{s-1}} \mathbb{1}_\mathcal{A}\big((\boldsymbol{v}\ z)^\mathsf{T}\big) Jh\big((\boldsymbol{v}\ z)^\mathsf{T}\big)\, \mathrm{d}\lambda^{s-1}(\boldsymbol{v}) \right) \mathrm{d}\lambda^1(z) && (287)\\ &= \int_\mathbb{R} \left(\int_{\mathbb{R}^{s-1}} \mathbb{1}_{\mathcal{A}_z}(\boldsymbol{v}) Jh\big((\boldsymbol{v}\ z)^\mathsf{T}\big)\, \mathrm{d}\lambda^{s-1}(\boldsymbol{v}) \right) \mathrm{d}\lambda^1(z), && (288) \end{aligned}$$

where (286) follows from Lemma H.4 with $\lambda^s(\mathcal{A}) > 0$ and $Jh(\boldsymbol{z}) > 0$ for all $\boldsymbol{z} \in \mathcal{A}$, and in (287) we applied Theorem H.1. We conclude that there must exist a $z_0 \in \mathbb{R}$ such that

$$\int_{\mathbb{R}^{s-1}} \mathbb{1}_{\mathcal{A}_{z_0}}(\boldsymbol{v}) Jh\big((\boldsymbol{v}\ z_0)^\mathsf{T}\big)\, \mathrm{d}\lambda^{s-1}(\boldsymbol{v}) > 0. \tag{289}$$

Again using Lemma H.4 establishes that $\lambda^{s-1}(\mathcal{A}_{z_0}) > 0$. It remains to show that there exists a real analytic mapping $\tilde{h}\colon \mathbb{R}^{s-1} \to \mathbb{R}^m$ with $\tilde{h}(\mathcal{A}_{z_0}) \subseteq \mathcal{C}$ and of $(s-1)$-dimensional Jacobian $J\tilde{h} \not\equiv 0$. To this end, consider the mapping

$$\begin{aligned} \varphi\colon \mathbb{R}^{s-1} &\to \mathbb{R}^s && (290)\\ \boldsymbol{v} &\mapsto (\boldsymbol{v}\ z_0)^\mathsf{T} && (291) \end{aligned}$$

and set $\tilde{h} = h \circ \varphi\colon \mathbb{R}^{s-1} \to \mathbb{R}^m$. Now, φ is real analytic thanks to Corollary K.1. Thus, $\tilde{h}$ as a composition of real analytic mappings is real analytic by Corollary K.2. Furthermore, since $\varphi(\mathcal{A}_{z_0}) \subseteq \mathcal{A}$ by (285), it follows that $\tilde{h}(\mathcal{A}_{z_0}) \subseteq \mathcal{C}$. It remains to establish that $J\tilde{h} \not\equiv 0$. To this end, we first note that the chain rule implies $D\tilde{h}(\boldsymbol{v}) = Dh(\boldsymbol{v}, z_0) D\varphi(\boldsymbol{v})$. Therefore,

$$\begin{aligned} \operatorname{rank}(D\tilde{h}(\boldsymbol{v})) &= \operatorname{rank}(Dh(\boldsymbol{v}, z_0) D\varphi(\boldsymbol{v})) && (292)\\ &= \operatorname{rank}\big(Dh(\boldsymbol{v}, z_0)(\boldsymbol{I}_{s-1}\ \boldsymbol{0})^\mathsf{T}\big) && (293)\\ &\geq s-1 \quad \text{for all } \boldsymbol{v} \in \mathcal{A}_{z_0}, && (294) \end{aligned}$$

where (294) is by Lemma K.12 upon noting that $Dh(\boldsymbol{v}, z_0) \in \mathbb{R}^{m\times s}$, with $\operatorname{rank}(Dh(\boldsymbol{v}, z_0)) = s$ (recall that $(\boldsymbol{v}\ z_0)^\mathsf{T} \in \mathcal{A}$ for all $\boldsymbol{v} \in \mathcal{A}_{z_0}$ by (285) and that the s-dimensional Jacobian $Jh(\boldsymbol{z}) > 0$ for all $\boldsymbol{z} \in \mathcal{A}$), and $(\boldsymbol{I}_{s-1}\ \boldsymbol{0})^\mathsf{T} \in \mathbb{R}^{s\times(s-1)}$. Since $\operatorname{rank}(D\tilde{h}(\boldsymbol{v})) \geq s-1$ and $D\tilde{h}(\boldsymbol{v}) \in \mathbb{R}^{m\times(s-1)}$, we conclude that $J\tilde{h}(\boldsymbol{v}) > 0$ for all $\boldsymbol{v} \in \mathcal{A}_{z_0}$. Thus, $J\tilde{h} \not\equiv 0$, which finalizes the proof.

Proof of ii). Suppose that μ is s-analytic and consider $\mathcal{C} \in \mathscr{B}(\mathbb{R}^m)$ with $\mu(\mathcal{C}) > 0$. We have to show that $\mathscr{H}^s(\mathcal{C}) > 0$. By Definition IV.1, there exist a set $\mathcal{A} \in \mathscr{B}(\mathbb{R}^s)$ of positive Lebesgue measure and a real analytic mapping $h\colon \mathbb{R}^s \to \mathbb{R}^m$ of s-dimensional Jacobian $Jh \not\equiv 0$ such that $h(\mathcal{A}) \subseteq \mathcal{C}$. By Property ii) of Lemma K.8, we can assume, w.l.o.g., that $h|_\mathcal{A}$ is an embedding. We now use the area formula Corollary H.3 to conclude that $\mathscr{H}^s(\mathcal{C}) > 0$. Concretely,

$$\begin{aligned} 0 &< \int_\mathcal{A} Jh(\boldsymbol{z})\, \mathrm{d}\lambda^s(\boldsymbol{z}) && (295)\\ &= \mathscr{H}^s\big(h(\mathcal{A})\big) && (296)\\ &\leq \mathscr{H}^s(\mathcal{C}), && (297) \end{aligned}$$

where (295) is by Lemma H.4, $\lambda^s(\mathcal{A}) > 0$, and $Jh(\boldsymbol{z}) > 0$ for all $\boldsymbol{z} \in \mathcal{A}$, in (296) we applied the area formula Corollary H.3 upon noting that $h|_\mathcal{A}$ is one-to-one as an embedding and locally Lipschitz by real analyticity of h, and (297) is by monotonicity of $\mathscr{H}^s$ together with $h(\mathcal{A}) \subseteq \mathcal{C}$.

Proof of iii). „$\Rightarrow$": Suppose that μ is s-analytic and there exists a set $\mathcal{U} \subseteq \mathbb{R}^m$ such that $\mu = \mu|_\mathcal{U}$ and $\mathscr{H}^s|_\mathcal{U}$ is σ-finite. Since $\mathscr{H}^s$ is Borel regular (see Definition H.2), we may assume, w.l.o.g., that $\mathcal{U} \in \mathscr{B}(\mathbb{R}^m)$. By σ-finiteness of $\mathscr{H}^s|_\mathcal{U}$, there exist sets $\mathcal{V}_i \in \mathscr{B}(\mathbb{R}^m)$, $i \in \mathbb{N}$, such that

$$\mathcal{U} = \bigcup_{i\in\mathbb{N}} \mathcal{U} \cap \mathcal{V}_i \tag{298}$$

and $\mathscr{H}^s(\mathcal{U} \cap \mathcal{V}_i) < \infty$ for all $i \in \mathbb{N}$. For every $i \in \mathbb{N}$, since $\mathcal{U} \cap \mathcal{V}_i$ as the intersection of two Borel sets is in $\mathscr{B}(\mathbb{R}^m)$ and $\mathscr{H}^s(\mathcal{U} \cap \mathcal{V}_i) < \infty$, we can write [47, p. 83] $\mathcal{U} \cap \mathcal{V}_i = \mathcal{W}_i \cup \widetilde{\mathcal{W}}_i$, where $\mathcal{W}_i$ is countably $(\mathscr{H}^s, s)$-rectifiable and $\widetilde{\mathcal{W}}_i$ is purely $\mathscr{H}^s$-unrectifiable, i.e., $\mathscr{H}^s(\widetilde{\mathcal{W}}_i \cap \mathcal{E}) = 0$ for all countably $(\mathscr{H}^s, s)$-rectifiable sets $\mathcal{E} \subseteq \mathbb{R}^m$. This allows us to decompose $\mathcal{U}$ according to $\mathcal{U} = \mathcal{W} \cup \widetilde{\mathcal{W}}$ with $\mathcal{W} = \bigcup_{i\in\mathbb{N}} \mathcal{W}_i$ and $\widetilde{\mathcal{W}} = \bigcup_{i\in\mathbb{N}} \widetilde{\mathcal{W}}_i$. Since all the $\mathcal{W}_i$ are countably $(\mathscr{H}^s, s)$-rectifiable, so is $\mathcal{W}$; and since all the $\widetilde{\mathcal{W}}_i$ are purely $\mathscr{H}^s$-unrectifiable, so is $\widetilde{\mathcal{W}}$.

As $\mu = \mu|_\mathcal{U}$ by assumption, and $\mathcal{U} = \mathcal{W} \cup \widetilde{\mathcal{W}}$, it remains to show that $\mu(\widetilde{\mathcal{W}}) = 0$ to conclude that $\mu = \mu|_\mathcal{W}$ for the countably $(\mathscr{H}^s, s)$-rectifiable set $\mathcal{W}$. Towards a contradiction, suppose that $\mu(\widetilde{\mathcal{W}}) > 0$. Analyticity of μ then implies that there exist a set $\mathcal{A} \in \mathscr{B}(\mathbb{R}^s)$ of positive Lebesgue measure and a real analytic mapping $h\colon \mathbb{R}^s \to \mathbb{R}^m$ of s-dimensional Jacobian $Jh \not\equiv 0$ such that $h(\mathcal{A}) \subseteq \widetilde{\mathcal{W}}$. By countable subadditivity of λ^s, we can assume, w.l.o.g., that $\mathcal{A}$ is bounded; and by Property ii) in Lemma K.8, we can assume, w.l.o.g., that $h|_\mathcal{A}$ is an embedding. It follows that

$$\begin{aligned} 0 &< \int_\mathcal{A} Jh(\boldsymbol{z})\, \mathrm{d}\lambda^s(\boldsymbol{z}) && (299)\\ &= \mathscr{H}^s\big(h(\mathcal{A})\big), && (300) \end{aligned}$$

where (299) is by Lemma H.4, $\lambda^s(\mathcal{A}) > 0$, and $Jh(\boldsymbol{z}) > 0$ for all $\boldsymbol{z} \in \mathcal{A}$, and in (300) we applied the area formula Corollary H.3 upon noting that $h|_\mathcal{A}$ is one-to-one as an embedding and locally Lipschitz by real analyticity of h. Moreover, as $\mathscr{H}^s$ is Borel regular (see Property ii) in Definition H.2), there must exist a set $\mathcal{C} \in \mathscr{B}(\mathbb{R}^m)$ with $h(\mathcal{A}) \subseteq \mathcal{C}$ and $\mathscr{H}^s(\mathcal{C}\backslash h(\mathcal{A})) = 0$. It follows that $\mathcal{C}$ is countably $(\mathscr{H}^s, s)$-rectifiable as $\mathscr{H}^s(\mathcal{C}\backslash$

$h(\overline{\mathcal{A}}))=0$ and h is Lipschitz on the compact set $\overline{\mathcal{A}}$ by Lemma H.12. But this implies

$$\mathscr{H}^s(\widetilde{\mathcal{W}}\cap\mathcal{C})\geq\mathscr{H}^s(h(\mathcal{A})\cap\mathcal{C}) \tag{301}$$
$$=\mathscr{H}^s(h(\mathcal{A})) \tag{302}$$
$$>0, \tag{303}$$

which is not possible as $\widetilde{\mathcal{W}}$ is purely $\mathscr{H}^s$-unrectifiable, thereby concluding the proof.
„$\Leftarrow$": Suppose that there exists a countably $(\mathscr{H}^s,s)$-rectifiable set $\mathcal{W}\subseteq\mathbb{R}^m$ such that $\mu=\mu|_{\mathcal{W}}$. Since $\mathscr{H}^s$ is Borel regular (see Definition H.2), we may assume, w.l.o.g., that $\mathcal{W}\in\mathscr{B}(\mathbb{R}^m)$. As this $\mathcal{W}$ is countably $(\mathscr{H}^s,s)$-rectifiable, [46, Lemma 15.5] implies that $\mathscr{H}^s|_{\mathcal{W}}$ is σ-finite. □

Appendix F
Proof of Lemma IV.4

Suppose that $\mathbf{x}\in\mathbb{R}^m$ is s-analytic and $\mathbf{u}=f(\mathbf{x})$, where $f\colon\mathbb{R}^m\to\mathbb{R}^k$ is a real analytic immersion, and consider $\mathcal{C}\in\mathscr{B}(\mathbb{R}^k)$ with $\mu_{\mathbf{u}}(\mathcal{C})>0$. We have to show that there exist a set $\mathcal{A}\in\mathscr{B}(\mathbb{R}^s)$ of positive Lebesgue measure and a real analytic mapping $g\colon\mathbb{R}^s\to\mathbb{R}^k$ of s-dimensional Jacobian $Jg\not\equiv0$ such that $g(\mathcal{A})\subseteq\mathcal{C}$. Set $\mathcal{D}=f^{-1}(\mathcal{C})\in\mathscr{B}(\mathbb{R}^m)$. Since $\mu_{\mathbf{x}}(\mathcal{D})=\mu_{\mathbf{u}}(\mathcal{C})>0$ and $\mathbf{x}$ is s-analytic, there exist a set $\mathcal{A}\in\mathscr{B}(\mathbb{R}^s)$ of positive Lebesgue measure and a real analytic mapping $h\colon\mathbb{R}^s\to\mathbb{R}^m$ of s-dimensional Jacobian $Jh\not\equiv0$ such that $h(\mathcal{A})\subseteq\mathcal{D}$. We set $g=f\circ h$. Now, $g(\mathcal{A})=f(h(\mathcal{A}))\subseteq f(\mathcal{D})\subseteq\mathcal{C}$. Furthermore, g as the composition of real analytic mappings is real analytic by Corollary K.2. It remains to show that $Jg\not\equiv0$. To this end, we first note that the chain rule implies $Dg(\boldsymbol{z})=(Df)(h(\boldsymbol{z}))Dh(\boldsymbol{z})$. Since $Jh\not\equiv0$, there exists a $\boldsymbol{z}_0\in\mathbb{R}^s$ such that $Jh(\boldsymbol{z}_0)\neq0$. Thus

$$\operatorname{rank}(Dh(\boldsymbol{z}_0))=s. \tag{304}$$

Now, as f is an immersion, it follows that $k\geq m$ and $Jf>0$. Thus,

$$\operatorname{rank}((Df)(h(\boldsymbol{z}_0)))=m. \tag{305}$$

Applying Lemma K.12 to $(Df)(h(\boldsymbol{z}_0))\in\mathbb{R}^{k\times m}$ and $Dh(\boldsymbol{z}_0)\in\mathbb{R}^{m\times s}$ and using (304) and (305) establishes that $\operatorname{rank}(Dg(\boldsymbol{z}_0))\geq s$, which in turn implies $Jg(\boldsymbol{z}_0)\neq0$ as $Dg(\boldsymbol{z}_0)\in\mathbb{R}^{k\times s}$. □

Appendix G
Proof of Lemma IV.6

Suppose that $\mathbf{a}\in\mathbb{R}^k$ and $\mathbf{b}\in\mathbb{R}^l$ with $\mu_{\mathbf{a}}\times\mu_{\mathbf{b}}\ll\lambda^{k+l}$, set $\mathbf{x}=\mathbf{a}\otimes\mathbf{b}$, and consider $\mathcal{C}\in\mathscr{B}(\mathbb{R}^{kl})$ with $\mu_{\mathbf{x}}(\mathcal{C})>0$. We have to show that there exist a set $\mathcal{A}\in\mathscr{B}(\mathbb{R}^{k+l-1})$ of positive Lebesgue measure and a real analytic mapping $h\colon\mathbb{R}^{k+l-1}\to\mathbb{R}^{kl}$ of $(k+l-1)$-dimensional Jacobian $Jh\not\equiv0$ such that $h(\mathcal{A})\subseteq\mathcal{C}$. Let

$$\mathcal{E}=\{(\boldsymbol{a}^\top\ \boldsymbol{b}^\top)^\top:\boldsymbol{a}\in\mathbb{R}^k,\boldsymbol{b}\in\mathbb{R}^l,\boldsymbol{a}\otimes\boldsymbol{b}\in\mathcal{C}\}. \tag{306}$$

Since $\mathcal{C}\in\mathscr{B}(\mathbb{R}^{kl})$ and $\otimes$ is Borel measurable, $\mathcal{E}$ as the inverse image of $\mathcal{C}$ under $\otimes$ is Borel measurable. Furthermore, as $\mathbf{x}=\mathbf{a}\otimes\mathbf{b}$, it follows that

$$(\mu_{\mathbf{a}}\times\mu_{\mathbf{b}})(\mathcal{E})=\mu_{\mathbf{x}}(\mathcal{C}) \tag{307}$$
$$>0, \tag{308}$$

which implies $\lambda^{k+l}(\mathcal{E})>0$ as $\mu_{\mathbf{a}}\times\mu_{\mathbf{b}}\ll\lambda^{k+l}$ by assumption. Using Corollary H.1, we can write

$$\lambda^{k+l}(\mathcal{E})=\int_{\mathbb{R}}\lambda^{k+l-1}(\mathcal{E}_z)\,\mathrm{d}\lambda^1(z), \tag{309}$$

where, for each $z\in\mathbb{R}$,

$$\mathcal{E}_z=\left\{\left(\boldsymbol{a}^\top\ \boldsymbol{v}^\top\right)^\top:\boldsymbol{a}\in\mathbb{R}^k,\boldsymbol{v}\in\mathbb{R}^{l-1},(\boldsymbol{a}^\top\ \boldsymbol{v}^\top z)^\top\in\mathcal{E}\right\}. \tag{310}$$

As $\lambda^{k+l}(\mathcal{E})>0$ we can conclude that there must exist a $z_0\in\mathbb{R}\setminus\{0\}$ such that $\lambda^{k+l-1}(\mathcal{E}_{z_0})>0$. We set $\mathcal{A}=\mathcal{E}_{z_0}$. Next, we define the mapping

$$h\colon\mathbb{R}^{k+l-1}\to\mathbb{R}^{kl} \tag{311}$$
$$(c_1\dots c_{k+l-1})^\top\mapsto(c_1\dots c_k)^\top\otimes(c_{k+1}\dots c_{k+l-1}\ z_0)^\top, \tag{312}$$

which is real analytic thanks to Corollary K.1, and we write

$$\mathcal{A}=\left\{\left(\boldsymbol{a}^\top\ \boldsymbol{v}^\top\right)^\top:\boldsymbol{a}\in\mathbb{R}^k,\boldsymbol{v}\in\mathbb{R}^{l-1},(\boldsymbol{a}^\top\ \boldsymbol{v}^\top z_0)^\top\in\mathcal{E}\right\} \tag{313}$$
$$=\left\{\left(\boldsymbol{a}^\top\ \boldsymbol{v}^\top\right)^\top:\boldsymbol{a}\in\mathbb{R}^k,\boldsymbol{v}\in\mathbb{R}^{l-1},\boldsymbol{a}\otimes(\boldsymbol{v}^\top\ z_0)^\top\in\mathcal{C}\right\} \tag{314}$$
$$=h^{-1}(\mathcal{C}), \tag{315}$$

where (314) follows from (306), and (315) is by (311)–(312). By construction, $h(\mathcal{A})\subseteq\mathcal{C}$. Furthermore, $\mathcal{A}$ as the inverse image of $\mathcal{C}\in\mathscr{B}(\mathbb{R}^{kl})$ under a real analytic and, therefore, Borel measurable mapping is in $\mathscr{B}(\mathbb{R}^{k+l-1})$. It remains to show that there exist an $\boldsymbol{a}_0\in\mathbb{R}^k$ and a $\boldsymbol{v}_0\in\mathbb{R}^{l-1}$ such that

$$Jh\Big(\big(\boldsymbol{a}_0^\top\ \boldsymbol{v}_0^\top\big)^\top\Big) \tag{316}$$
$$=\sqrt{\det\left(\left(Dh\Big(\big(\boldsymbol{a}_0^\top\ \boldsymbol{v}_0^\top\big)^\top\Big)\right)^\top Dh\Big(\big(\boldsymbol{a}_0^\top\ \boldsymbol{v}_0^\top\big)^\top\Big)\right)} \tag{317}$$
$$>0. \tag{318}$$

This will be accomplished by showing that there exist an $\boldsymbol{a}_0\in\mathbb{R}^k$ and a $\boldsymbol{v}_0\in\mathbb{R}^{l-1}$ such that

$$\operatorname{rank}\Big(Dh\Big(\big(\boldsymbol{a}_0^\top\ \boldsymbol{v}_0^\top\big)^\top\Big)\Big)=k+l-1. \tag{319}$$

Now,

$$Dh\Big(\big(\boldsymbol{a}^\top\ \boldsymbol{v}^\top\big)^\top\Big)=\begin{pmatrix}\boldsymbol{v}&\mathbf{0}&\dots&\mathbf{0}&\mathbf{0}&a_1\boldsymbol{I}_{l-1}\\ z_0&0&\dots&0&0&\mathbf{0}\\ \mathbf{0}&\boldsymbol{v}&\dots&\mathbf{0}&\mathbf{0}&a_2\boldsymbol{I}_{l-1}\\ 0&z_0&\dots&0&0&\mathbf{0}\\ \vdots&\vdots&\vdots&\vdots&\vdots&\vdots\\ \mathbf{0}&\mathbf{0}&\dots&\boldsymbol{v}&\mathbf{0}&a_{k-1}\boldsymbol{I}_{l-1}\\ 0&0&\dots&z_0&0&\mathbf{0}\\ \mathbf{0}&\mathbf{0}&\dots&\mathbf{0}&\boldsymbol{v}&a_k\boldsymbol{I}_{l-1}\\ 0&0&\dots&0&z_0&\mathbf{0}\end{pmatrix} \tag{320}$$

for general $\boldsymbol{a} \in \mathbb{R}^k$ and $\boldsymbol{v} \in \mathbb{R}^{l-1}$, and

$$Dh\Big(\big(\boldsymbol{a}_0^{\mathsf{T}}\ \boldsymbol{v}_0^{\mathsf{T}}\big)^{\mathsf{T}}\Big) = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \boldsymbol{I}_{l-1} \\ z_0 & 0 & \dots & 0 & 0 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \boldsymbol{I}_{l-1} \\ 0 & z_0 & \dots & 0 & 0 & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \boldsymbol{I}_{l-1} \\ 0 & 0 & \dots & z_0 & 0 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \boldsymbol{I}_{l-1} \\ 0 & 0 & \dots & 0 & z_0 & \mathbf{0} \end{pmatrix} \tag{321}$$

for the specific choices $\boldsymbol{a}_0 = (1 \dots 1)^{\mathsf{T}} \in \mathbb{R}^k$ and $\boldsymbol{v}_0 = (0 \dots 0)^{\mathsf{T}} \in \mathbb{R}^{l-1}$. Since the $(kl) \times (k+l-1)$ matrix $Dh\big((\boldsymbol{a}_0^{\mathsf{T}}\ \boldsymbol{v}_0^{\mathsf{T}})^{\mathsf{T}}\big)$ has the regular $(k+l-1) \times (k+l-1)$ submatrix (recall that $\boldsymbol{z}_0 \neq 0$)

$$\begin{pmatrix} \mathbf{0} & \boldsymbol{I}_{l-1} \\ z_0 \boldsymbol{I}_k & \mathbf{0} \end{pmatrix}, \tag{322}$$

(319) indeed holds, which concludes the proof. □

APPENDIX H
TOOLS FROM (GEOMETRIC) MEASURE THEORY

In this appendix, we state some basic definitions and results from measure theory and geometric measure theory used throughout the paper. For an excellent in-depth treatment of geometric measure theory, the interested reader is referred to [36], [40], [48], [49].

A. Preliminaries from Measure Theory

Definition H.1. [36, Definition 1.2.1] A measure (sometimes called outer measure, see [36, Remark 1.2.6]) on a nonempty set $\mathcal{X}$ is a nonnegative function μ defined on all subsets of $\mathcal{X}$ with the following properties:

i) $\mu(\emptyset) = 0$.
ii) Monotonicity:

$$\mu(\mathcal{A}) \leq \mu(\mathcal{B}) \quad \text{for all } \mathcal{A} \subseteq \mathcal{B} \subseteq \mathcal{X}. \tag{323}$$

iii) Countable subadditivity:

$$\mu\left(\bigcup_{i\in\mathbb{N}} \mathcal{A}_i\right) \leq \sum_{i\in\mathbb{N}} \mu(\mathcal{A}_i) \tag{324}$$

for all sequences $\{\mathcal{A}_i\}_{i\in\mathbb{N}}$ of sets $\mathcal{A}_i \subseteq \mathcal{X}$.

A set $\mathcal{A} \subseteq \mathcal{X}$ is μ-measurable if it satisfies

$$\mu(\mathcal{E}) = \mu(\mathcal{E} \cap \mathcal{A}) + \mu(\mathcal{E} \backslash \mathcal{A}) \quad \text{for all } \mathcal{E} \subseteq \mathcal{X}. \tag{325}$$

For a probability measure $\mu_{\mathbf{x}}$, countable subadditivity is equivalent to the union bound

$$\mathrm{P}\left[\mathbf{x} \in \bigcup_{i\in\mathbb{N}} \mathcal{A}_i\right] \leq \sum_{i\in\mathbb{N}} \mathrm{P}[\mathbf{x} \in \mathcal{A}_i]. \tag{326}$$

The collection of μ-measurable sets forms a σ-algebra [36, Theorem 1.2.4], which we denote by $\mathscr{S}_\mu(\mathcal{X})$. If $\mathcal{X}$ is endowed with a topology,[3] the smallest σ-algebra containing the open sets is the Borel σ-algebra $\mathscr{B}(\mathcal{X})$. A measurable space $(\mathcal{X}, \mathscr{S}(\mathcal{X}))$ is a set $\mathcal{X}$ equipped with a σ-algebra $\mathscr{S}(\mathcal{X})$. A measure space $(\mathcal{X}, \mathscr{S}(\mathcal{X}), \mu)$ is a set $\mathcal{X}$ with a measure μ and a σ-algebra $\mathscr{S}(\mathcal{X}) \subseteq \mathscr{S}_\mu(\mathcal{X})$.

[3] A topology for $\mathcal{X}$ is a collection of subsets of $\mathcal{X}$ that contains $\emptyset$ and $\mathcal{X}$ and is closed under finite intersections and arbitrary unions. The members of a topology are called open sets.

Lemma H.1. [36, Theorem 1.2.5] If $(\mathcal{X}, \mathscr{S}(\mathcal{X}), \mu)$ is a measure space and $\{\mathcal{A}_n\}_{n\in\mathbb{N}}$ is a sequence of sets $\mathcal{A}_n \in \mathscr{S}(\mathcal{X})$, then the following properties hold.

i) If the sets $\mathcal{A}_n$, $n \in \mathbb{N}$, are pairwise disjoint, then μ is countably additive on their union. That is,

$$\mu\Big(\bigcup_{n\in\mathbb{N}} \mathcal{A}_n\Big) = \sum_{n\in\mathbb{N}} \mu(\mathcal{A}_n). \tag{327}$$

ii) If $\mathcal{A}_n \subseteq \mathcal{A}_{n+1}$ for all $n \in \mathbb{N}$, then

$$\mu\Big(\bigcup_{n\in\mathbb{N}} \mathcal{A}_n\Big) = \lim_{n\to\infty} \mu(\mathcal{A}_n). \tag{328}$$

iii) If $\mathcal{A}_{n+1} \subseteq \mathcal{A}_n$ for all $n \in \mathbb{N}$ and $\mu(\mathcal{A}_1) < \infty$, then

$$\mu\Big(\bigcap_{n\in\mathbb{N}} \mathcal{A}_n\Big) = \lim_{n\to\infty} \mu(\mathcal{A}_n). \tag{329}$$

Definition H.2. [36, Definition 1.2.10] A measure μ on a nonempty set $\mathcal{X}$ endowed with a topology is

i) Borel if all open sets are μ-measurable;
ii) Borel regular if it is Borel and, for each $\mathcal{A} \subseteq \mathcal{X}$, there exists a set $\mathcal{B} \in \mathscr{B}(\mathcal{X})$ such that $\mathcal{A} \subseteq \mathcal{B}$ and $\mu(\mathcal{A}) = \mu(\mathcal{B})$.

Lemma H.2. Let $s \in \mathbb{N}$ and consider a measure μ on $\mathbb{R}^m$. For every nonempty set $\mathcal{U} \subseteq \mathbb{R}^m$, the following statements are equivalent:

i) There exist compact sets $\mathcal{A}_i \subseteq \mathbb{R}^s$ and Lipschitz mappings $\varphi_i \colon \mathcal{A}_i \to \mathbb{R}^m$, $i \in \mathbb{N}$, such that

$$\mu\left(\mathcal{U} \setminus \bigcup_{i\in\mathbb{N}} \varphi_i(\mathcal{A}_i)\right) = 0. \tag{330}$$

ii) There exist bounded sets $\mathcal{A}_i \subseteq \mathbb{R}^s$ and Lipschitz mappings $\varphi_i \colon \mathcal{A}_i \to \mathbb{R}^m$, $i \in \mathbb{N}$, such that

$$\mu\left(\mathcal{U} \setminus \bigcup_{i\in\mathbb{N}} \varphi_i(\mathcal{A}_i)\right) = 0. \tag{331}$$

iii) There exist Lipschitz mappings $\varphi_i \colon \mathbb{R}^s \to \mathbb{R}^m$, $i \in \mathbb{N}$, such that

$$\mu\left(\mathcal{U} \setminus \bigcup_{i\in\mathbb{N}} \varphi_i(\mathbb{R}^s)\right) = 0. \tag{332}$$

Proof. We show that i) ⇒ ii) ⇒ iii) ⇒ i). i) ⇒ ii) is trivial since every compact set $\mathcal{A}_i \subseteq \mathbb{R}^s$ is bounded. ii) ⇒ iii) follows from the fact that, by [46, Theorem 7.2], for every bounded set $\mathcal{A}_i \subseteq \mathbb{R}^s$ and corresponding Lipschitz mapping $\varphi_i \colon \mathcal{A}_i \to \mathbb{R}^m$, there exists a Lipschitz mapping $\tilde{\varphi}_i \colon \mathbb{R}^s \to \mathbb{R}^m$ such that $\tilde{\varphi}_i|_{\mathcal{A}_i} = \varphi_i$. Finally, iii) ⇒ i) as

$$\varphi_i(\mathbb{R}^s) = \bigcup_{j\in\mathbb{N}} \varphi_i(\overline{\mathcal{B}_s(\mathbf{0}, j)}) \quad \text{for all } i \in \mathbb{N} \tag{333}$$

with $\overline{\mathcal{B}_s(\mathbf{0}, j)}$ compact for all $j \in \mathbb{N}$. □

Definition H.3. (Hausdorff measures) [47, Definition 2.46] Let $d \in [0, \infty)$ and $\mathcal{U} \subseteq \mathbb{R}^m$. The d-dimensional Hausdorff measure of $\mathcal{U}$, denoted by $\mathscr{H}^d(\mathcal{U})$, is defined according to

$$\mathscr{H}^d(\mathcal{U}) = \lim_{\delta \to 0} \mathscr{H}_\delta^d(\mathcal{U}), \tag{334}$$

where

$$\mathscr{H}_\delta^d(\mathcal{U}) \tag{335}$$

$$= \frac{\pi^{d/2}}{2^d \Gamma(d/2+1)} \inf \Bigg\{ \sum_{i \in \mathbb{N}} \operatorname{diam}(\mathcal{U}_i)^d : \operatorname{diam}(\mathcal{U}_i) < \delta,$$

$$\mathcal{U} \subseteq \bigcup_{i \in \mathbb{N}} \mathcal{U}_i \Bigg\} \tag{336}$$

for all $\delta > 0$, and the diameter $\operatorname{diam}(\cdot)$ of an arbitrary set $\mathcal{U} \subseteq \mathbb{R}^m$ is defined according to

$$\operatorname{diam}(\mathcal{U}) = \begin{cases} \sup\{\|\boldsymbol{u} - \boldsymbol{v}\|_2 : \boldsymbol{u}, \boldsymbol{v} \in \mathcal{U}\} & \text{if } \mathcal{U} \neq \emptyset \\ 0 & \text{else} \end{cases} \tag{337}$$

with $\Gamma(\cdot)$ denoting the Gamma function.

Lemma H.3. (Main properties of Hausdorff measures)

i) If $a > b \geq 0$, then $\mathscr{H}^a(\mathcal{E}) > 0$ implies $\mathscr{H}^b(\mathcal{E}) = \infty$ for all $\mathcal{E} \subseteq \mathbb{R}^m$ (see Fig. 3).
ii) If $f \colon \mathbb{R}^m \to \mathbb{R}^n$ is Lipschitz with Lipschitz constant L, then

$$\mathscr{H}^d(f(\mathcal{E})) \leq L^d \mathscr{H}^d(\mathcal{E}) \quad \text{for all } \mathcal{E} \subseteq \mathbb{R}^m. \tag{338}$$

iii) $\mathscr{H}^m(\mathcal{E}) = \lambda^m(\mathcal{E})$ for all $\mathcal{E} \in \mathscr{B}(\mathbb{R}^m)$.
iv) $\mathscr{H}^0$ is the counting measure.

Proof. See [47, Proposition 2.49] and [47, Theorem 2.53] for Properties i)–iii). Property iv) follows immediately from the definition of the 0-dimensional Hausdorff measure. □

Definition H.4. (Hausdorff dimension) [47, Definition 2.51] The Hausdorff dimension of $\mathcal{U} \subseteq \mathbb{R}^m$, denoted by $\dim_{\mathrm{H}}(\mathcal{U})$, is defined according to

$$\dim_{\mathrm{H}}(\mathcal{U}) := \sup\{d \geq 0 : \mathscr{H}^d(\mathcal{U}) = \infty\} \tag{339}$$

$$= \inf\{d \geq 0 : \mathscr{H}^d(\mathcal{U}) = 0\}, \tag{340}$$

i.e., $\dim_{\mathrm{H}}(\mathcal{U})$ is the value of d for which the sharp transition from ∞ to 0 in Figure 3 occurs. Depending on the set $\mathcal{U}$, for $d = \dim_{\mathrm{H}}(\mathcal{U})$, $\mathscr{H}^d(\mathcal{U})$ can take on any value in $[0, \infty]$.

As a consequence of Carathéodory's criterion [36, Theorem 1.2.13], Lebesgue and Hausdorff measures are Borel regular.

Definition H.5. (Measurable mapping) [50, Chapter 2]

i) Let $(\mathcal{X}, \mathscr{S}(\mathcal{X}))$ and $(\mathcal{Y}, \mathscr{S}(\mathcal{Y}))$ be measurable spaces. A mapping $f \colon \mathcal{D} \to \mathcal{Y}$, with $\mathcal{D} \in \mathscr{S}(\mathcal{X})$, is $(\mathscr{S}(\mathcal{X}), \mathscr{S}(\mathcal{Y}))$-measurable if $f^{-1}(\mathcal{A}) \in \mathscr{S}(\mathcal{X})$ for all $\mathcal{A} \in \mathscr{S}(\mathcal{Y})$.
ii) Let $(\mathcal{X}, \mathscr{S}(\mathcal{X}))$ be a measurable space and $\mathcal{Y}$ endowed with a topology. A mapping $f \colon \mathcal{D} \to \mathcal{Y}$, with $\mathcal{D} \in \mathscr{S}(\mathcal{X})$, is $\mathscr{S}(\mathcal{X})$-measurable if $f^{-1}(\mathcal{A}) \in \mathscr{S}(\mathcal{X})$ for all $\mathcal{A} \in \mathscr{B}(\mathcal{Y})$.

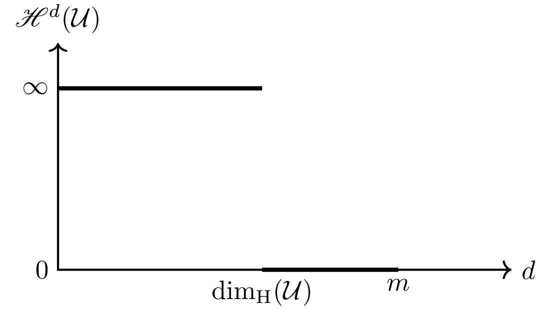

Figure 3. ([38, Figure 3.3]) Graph of $\mathscr{H}^d(\mathcal{U})$ as a function of $d \in [0, m]$ for a set $\mathcal{U} \subseteq \mathbb{R}^m$.

iii) Let $\mathcal{X}$ and $\mathcal{Y}$ both be endowed with a topology. A mapping $f \colon \mathcal{D} \to \mathcal{Y}$, with $\mathcal{D} \in \mathscr{B}(\mathcal{X})$, is Borel measurable if $f^{-1}(\mathcal{A}) \in \mathscr{B}(\mathcal{X})$ for all $\mathcal{A} \in \mathscr{B}(\mathcal{Y})$.

Lemma H.4. [50, Corollary 4.10] Let $(\mathcal{X}, \mathscr{S}(\mathcal{X}), \mu)$ be a measure space and consider a nonnegative measurable function $f \colon \mathcal{X} \to \mathbb{R}$. Then, $f(x) = 0$ μ-almost everywhere, i.e., $\mu\{x \in \mathcal{X} : f(x) \neq 0\} = 0$, if and only if

$$\int_{\mathcal{X}} f(x) \, \mathrm{d}\mu(x) = 0. \tag{341}$$

Definition H.6. [36, Definition 1.3.25] The measure space $(\mathcal{X}, \mathscr{S}(\mathcal{X}), \mu)$ is σ-finite if $\mathcal{X} = \bigcup_{i \in \mathbb{N}} \mathcal{A}_i$, with $\mathcal{A}_i \in \mathscr{S}(\mathcal{X})$ and $\mu(\mathcal{A}_i) < \infty$ for all $i \in \mathbb{N}$. The measure space is finite if $\mu(\mathcal{X}) < \infty$. A Borel measure μ on a topological space $\mathcal{X}$ is *σ-finite* (respectively *finite*) if $(\mathcal{X}, \mathscr{B}(\mathcal{X}), \mu)$ is a σ-finite (respectively finite) measure space.

For example, all probability measures are finite and the Lebesgue measure is σ-finite. However, the s-dimensional Hausdorff measure with $s < m$ is not σ-finite.

Theorem H.1. (Fubini's theorem) [50, Theorem 10.9] Let $(\mathcal{X}, \mathscr{S}(\mathcal{X}), \mu)$ and $(\mathcal{Y}, \mathscr{S}(\mathcal{Y}), \nu)$ be σ-finite measure spaces and suppose that $f \colon \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ is nonnegative and $(\mathscr{S}(\mathcal{X}) \otimes \mathscr{S}(\mathcal{Y}))$-measurable. Then,

$$\varphi(x) = \int_{\mathcal{Y}} f(x, y) \, \mathrm{d}\nu(y) \quad \text{is } \mathscr{S}(\mathcal{X})\text{-measurable}, \tag{342}$$

$$\psi(y) = \int_{\mathcal{X}} f(x, y) \, \mathrm{d}\mu(x) \quad \text{is } \mathscr{S}(\mathcal{Y})\text{-measurable}, \tag{343}$$

and

$$\int_{\mathcal{X}} \varphi(x) \, \mathrm{d}\mu(x) = \int_{\mathcal{X} \times \mathcal{Y}} f(x, y) \, \mathrm{d}(\mu \times \nu)(x \times y) \tag{344}$$

$$= \int_{\mathcal{Y}} \psi(y) \, \mathrm{d}\nu(y). \tag{345}$$

Corollary H.1. Let $(\mathcal{X}, \mathscr{S}(\mathcal{X}), \mu)$ and $(\mathcal{Y}, \mathscr{S}(\mathcal{Y}), \nu)$ be σ-finite measure spaces and suppose that $\mathcal{A} \in \mathscr{S}(\mathcal{X}) \otimes \mathscr{S}(\mathcal{Y})$. Then,

$$\int_{\mathcal{X}} \nu(\{y : (x, y) \in \mathcal{A}\}) \, \mathrm{d}\mu(x) \tag{346}$$

$$= \int_{\mathcal{X} \times \mathcal{Y}} \mathbb{1}_{\mathcal{A}}(x, y) \, \mathrm{d}(\mu \times \nu)(x \times y) \tag{347}$$

$$= \int_{\mathcal{Y}} \mu(\{x : (x, y) \in \mathcal{A}\}) \, \mathrm{d}\nu(y). \tag{348}$$

Proof. Follows from Theorem H.1 with $f(x,y) = \mathbb{1}_{\mathcal{A}}(x,y)$, noting that

$$\int_{\mathcal{Y}} \mathbb{1}_{\mathcal{A}}(x,y)\,\mathrm{d}\nu(y) = \nu(\{y : (x,y) \in \mathcal{A}\}), \tag{349}$$

$$\int_{\mathcal{X}} \mathbb{1}_{\mathcal{A}}(x,y)\,\mathrm{d}\mu(x) = \mu(\{x : (x,y) \in \mathcal{A}\}). \tag{350}$$

□

Lemma H.5. [51, Exercise 1.7.19] If $\mathcal{X}$ and $\mathcal{Y}$ are Euclidean spaces, then $\mathscr{B}(\mathcal{X} \times \mathcal{Y}) = \mathscr{B}(\mathcal{X}) \otimes \mathscr{B}(\mathcal{Y})$.

Lemma H.6. [50, Corollary 2.10] Let $(\mathcal{X}, \mathscr{S}(\mathcal{X}))$ be a measurable space and suppose that $\{f_i\}_{i\in\mathbb{N}}$ is a sequence of $\mathscr{S}(\mathcal{X})$-measurable functions $f_i \colon \mathcal{X} \to \mathbb{R}$ converging pointwise to $f \colon \mathcal{X} \to \mathbb{R}$. Then, f is $\mathscr{S}(\mathcal{X})$-measurable.

Lemma H.7. Let $\mathcal{X}$ and $\mathcal{Y}$ be Euclidean spaces and suppose that $f_i \colon \mathcal{X} \to \mathcal{Y}$, $i = 1, \dots, n$, are Borel measurable mappings. Then,

$$f \colon \mathcal{X} \to \tilde{\mathcal{Y}} = \underbrace{\mathcal{Y} \times \cdots \times \mathcal{Y}}_{n \text{ times}} \tag{351}$$

$$x \mapsto (f_1(x), \dots, f_n(x)) \tag{352}$$

is Borel measurable.

Proof. Lemma H.5 implies $\mathscr{B}(\tilde{\mathcal{Y}}) = \mathscr{B}(\mathcal{Y}) \otimes \cdots \otimes \mathscr{B}(\mathcal{Y})$. Thus, $\mathscr{B}(\tilde{\mathcal{Y}})$ is generated by sets $\mathcal{U}_1 \times \cdots \times \mathcal{U}_n$, where $\mathcal{U}_i \in \mathscr{B}(\mathcal{Y})$ for $i = 1, \dots, n$. It is therefore sufficient to show that $f^{-1}(\mathcal{U}_1 \times \cdots \times \mathcal{U}_n) \in \mathscr{B}(\mathcal{X})$ for all $\mathcal{U}_i \in \mathscr{B}(\mathcal{X})$, $i = 1, \dots, n$. But $f^{-1}(\mathcal{U}_1 \times \cdots \times \mathcal{U}_n) = f_1^{-1}(\mathcal{U}_1) \cap \cdots \cap f_n^{-1}(\mathcal{U}_n) \in \mathscr{B}(\mathcal{X})$ for all $\mathcal{U}_i \in \mathscr{B}(\mathcal{Y})$, $i = 1, \dots, n$, as all the f_i are Borel measurable and every σ-algebra is closed under finite or countably infinite intersections. □

Corollary H.2. Let $\mathcal{X}$ be a Euclidean space and suppose that $\{f_i\}_{i\in\mathbb{N}}$ is a sequence of Borel measurable mappings $f_i \colon \mathcal{X} \to \mathbb{R}^n$ converging pointwise to $f \colon \mathcal{X} \to \mathbb{R}^n$. Then, f is Borel measurable.

Proof. We can write $f(x) = (f^{(1)}(x), \dots, f^{(n)}(x))^{\mathsf{T}}$, where $f^{(j)} \colon \mathcal{X} \to \mathbb{R}$ is Borel measurable for $j = 1, \dots, n$, and $f_i(x) = \big(f_i^{(1)}(x), \dots, f_i^{(n)}(x)\big)^{\mathsf{T}}$, where $f_i^{(j)} \colon \mathcal{X} \to \mathbb{R}$ is Borel measurable for $j = 1, \dots, n$ and all $i \in \mathbb{N}$. Then, for each $j = 1, \dots, n$, the sequence $(f_i^{(j)})_{i\in\mathbb{N}}$ of Borel measurable functions converges pointwise to $f^{(j)}$, which is Borel measurable thanks to Lemma H.6. Finally, Lemma H.7 implies that f is Borel measurable as its individual components $f^{(1)}, \dots, f^{(n)}$ are Borel measurable. □

Lemma H.8. Let $\mathcal{X}$ and $\mathcal{Y}$ be topological spaces, $\mathcal{C} \in \mathscr{B}(\mathcal{X})$, $f \colon \mathcal{C} \to \mathcal{Y}$ Borel measurable, and $y_0 \in \mathcal{Y} \setminus f(\mathcal{C})$. Consider $h \colon \mathcal{X} \to \mathcal{Y}$ with $h|_{\mathcal{C}} = f$ and $h|_{\mathcal{X}\setminus\mathcal{C}} = y_0$. Then, h is Borel measurable.

Proof. We have to show that $h^{-1}(\mathcal{A}) \in \mathscr{B}(\mathcal{X})$ for all $\mathcal{A} \in \mathscr{B}(\mathcal{Y})$. Consider an arbitrary but fixed $\mathcal{A} \in \mathscr{B}(\mathcal{Y})$ and suppose first that $y_0 \in \mathcal{A}$. Now, $\mathcal{A} \in \mathscr{B}(\mathcal{Y})$ and $\{y_0\} \in \mathscr{B}(\mathcal{Y})$ imply $\mathcal{A}\setminus\{y_0\} \in \mathscr{B}(\mathcal{Y})$. Therefore,

$$h^{-1}(\mathcal{A}) = h^{-1}(\mathcal{A}\setminus\{y_0\}) \cup h^{-1}(\{y_0\}) \tag{353}$$

$$= f^{-1}(\mathcal{A}\setminus\{y_0\}) \cup (\mathcal{X}\setminus\mathcal{C}) \in \mathscr{B}(\mathcal{X}) \tag{354}$$

because f is Borel measurable, $\mathcal{A}\setminus\{y_0\} \in \mathscr{B}(\mathcal{Y})$, and $\mathcal{X}\setminus\mathcal{C} \in \mathscr{B}(\mathcal{X})$. If $y_0 \notin \mathcal{A}$, then

$$h^{-1}(\mathcal{A}) = f^{-1}(\mathcal{A}) \in \mathscr{B}(\mathcal{X}) \tag{355}$$

as f is Borel measurable and $\mathcal{A} \in \mathscr{B}(\mathcal{Y})$. □

Lemma H.9. Let μ be a measure on $\mathbb{R}^k$ and consider $\mathcal{A} \subseteq \mathbb{R}^k$ with $\mu(\mathcal{A}) > 0$. Then, there exists a $\boldsymbol{z}_0 \in \mathcal{A}$ such that

$$\mu\big(\mathcal{B}_k(\boldsymbol{z}_0, r) \cap \mathcal{A}\big) > 0 \quad \text{for all } r > 0. \tag{356}$$

Proof. Suppose, to the contrary, that such a $\boldsymbol{z}_0$ does not exist. Then, for every $\boldsymbol{z} \in \mathcal{A}$, there must exist a $r_{\boldsymbol{z}} > 0$ such that $\mu\big(\mathcal{B}_k(\boldsymbol{z}, r_{\boldsymbol{z}}) \cap \mathcal{A}\big) = 0$. With these $r_{\boldsymbol{z}}$, we write

$$\mathcal{A} = \bigcup_{\boldsymbol{z}\in\mathcal{A}} (\mathcal{B}_k(\boldsymbol{z}, r_{\boldsymbol{z}}) \cap \mathcal{A}). \tag{357}$$

It now follows from the Lindelöf property of $\mathbb{R}^k$ [45, Definition 5.6.19][4] that there must exist a countable subset $\{\boldsymbol{z}_i : i \in \mathbb{N}\} \subseteq \mathcal{A}$ such that

$$\mathcal{A} = \bigcup_{i\in\mathbb{N}} (\mathcal{B}_k(\boldsymbol{z}_i, r_{\boldsymbol{z}_i}) \cap \mathcal{A}). \tag{358}$$

Since $\mu\big(\mathcal{B}_k(\boldsymbol{z}_i, r_{\boldsymbol{z}_i}) \cap \mathcal{A}\big) = 0$ for all $i \in \mathbb{N}$, the countable subadditivity of μ implies

$$\mu(\mathcal{A}) \le \sum_{i\in\mathbb{N}} \mu\big(\mathcal{B}_k(\boldsymbol{z}_i, r_{\boldsymbol{z}_i}) \cap \mathcal{A}\big) \tag{359}$$

$$= 0, \tag{360}$$

which contradicts the assumption $\mu(\mathcal{A}) > 0$. □

Lemma H.10. Let $(\mathcal{X}, \mathscr{B}(\mathcal{X}), \mu)$ be a σ-finite measure space and $\mathcal{B} \in \mathscr{B}(\mathcal{X})$. Then,

$$\mathcal{B} = \mathcal{N} \cup \mathcal{A}, \tag{361}$$

where $\mu(\mathcal{N}) = 0$ and $\mathcal{A} = \bigcup_{i\in\mathbb{N}} \mathcal{A}_i$ with $\mathcal{A}_i \subseteq \mathbb{R}^m$ compact for all $i \in \mathbb{N}$.

Proof. By [47, Proposition 1.43], we can find, for each $i \in \mathbb{N}$, a compact set $\mathcal{K}_i$ such that $\mathcal{K}_i \subseteq \mathcal{B}$ and $\mu(\mathcal{B}\setminus\mathcal{K}_i) \le 1/i$. For $j \in \mathbb{N}$, let $\mathcal{A}_j = \bigcup_{i=1}^{j} \mathcal{K}_i$. It follows that $\{\mathcal{A}_j\}_{j\in\mathbb{N}}$ is an increasing (in terms of $\subseteq$) sequence of compact sets $\mathcal{A}_j \subseteq \mathcal{B}$ satisfying $\mu(\mathcal{B}) - \mu(\mathcal{A}_j) = \mu(\mathcal{B}\setminus\mathcal{A}_j) \le 1/j$ for all $j \in \mathbb{N}$. We set $\mathcal{A} = \bigcup_{j\in\mathbb{N}} \mathcal{A}_j \subseteq \mathcal{B}$. Thus, $\mu(\mathcal{A}) \le \mu(\mathcal{B})$ by monotonicity of μ. Now,

$$\mu(\mathcal{A}) = \lim_{j\to\infty} \mu(\mathcal{A}_j) \tag{362}$$

$$= \mu(\mathcal{B}) - \lim_{j\to\infty} \mu(\mathcal{B}\setminus\mathcal{A}_j) \tag{363}$$

$$\ge \mu(\mathcal{B}) - \lim_{j\to\infty} \frac{1}{j} \tag{364}$$

$$= \mu(\mathcal{B}), \tag{365}$$

where (362) follows from Property ii) in Lemma H.1. We conclude that $\mu(\mathcal{A}) = \mu(\mathcal{B})$, which, together with $\mathcal{A} \subseteq \mathcal{B}$, yields (361). □

[4]Recall that $\mathbb{R}^k$ with the Euclidean distance metric is a separable metric space, i.e., $\mathbb{R}^k$ includes a countable dense subset, and it is therefore a Lindelöf space by [45, Proposition 5.6.22].

Lemma H.11. Let $x_i > 0$ for all $i \in \mathcal{I}$ and set

$$M = \sup_{\mathcal{J}\subseteq\mathcal{I}\,:\,|\mathcal{J}|<\infty} \sum_{i\in\mathcal{J}} x_i. \tag{366}$$

Suppose that $M < \infty$. Then, $\mathcal{I}$ is finite or countably infinite.

Proof. For every $k \in \mathbb{N}_0$, set $I_k = \{i \in \mathcal{I} : 1/(k+1) \le x_i < 1/k\}$ and

$$M_k = \sup_{\mathcal{J}\subseteq\mathcal{I}_k:|\mathcal{J}|<\infty} \sum_{i\in\mathcal{J}} x_i. \tag{367}$$

Since $M < \infty$ and $M_k \le M$, we must have $M_k < \infty$ for all $k \in \mathbb{N}_0$. But for every $k \in \mathbb{N}_0$, by the definition of $\mathcal{I}_k$, M_k can only be finite if $|\mathcal{I}_k| < \infty$. Thus, $|\mathcal{I}_k| < \infty$ for all $k \in \mathbb{N}_0$. Since $\mathcal{I} = \bigcup_{k\in\mathbb{N}} \mathcal{I}_k$, we conclude that $\mathcal{I}$ is finite or countably infinite. □

B. Properties of Locally Lipschitz and Differentiable Mappings

Definition H.7. (Locally Lipschitz mapping) [52, Definition 3.118]

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

$$\|f(\boldsymbol{u}) - f(\boldsymbol{v})\|_2 \le L\|\boldsymbol{u} - \boldsymbol{v}\|_2 \quad \text{for all } \boldsymbol{u}, \boldsymbol{v} \in \mathcal{U}. \tag{368}$$

The smallest constant L for (368) to hold is the Lipschitz constant of f.

ii) A mapping $f\colon \mathbb{R}^k \to \mathbb{R}^l$ is locally Lipschitz if every $\boldsymbol{x} \in \mathbb{R}^k$ admits an open neighborhood $\mathcal{U}_{\boldsymbol{x}} \subseteq \mathbb{R}^k$ containing $\boldsymbol{x}$ such that $f|_{\mathcal{U}_{\boldsymbol{x}}}$ is Lipschitz.

The following result establishes a necessary and sufficient condition for a mapping to be locally Lipschitz. Since we could not find a proof for this statement in the literature, we present one here for completeness.

Lemma H.12. The mapping $f\colon \mathbb{R}^k \to \mathbb{R}^l$ is locally Lipschitz if and only if $f|_{\mathcal{K}}$ is Lipschitz for all compact sets $\mathcal{K} \subseteq \mathbb{R}^k$.

Proof. „⇒": Suppose that $f\colon \mathbb{R}^k \to \mathbb{R}^l$ is locally Lipschitz and consider a compact set $\mathcal{K} \subseteq \mathbb{R}^k$. We have to show that $f|_{\mathcal{K}}$ is Lipschitz. For every $\boldsymbol{x} \in \mathcal{K}$, by the local Lipschitz property of f, there exists an open neighborhood $\mathcal{U}_{\boldsymbol{x}}$ containing $\boldsymbol{x}$ such that $f|_{\mathcal{U}_{\boldsymbol{x}}}$ is Lipschitz. Since $\mathcal{K} \subseteq \bigcup_{\boldsymbol{x}\in\mathcal{K}} \mathcal{U}_{\boldsymbol{x}}$ is a cover of the compact set $\mathcal{K}$ by open sets, there must exist $\mathcal{U}_{\boldsymbol{x}_i}$, denoted by $\mathcal{U}_i$, $i = 1, \dots, n$, such that $\mathcal{K} \subseteq \bigcup_{i=1}^n \mathcal{U}_i$. For $i = 1, \dots, n$, let L_i denote the Lipschitz constant of $f|_{\mathcal{U}_i}$. Now, as shown below, there exists a $\delta > 0$ such that, for every $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{K}$ with $\|\boldsymbol{x} - \boldsymbol{y}\|_2 < \delta$, we can find a $\mathcal{U}_i$ in the finite subcover of $\mathcal{U}$ with $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{U}_i$. With this δ we let

$$L = \max\left\{L_1, \dots, L_n, \frac{2\Delta}{\delta}\right\}, \tag{369}$$

where $\Delta := \max_{\boldsymbol{x}\in\mathcal{K}} f(\boldsymbol{x})$ (Note that the local Lipschitz property of f implies its continuity and, therefore, f attains its maximum on the compact set $\mathcal{K}$.) Consider an arbitrary pair $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{K}$. If $\|\boldsymbol{x} - \boldsymbol{y}\|_2 < \delta$, then this pair must be in the same $\mathcal{U}_i$, which implies $\|f(\boldsymbol{x}) - f(\boldsymbol{y})\| \le L_i\|\boldsymbol{x} - \boldsymbol{y}\| \le L\|\boldsymbol{x} - \boldsymbol{y}\|$. If $\|\boldsymbol{x} - \boldsymbol{y}\|_2 \ge \delta$, then $\|f(\boldsymbol{x}) - f(\boldsymbol{y})\| \le 2\Delta = 2\Delta\delta/\delta \le (2\Delta/\delta)\|\boldsymbol{x} - \boldsymbol{y}\|_2 \le L\|\boldsymbol{x} - \boldsymbol{y}\|_2$. Thus, $f|_{\mathcal{K}}$ is Lipschitz with Lipschitz constant at most L. It remains to establish the existence of a δ with the desired properties. Towards a contradiction, suppose that such a δ does not exist. This implies that, for every $m \in \mathbb{N}$, we can find a pair $\boldsymbol{x}_m, \boldsymbol{y}_m \in \mathcal{K}$ such that $\|\boldsymbol{x}_m - \boldsymbol{y}_m\|_2 < 1/m$, but there is no $\mathcal{U}_i$ containing this pair. Since $\mathcal{K} \times \mathcal{K}$ is compact by Tychonoff's Theorem [44, Theorem 4.42], there exists a convergent subsequence $\{\boldsymbol{x}_{m_j}, \boldsymbol{y}_{m_j}\}_{j\in\mathbb{N}}$ of the sequence $\{\boldsymbol{x}_m, \boldsymbol{y}_m\}_{m\in\mathbb{N}}$ [42, Theorem 2.41]. Let $(\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}})$ denote the limit point of this subsequence. Note that $\bar{\boldsymbol{x}}$ and $\bar{\boldsymbol{y}}$ cannot be in the same $\mathcal{U}_i$, for if they were there would be an $M \in \mathbb{N}$ such that $(\boldsymbol{x}_{m_j}, \boldsymbol{y}_{m_j}) \in \mathcal{U}_i \times \mathcal{U}_i$ for all $j \ge M$, which is not possible as $\boldsymbol{x}_m$ and $\boldsymbol{y}_m$ are in different sets $\mathcal{U}_i$ for all $m \in \mathbb{N}$. But

$$\bar{\boldsymbol{x}} - \bar{\boldsymbol{y}} = \lim_{j\to\infty} (\boldsymbol{x}_{m_j} - \boldsymbol{y}_{m_j}) \tag{370}$$

$$= \boldsymbol{0}, \tag{371}$$

where the second equality follows from $\lim_{j\to\infty} \|\boldsymbol{x}_{m_j} - \boldsymbol{y}_{m_j}\|_2 \le \lim_{j\to\infty} 1/m_j = 0$ and the continuity of $\|\cdot\|_2$. Thus, $\bar{\boldsymbol{x}} = \bar{\boldsymbol{y}}$, which is not possible as there is no $\mathcal{U}_i$ containing $\bar{\boldsymbol{x}}$ and $\bar{\boldsymbol{y}}$.

„⇐": Suppose that $f|_{\mathcal{K}}$ is Lipschitz for all compact sets $\mathcal{K} \subseteq \mathbb{R}^k$. It is sufficient to show that $f|_{\mathcal{B}_k(\boldsymbol{x},1)}$ is Lipschitz for every $\boldsymbol{x} \in \mathbb{R}^k$. As $\overline{\mathcal{B}_k(\boldsymbol{x}, 1)}$ is compact, $f|_{\overline{\mathcal{B}_k(\boldsymbol{x},1)}}$ is Lipschitz. The Lipschitz property of $f|_{\mathcal{B}_k(\boldsymbol{x},1)}$ therefore follows immediately from $\mathcal{B}_k(\boldsymbol{x}, 1) \subseteq \overline{\mathcal{B}_k(\boldsymbol{x}, 1)}$. □

Lemma H.13. If $f\colon \mathbb{R}^m \to \mathbb{R}^n$ is locally Lipschitz and $\mathcal{K} \subseteq \mathbb{R}^m$ is compact, then $f(\mathcal{K})$ is compact.

Proof. Follows from the continuity of f and [53, Theorem 2-7.2]. □

We will need the following composition property of locally Lipschitz mappings.

Lemma H.14. Suppose that $f\colon \mathbb{R}^k \to \mathbb{R}^m$ and $g\colon \mathbb{R}^m \to \mathbb{R}^n$ are both locally Lipschitz. Then, $h = g \circ f\colon \mathbb{R}^k \to \mathbb{R}^n$ is locally Lipschitz.

Proof. To prove that $h = g \circ f$ is locally Lipschitz, by Lemma H.12, it is sufficient to show that $h|_{\mathcal{K}}$ is Lipschitz for all compact sets $\mathcal{K} \subseteq \mathbb{R}^k$. Let $\mathcal{K} \subseteq \mathbb{R}^k$ be an arbitrary but fixed compact set and note that $\mathcal{Q} = f(\mathcal{K})$ is compact owing to Lemma H.13. Thus, by Lemma H.12, $f|_{\mathcal{K}}$ and $g|_{\mathcal{Q}}$ are both Lipschitz with Lipschitz constants, say, L and M, respectively. It follows that

$$\|h(\boldsymbol{u}) - h(\boldsymbol{v})\|_2 \le M\|f(\boldsymbol{u}) - f(\boldsymbol{v})\|_2 \tag{372}$$

$$\le LM\|\boldsymbol{u} - \boldsymbol{v}\|_2 \tag{373}$$

for all $\boldsymbol{u}$ and $\boldsymbol{v}$ in $\mathcal{K}$, which implies that $h|_{\mathcal{K}}$ is Lipschitz. As $\mathcal{K}$ was arbitrary, we can conclude that h is locally Lipschitz. □

Definition H.8. (Differentiable mapping) [54, Definition 1.2] Let $\mathcal{U}$ be an open set in $\mathbb{R}^m$. A function $f\colon \mathcal{U} \to \mathbb{R}$ is

i) C^0 if it is continuous.

ii) C^r with $r \in \mathbb{N}$ if, for all possible choices of $r_1, \ldots, r_m \in \mathbb{N}_0$ with $\sum_{i=1}^m r_i = r$, the partial derivatives

$$\frac{\partial^r}{\partial x_1^{r_1} \ldots \partial x_m^{r_m}} f(\boldsymbol{x}) \tag{374}$$

exist and are continuous on $\mathcal{U}$.

iii) C^∞ if it is C^r for all $r \in \mathbb{N}$.

For $r \in \mathbb{N}_0 \cup \{\infty\}$, a mapping $f\colon \mathcal{U} \to \mathbb{R}^n$, $\boldsymbol{x} \mapsto (f_1(\boldsymbol{x}) \ldots f_n(\boldsymbol{x}))^\mathsf{T}$ is C^r if every component f_i, $i = 1, \ldots, n$, is C^r.

It follows from the mean value theorem [55, Theorem 3.4] that C^1 mappings are locally Lipschitz. Conversely, by Rademacher's Theorem [36, Theorem. 5.1.11], every locally Lipschitz mapping $f\colon \mathbb{R}^m \to \mathbb{R}^n$ has a $\mathscr{L}(\mathbb{R}^m)$-measurable (but not necessarily continuous) differential Df, which is defined λ^m-almost everywhere.

Theorem H.2. (Sard's theorem) [56, Theorems 4.1 and 7.2] Let $f\colon \mathbb{R}^m \to \mathbb{R}^n$, $\boldsymbol{x} \mapsto (f_1(\boldsymbol{x}) \ldots f_n(\boldsymbol{x}))^\mathsf{T}$ be C^r and set $\mathcal{A} = \{\boldsymbol{x} : Jf(\boldsymbol{x}) = 0\}$ with $Jf(\boldsymbol{x})$ as in (1). The following statements hold.

i) If $m \le n$, then $\mathscr{H}^m(f(\mathcal{A})) = 0$.

ii) If $m > n$ and $r \ge m - n + 1$, then $\lambda^n(f(\mathcal{A})) = 0$.

C. Area and Coarea Formula

Next, we state two fundamental results from geometric measure theory that are used frequently in the paper, namely, the area and the coarea formula for locally Lipschitz mappings.

Theorem H.3. (Area formula) [36, Theorem 5.1.1] If $\mathcal{A} \in \mathscr{B}(\mathbb{R}^m)$, $f\colon \mathbb{R}^m \to \mathbb{R}^n$ is Lipschitz, and $g\colon \mathcal{A} \to \mathbb{R}$ is nonnegative and Lebesgue measurable, and $m \le n$, then

$$\int_{\mathcal{A}} Jf(\boldsymbol{x})\,\mathrm{d}\lambda^m(\boldsymbol{x}) = \int_{\mathbb{R}^n} \operatorname{card}\big(\mathcal{A} \cap f^{-1}(\{\boldsymbol{y}\})\big)\,\mathrm{d}\mathscr{H}^m(\boldsymbol{y}). \tag{375}$$

Corollary H.3. (Area formula) If $\mathcal{A} \in \mathscr{B}(\mathbb{R}^m)$, $f\colon \mathbb{R}^m \to \mathbb{R}^n$ is locally Lipschitz and one-to-one on $\mathcal{A}$, and $m \le n$, then

$$\int_{\mathcal{A}} Jf(\boldsymbol{x})\,\mathrm{d}\lambda^m(\boldsymbol{x}) = \mathscr{H}^m(f(\mathcal{A})). \tag{376}$$

Proof. We have

$$\int_{\mathcal{A}} Jf(\boldsymbol{x})\,\mathrm{d}\lambda^m(\boldsymbol{x}) \tag{377}$$

$$= \lim_{i\to\infty} \int_{\mathcal{A}} \mathbb{1}_{\mathcal{B}_m(\mathbf{0},i)}(\boldsymbol{x}) Jf(\boldsymbol{x})\,\mathrm{d}\lambda^m(\boldsymbol{x}) \tag{378}$$

$$= \lim_{i\to\infty} \int_{\mathcal{A}} J(f|_{\mathcal{B}_m(\mathbf{0},i)})(\boldsymbol{x})\,\mathrm{d}\lambda^m(\boldsymbol{x}) \tag{379}$$

$$= \lim_{i\to\infty} \int_{\mathbb{R}^n} \operatorname{card}\Big(\mathcal{A} \cap f|_{\mathcal{B}_m(\mathbf{0},i)}^{-1}(\{\boldsymbol{y}\})\Big)\,\mathrm{d}\mathscr{H}^m(\boldsymbol{y}) \tag{380}$$

$$= \lim_{i\to\infty} \mathscr{H}^m\big(f\big(\mathcal{A} \cap \mathcal{B}_m(\mathbf{0},i)\big)\big) \tag{381}$$

$$= \mathscr{H}^m\Bigg(\bigcup_{i\in\mathbb{N}} f\big(\mathcal{A} \cap \mathcal{B}_m(\mathbf{0},i)\big)\Bigg) \tag{382}$$

$$= \mathscr{H}^m(f(\mathcal{A})), \tag{383}$$

where (378) is by the Lebesgue Monotone Convergence Theorem [50, Theorem 4.6] upon noting that $(\mathbb{1}_{\mathcal{B}_m(\mathbf{0},i)})_{i\in\mathbb{N}}$ is an increasing sequence of nonnegative Lebesgue measurable functions converging pointwise to the constant function 1, in (380) we applied Theorem H.3 to $f|_{\mathcal{B}_m(\mathbf{0},i)}$, which is Lipschitz by Lemma H.12 as $\mathcal{B}_m(\mathbf{0},i) \subseteq \overline{\mathcal{B}_m(\mathbf{0},i)}$ and $\overline{\mathcal{B}_m(\mathbf{0},i)}$ is compact for all $i \in \mathbb{N}$, in (381) we used that f is one-to-one, by assumption, which implies

$$\operatorname{card}\Big(\mathcal{A} \cap f|_{\mathcal{B}_m(\mathbf{0},i)}^{-1}(\{\boldsymbol{y}\})\Big) = \begin{cases} 1 & \text{if } \boldsymbol{y} \in f\big(\mathcal{A} \cap \mathcal{B}_m(\mathbf{0},i)\big) \\ 0 & \text{else,} \end{cases} \tag{384}$$

and (382) is by Property ii) in Lemma H.1. □

If $m = n$, by Property iii) in Lemma H.3, $\mathscr{H}^m$ can be replaced by λ^m in Theorem H.3 and in Corollary H.3.

Theorem H.4. (Coarea formula) [36, Theorem 5.2.1] If $f\colon \mathbb{R}^m \to \mathbb{R}^n$ is Lipschitz, $\mathcal{A} \in \mathscr{B}(\mathbb{R}^m)$, and $m \ge n$, then

$$\int_{\mathcal{A}} Jf(\boldsymbol{x})\,\mathrm{d}\lambda^m(\boldsymbol{x}) = \int_{\mathbb{R}^n} \mathscr{H}^{m-n}\big(\mathcal{A} \cap f^{-1}(\{\boldsymbol{y}\})\big)\,\mathrm{d}\lambda^n(\boldsymbol{y}). \tag{385}$$

Corollary H.4. (Coarea formula) If $f\colon \mathbb{R}^m \to \mathbb{R}^n$ is locally Lipschitz, $\mathcal{A} \in \mathscr{B}(\mathbb{R}^m)$, and $m \ge n$, then

$$\int_{\mathcal{A}} Jf(\boldsymbol{x})\,\mathrm{d}\lambda^m(\boldsymbol{x}) = \int_{\mathbb{R}^n} \mathscr{H}^{m-n}\big(\mathcal{A} \cap f^{-1}(\{\boldsymbol{y}\})\big)\,\mathrm{d}\lambda^n(\boldsymbol{y}). \tag{386}$$

Proof. We have

$$\int_{\mathcal{A}} Jf(\boldsymbol{x})\,\mathrm{d}\lambda^m(\boldsymbol{x}) \tag{387}$$

$$= \lim_{i\to\infty} \int_{\mathcal{A}} J(f|_{\mathcal{B}_m(\mathbf{0},i)})(\boldsymbol{x})\,\mathrm{d}\lambda^m(\boldsymbol{x}) \tag{388}$$

$$= \lim_{i\to\infty} \int_{\mathbb{R}^n} \mathscr{H}^{m-n}\Big(\mathcal{A} \cap f|_{\mathcal{B}_m(\mathbf{0},i)}^{-1}(\{\boldsymbol{y}\})\Big)\,\mathrm{d}\lambda^n(\boldsymbol{y}) \tag{389}$$

$$= \int_{\mathbb{R}^n} \mathscr{H}^{m-n}\Bigg(\bigcup_{i\in\mathbb{N}} \mathcal{A} \cap f|_{\mathcal{B}_m(\mathbf{0},i)}^{-1}(\{\boldsymbol{y}\})\Bigg)\,\mathrm{d}\lambda^n(\boldsymbol{y}) \tag{390}$$

$$= \int_{\mathbb{R}^n} \mathscr{H}^{m-n}\big(\mathcal{A} \cap f^{-1}(\{\boldsymbol{y}\})\big)\,\mathrm{d}\lambda^n(\boldsymbol{y}), \tag{391}$$

where (388) follows from (377)–(379), in (389) we applied Theorem H.4 to $f|_{\mathcal{B}_m(\mathbf{0},i)}$, which is Lipschitz by Lemma H.12 as $\mathcal{B}_m(\mathbf{0},i) \subseteq \overline{\mathcal{B}_m(\mathbf{0},i)}$ and $\overline{\mathcal{B}_m(\mathbf{0},i)}$ is compact for all $i \in \mathbb{N}$, and (390) is by the Lebesgue Monotone Convergence Theorem [50, Theorem 4.6] upon noting that, for every $i \in \mathbb{N}$, the function

$$g_i\colon \mathbb{R}^n \to \mathbb{R} \tag{392}$$

$$\boldsymbol{y} \mapsto \mathscr{H}^{m-n}\Big(\mathcal{A} \cap f|_{\mathcal{B}_m(\mathbf{0},i)}^{-1}(\{\boldsymbol{y}\})\Big) \tag{393}$$

is Lebesgue measurable [36, Lemma 5.2.5] and, therefore, $(g_i)_{i\in\mathbb{N}}$ is a sequence of nonnegative increasing Lebesgue measurable functions, with

$$\lim_{i\to\infty} g_i(\boldsymbol{y}) = \mathscr{H}^{m-n}\Bigg(\bigcup_{i\in\mathbb{N}} \mathcal{A} \cap f|_{\mathcal{B}_m(\mathbf{0},i)}^{-1}(\{\boldsymbol{y}\})\Bigg) \tag{394}$$

for all $\boldsymbol{y} \in \mathbb{R}^n$ by Property ii) in Lemma H.1. □

D. Properties of Modified Minkowski Dimension

In this section, we state some properties of modified Minkowski dimension (see Definitions II.1 and II.2).

Lemma H.15. Main properties of modified Minkowski dimension:

i) We have

$$\underline{\dim}_{\mathrm{MB}}(\mathcal{U}) = \inf\left\{\sup_{i\in\mathbb{N}} \underline{\dim}_{\mathrm{B}}(\mathcal{U}_i) : \mathcal{U} \subseteq \bigcup_{i\in\mathbb{N}} \mathcal{U}_i\right\} \quad (395)$$

and

$$\overline{\dim}_{\mathrm{MB}}(\mathcal{U}) = \inf\left\{\sup_{i\in\mathbb{N}} \overline{\dim}_{\mathrm{B}}(\mathcal{U}_i) : \mathcal{U} \subseteq \bigcup_{i\in\mathbb{N}} \mathcal{U}_i\right\}, \quad (396)$$

respectively, where in (395) and (396) the infima are over all possible coverings $\{\mathcal{U}_i\}_{i\in\mathbb{N}}$ of $\mathcal{U}$ by nonempty compact sets $\mathcal{U}_i$.

ii) Every s-dimensional C^1-submanifold [36, Definition 5.3.1] $\mathcal{U}$ of $\mathbb{R}^m$ has $\dim_{\mathrm{MB}}(\mathcal{U}) = s$.

iii) $\underline{\dim}_{\mathrm{MB}}(\cdot)$ and $\overline{\dim}_{\mathrm{MB}}(\cdot)$ are monotonically nondecreasing with respect to $\subseteq$.

iv) $\underline{\dim}_{\mathrm{MB}}(\mathcal{U}) \leq \underline{\dim}_{\mathrm{B}}(\mathcal{U})$ and $\overline{\dim}_{\mathrm{MB}}(\mathcal{U}) \leq \overline{\dim}_{\mathrm{B}}(\mathcal{U})$ for all nonempty sets $\mathcal{U}$.

v) If f is Lipschitz, then

$$\overline{\dim}_{\mathrm{MB}}(f(\mathcal{U})) \leq \overline{\dim}_{\mathrm{MB}}(\mathcal{U}) \quad (397)$$
$$\underline{\dim}_{\mathrm{MB}}(f(\mathcal{U})) \leq \underline{\dim}_{\mathrm{MB}}(\mathcal{U}) \quad (398)$$

for all nonempty subsets $\mathcal{U}$ in the domain of f.

vi) $\overline{\dim}_{\mathrm{MB}}$ and $\underline{\dim}_{\mathrm{MB}}$ are countably stable, i.e.,

$$\overline{\dim}_{\mathrm{MB}}\left(\bigcup_{i\in\mathbb{N}} \mathcal{U}_i\right) = \sup_{i\in\mathbb{N}} \overline{\dim}_{\mathrm{MB}}(\mathcal{U}_i) \quad (399)$$

and

$$\underline{\dim}_{\mathrm{MB}}\left(\bigcup_{i\in\mathbb{N}} \mathcal{U}_i\right) = \sup_{i\in\mathbb{N}} \underline{\dim}_{\mathrm{MB}}(\mathcal{U}_i) \quad (400)$$

for all countable collections of nonempty sets $\mathcal{U}_i$, $i \in \mathbb{N}$.

vii) Let $f\colon \mathbb{R}^m \to \mathbb{R}^n$ be locally Lipschitz. Then,

$$\overline{\dim}_{\mathrm{MB}}(f(\mathcal{U})) \leq \overline{\dim}_{\mathrm{MB}}(\mathcal{U}) \quad (401)$$
$$\underline{\dim}_{\mathrm{MB}}(f(\mathcal{U})) \leq \underline{\dim}_{\mathrm{MB}}(\mathcal{U}) \quad (402)$$

for all nonempty subsets $\mathcal{U} \subseteq \mathbb{R}^m$.

Proof. Property i) is by invariance of $\overline{\dim}_{\mathrm{B}}$ and $\underline{\dim}_{\mathrm{B}}$ under set closure [38, Chapter 2.2] and the Heine-Borel theorem [42, Theorem 2.41].

Properties ii)–iv) follow from Definition II.2 and the properties of lower and upper Minkowski dimension listed in [38, Chapter 2.2].

We prove Properties v)–vii) for $\underline{\dim}_{\mathrm{MB}}$ only. The corresponding arguments for $\overline{\dim}_{\mathrm{MB}}$ are along the same lines.

To prove Property v) for $\underline{\dim}_{\mathrm{MB}}$, let $f\colon \mathcal{D} \to \mathbb{R}^n$ be Lipschitz with domain $\mathcal{D} \subseteq \mathbb{R}^m$ and consider $\mathcal{U} \subseteq \mathbb{R}^m$. By [53, Theorem 2], there exists a Lipschitz mapping $g\colon \mathbb{R}^m \to \mathbb{R}^n$ with $g|_{\mathcal{D}} = f$. We have

$$\begin{aligned}\underline{\dim}_{\mathrm{MB}}(f(\mathcal{U})) &= \underline{\dim}_{\mathrm{MB}}(g(\mathcal{U})) && (403)\\ &= \inf\left\{\sup_{i\in\mathbb{N}} \underline{\dim}_{\mathrm{B}}(\mathcal{V}_i) : g(\mathcal{U}) \subseteq \bigcup_{i\in\mathbb{N}} \mathcal{V}_i\right\} && (404)\\ &\leq \inf\left\{\sup_{i\in\mathbb{N}} \underline{\dim}_{\mathrm{B}}(g(\mathcal{U}_i)) : \mathcal{U} \subseteq \bigcup_{i\in\mathbb{N}} \mathcal{U}_i\right\} && (405)\\ &\leq \inf\left\{\sup_{i\in\mathbb{N}} \underline{\dim}_{\mathrm{B}}(\mathcal{U}_i) : \mathcal{U} \subseteq \bigcup_{i\in\mathbb{N}} \mathcal{U}_i\right\} && (406)\\ &= \underline{\dim}_{\mathrm{MB}}(\mathcal{U}), && (407)\end{aligned}$$

where (404) is by Property i) with the infimum over all possible coverings $\{\mathcal{V}_i\}_{i\in\mathbb{N}}$ of $g(\mathcal{U})$ by nonempty compact sets $\mathcal{V}_i \subseteq \mathbb{R}^n$, in (405) we used that, by Lemma H.13, Lipschitz images of compact sets are again compact with the infimum over all possible coverings $\{\mathcal{U}_i\}_{i\in\mathbb{N}}$ of $\mathcal{U}$ by nonempty compact sets $\mathcal{U}_i \subseteq \mathbb{R}^m$, in (406) we used that $\underline{\dim}_{\mathrm{B}}(g(\mathcal{U}_i)) \leq \underline{\dim}_{\mathrm{B}}(\mathcal{U}_i)$ for all $i \in \mathbb{N}$ as g is Lipschitz [38, Proposition 2.5, Property (a)], and (407) is again by Property i).

Property vi) is stated in [38, Chapter 2.3] without proof. For the sake of completeness, we prove (400). Let

$$\mathcal{U} = \bigcup_{i\in\mathbb{N}} \mathcal{U}_i \quad (408)$$

with $\mathcal{U}_i \subset \mathbb{R}^m$ nonempty for all $i \in \mathbb{N}$. By the monotonicity of $\underline{\dim}_{\mathrm{MB}}$, we have

$$\underline{\dim}_{\mathrm{MB}}(\mathcal{U}) \geq \sup_{i\in\mathbb{N}} \underline{\dim}_{\mathrm{MB}}(\mathcal{U}_i). \quad (409)$$

It remains to show that

$$\underline{\dim}_{\mathrm{MB}}(\mathcal{U}) \leq \sup_{i\in\mathbb{N}} \underline{\dim}_{\mathrm{MB}}(\mathcal{U}_i). \quad (410)$$

Suppose first that the $\mathcal{U}_i$ are all bounded. It follows that

$$\begin{aligned}\underline{\dim}_{\mathrm{MB}}(\mathcal{U}) &= \inf\left\{\sup_{i\in\mathbb{N}} \underline{\dim}_{\mathrm{B}}(\mathcal{V}_i) : \mathcal{U} \subseteq \bigcup_{i\in\mathbb{N}} \mathcal{V}_i\right\} && (411)\\ &= \inf\left\{\sup_{i\in\mathbb{N}} \underline{\dim}_{\mathrm{MB}}(\mathcal{V}_i) : \mathcal{U} \subseteq \bigcup_{i\in\mathbb{N}} \mathcal{V}_i\right\} && (412)\\ &\leq \sup_{i\in\mathbb{N}} \underline{\dim}_{\mathrm{MB}}(\mathcal{U}_i), && (413)\end{aligned}$$

where (411) is by Definition II.2 with the infimum over all possible coverings $\{\mathcal{V}_i\}_{i\in\mathbb{N}}$ of $\mathcal{U}$ by nonempty bounded sets $\mathcal{V}_i \subseteq \mathbb{R}^m$, in (412) we applied Lemma H.16 below, and in (413) we used that, by assumption, $\{\mathcal{U}_i\}_{i\in\mathbb{N}}$ is a covering of $\mathcal{U}$ consisting of nonempty bounded sets $\mathcal{U}_i \subseteq \mathbb{R}^m$, which establishes (410) for the case where the $\mathcal{U}_i$ are all bounded.

Next, suppose that the $\mathcal{U}_i$ are not necessarily all bounded. We can write

$$\mathcal{U} = \bigcup_{i,j\in\mathbb{N}} \mathcal{C}_{i,j} \quad (414)$$

with $\mathcal{C}_{i,j} = \mathcal{U}_i \cap \mathcal{B}_m(\mathbf{0}, j)$ for all $i, j \in \mathbb{N}$, so that

$$\underline{\dim}_{\mathrm{MB}}(\mathcal{U}) = \sup_{i,j\in\mathbb{N}} \underline{\dim}_{\mathrm{MB}}(\mathcal{C}_{i,j}) \tag{415}$$

$$\leq \sup_{i\in\mathbb{N}} \underline{\dim}_{\mathrm{MB}}(\mathcal{U}_i), \tag{416}$$

where in (415) we rely on Property vi) for nonempty bounded sets, and (416) follows from $\mathcal{C}_{i,j} \subseteq \mathcal{U}_i$ for all $i, j \in \mathbb{N}$ and the monotonicity of $\underline{\dim}_{\mathrm{MB}}$; this establishes (410) for the case where the $\mathcal{U}_i$ are not necessarily all bounded.

It remains to establish Property vii). Consider a locally Lipschitz mapping $f \colon \mathbb{R}^m \to \mathbb{R}^n$ and let $\mathcal{U} \subseteq \mathbb{R}^m$ be nonempty. We have

$$\underline{\dim}_{\mathrm{MB}}(f(\mathcal{U})) = \underline{\dim}_{\mathrm{MB}}\left(\bigcup_{j\in\mathbb{N}} f(\mathcal{U} \cap \mathcal{B}_m(\mathbf{0}, j))\right) \tag{417}$$

$$= \sup_{j\in\mathbb{N}}(\underline{\dim}_{\mathrm{MB}}(f(\mathcal{U} \cap \mathcal{B}_m(\mathbf{0}, j)))) \tag{418}$$

$$\leq \sup_{j\in\mathbb{N}}(\underline{\dim}_{\mathrm{MB}}(\mathcal{U} \cap \mathcal{B}_m(\mathbf{0}, j))) \tag{419}$$

$$= \underline{\dim}_{\mathrm{MB}}\left(\bigcup_{j\in\mathbb{N}} \mathcal{U} \cap \mathcal{B}_m(\mathbf{0}, j)\right) \tag{420}$$

$$= \underline{\dim}_{\mathrm{MB}}(\mathcal{U}), \tag{421}$$

where (418) and (420) are by countable stability of $\underline{\dim}_{\mathrm{MB}}$, and in (419) we applied Property v) to $f|_{\overline{\mathcal{U}\cap\mathcal{B}_m(\mathbf{0},j)}}$, which is Lipschitz thanks to Lemma H.12 with $\overline{\mathcal{U} \cap \mathcal{B}_m(\mathbf{0}, j)}$ compact by the Heine-Borel theorem [42, Theorem 2.41], for all $j \in \mathbb{N}$. ☐

Lemma H.16. Let $\mathcal{U} \subseteq \mathbb{R}^m$ be nonempty. Then, $\underline{\dim}_{\mathrm{MB}}(\mathcal{U}) = \underline{F}(\mathcal{U})$ with

$$\underline{F}(\mathcal{U}) = \inf\left\{\sup_{i\in\mathbb{N}} \underline{\dim}_{\mathrm{MB}}(\mathcal{U}_i) : \mathcal{U} \subseteq \bigcup_{i\in\mathbb{N}} \mathcal{U}_i\right\} \tag{422}$$

and $\overline{\dim}_{\mathrm{MB}}(\mathcal{U}) = \overline{F}(\mathcal{U})$ with

$$\overline{F}(\mathcal{U}) = \inf\left\{\sup_{i\in\mathbb{N}} \overline{\dim}_{\mathrm{MB}}(\mathcal{U}_i) : \mathcal{U} \subseteq \bigcup_{i\in\mathbb{N}} \mathcal{U}_i\right\}, \tag{423}$$

respectively, where in both cases the infimum is over all possible coverings $\{\mathcal{U}_i\}_{i\in\mathbb{N}}$ of $\mathcal{U}$ by nonempty bounded sets $\mathcal{U}_i \subseteq \mathbb{R}^m$.

Proof. We present a proof for $\underline{\dim}_{\mathrm{MB}}$ only, the arguments for $\overline{\dim}_{\mathrm{MB}}$ are along the same lines. First note that $\underline{F}(\mathcal{U}) \leq \underline{\dim}_{\mathrm{MB}}(\mathcal{U})$ as $\underline{\dim}_{\mathrm{MB}} \leq \underline{\dim}_{\mathrm{B}}$ by Property iv) in Lemma H.15. It remains to show that $\underline{F}(\mathcal{U}) \geq \underline{\dim}_{\mathrm{MB}}(\mathcal{U})$. Suppose, towards a contradiction, that $\underline{F}(\mathcal{U}) < \underline{\dim}_{\mathrm{MB}}(\mathcal{U})$ and set $\Delta = \underline{\dim}_{\mathrm{MB}}(\mathcal{U}) - \underline{F}(\mathcal{U}) > 0$. By definition of $\underline{F}(\mathcal{U})$, there must exist a collection $\{\mathcal{U}_i\}_{i\in\mathbb{N}}$ of nonempty bounded sets $\mathcal{U}_i \subseteq \mathbb{R}^m$ such that

$$\mathcal{U} \subseteq \bigcup_{i\in\mathbb{N}} \mathcal{U}_i \tag{424}$$

and

$$\sup_{i\in\mathbb{N}} \underline{\dim}_{\mathrm{MB}}(\mathcal{U}_i) \leq \underline{F}(\mathcal{U}) + \frac{\Delta}{3}. \tag{425}$$

Similarly, by Definition II.2, for every $i \in \mathbb{N}$, there must exist a collection $\{\mathcal{V}_j^{(i)}\}_{j\in\mathbb{N}}$ of nonempty bounded sets $\mathcal{V}_j^{(i)}$ such that

$$\mathcal{U}_i \subseteq \bigcup_{j\in\mathbb{N}} \mathcal{V}_j^{(i)} \tag{426}$$

and

$$\sup_{j\in\mathbb{N}} \underline{\dim}_{\mathrm{B}}\left(\mathcal{V}_j^{(i)}\right) \leq \underline{\dim}_{\mathrm{MB}}(\mathcal{U}_i) + \frac{\Delta}{3}. \tag{427}$$

Combining (424) with (426) yields

$$\mathcal{U} \subseteq \bigcup_{i,j\in\mathbb{N}} \mathcal{V}_j^{(i)}. \tag{428}$$

Next, note that

$$\underline{\dim}_{\mathrm{MB}}(\mathcal{U}) \leq \sup_{i,j\in\mathbb{N}} \underline{\dim}_{\mathrm{B}}\left(\mathcal{V}_j^{(i)}\right) \tag{429}$$

$$= \sup_{i\in\mathbb{N}} \sup_{j\in\mathbb{N}} \underline{\dim}_{\mathrm{B}}\left(\mathcal{V}_j^{(i)}\right) \tag{430}$$

$$\leq \sup_{i\in\mathbb{N}} \underline{\dim}_{\mathrm{MB}}(\mathcal{U}_i) + \frac{\Delta}{3} \tag{431}$$

$$\leq \underline{F}(\mathcal{U}) + \frac{2\Delta}{3} \tag{432}$$

$$< \underline{F}(\mathcal{U}) + \Delta, \tag{433}$$

where (429) follows from Definition II.2 and (428), in (431) we used (427), and (432) is by (425). This is in contradiction to $\Delta = \underline{\dim}_{\mathrm{MB}}(\mathcal{U}) - \underline{F}(\mathcal{U})$. ☐

Appendix I
Properties of Set-Valued Mappings

A set-valued mapping $\Phi \colon \mathcal{T} \to 2^{\mathbb{R}^m}$ associates to each $x \in \mathcal{T}$ a set $\Phi(x) \subseteq \mathbb{R}^m$. Many properties of ordinary mappings such as, e.g., measurability, can be extended to set-valued mappings. In this appendix, we first briefly review properties of set-valued mappings and then state a result needed in the existence proof of a measurable decoder in Section V-A.

Definition I.1. (Closed-valuedness of set-valued mappings) [57, Chapter 5] A set-valued mapping $\Phi \colon \mathcal{T} \to 2^{\mathbb{R}^m}$ is closed-valued if, for every $t \in \mathcal{T}$, the set $\Phi(t)$ is closed.

Definition I.2. (Inverse image of a set-valued mapping) [57, Chapter 14] For a set-valued mapping $\Phi \colon \mathcal{T} \to 2^{\mathbb{R}^m}$, the inverse image $\Phi^{-1}(\mathcal{A})$ of $\mathcal{A} \subseteq \mathbb{R}^m$ is

$$\Phi^{-1}(\mathcal{A}) = \{t \in \mathcal{T} : \Phi(t) \cap \mathcal{A} \neq \emptyset\}. \tag{434}$$

Definition I.3. (Measurable set-valued mapping) [57, Chapter 14] Let $(\mathcal{T}, \mathscr{S}(\mathcal{T}))$ be a measurable space. A set-valued mapping $\Phi \colon \mathcal{T} \to 2^{\mathbb{R}^m}$ is $\mathscr{S}(\mathcal{T})$-measurable if, for every open set $\mathcal{O} \subseteq \mathbb{R}^m$, the inverse image $\Phi^{-1}(\mathcal{O}) \in \mathscr{S}(\mathcal{T})$.

Lemma I.1. [57, Theorem 14.3] Let $(\mathcal{T}, \mathscr{S}(\mathcal{T}))$ be a measurable space and $\Phi \colon \mathcal{T} \to 2^{\mathbb{R}^m}$ closed-valued. Then, Φ is $\mathscr{S}(\mathcal{T})$-measurable if and only if $\Phi^{-1}(\mathcal{K}) \in \mathscr{S}(\mathcal{T})$ for all compact sets $\mathcal{K} \subseteq \mathbb{R}^m$.

Lemma I.2. [57, Corollary 14.6] Let $(\mathcal{T}, \mathscr{S}(\mathcal{T}))$ be a measurable space and $\Phi \colon \mathcal{T} \to 2^{\mathbb{R}^m}$ an $\mathscr{S}(\mathcal{T})$-measurable closed-valued mapping. Then, there exists a $\mathscr{S}(\mathcal{T})$-measurable mapping $f\colon \Phi^{-1}(\mathbb{R}^m) \to \mathbb{R}^m$ such that $f(t) \in \Phi(t)$ for all $t \in \Phi^{-1}(\mathbb{R}^m)$.

Definition I.4. (Normal integrand) [57, Definition 14.27] Let $(\mathcal{T}, \mathscr{S}(\mathcal{T}))$ be a measurable space. An extended real-valued function $f\colon \mathcal{T} \times \mathbb{R}^m \to \overline{\mathbb{R}}$ is a normal integrand with respect to $\mathscr{S}(\mathcal{T})$ if its epigraphical mapping

$$S_f\colon \mathcal{T} \to 2^{\mathbb{R}^m \times \mathbb{R}} \tag{435}$$

$$t \mapsto \{(\boldsymbol{x}, \alpha) \in \mathbb{R}^m \times \mathbb{R} : f(t, \boldsymbol{x}) \leq \alpha\} \tag{436}$$

is closed-valued and $\mathscr{S}(\mathcal{T})$-measurable.

Lemma I.3. [57, Example 14.31] Let $\mathcal{T} = \mathbb{R}^{n\times m} \times \mathbb{R}^n$ and suppose that $f\colon \mathcal{T} \times \mathbb{R}^m \to \overline{\mathbb{R}}$ is continuous. Then, f is a normal integrand with respect to $\mathscr{B}(\mathcal{T})$.

Lemma I.4. [57, Proposition 14.33] Let $(\mathcal{T}, \mathscr{S}(\mathcal{T}))$ be a measurable space. An extended real-valued function $f\colon \mathcal{T} \times \mathbb{R}^m \to \overline{\mathbb{R}}$ is a normal integrand with respect to $\mathscr{S}(\mathcal{T})$ if and only if, for every $\alpha \in \overline{\mathbb{R}}$, the level-set mapping

$$L_\alpha\colon \mathcal{T} \to 2^{\mathbb{R}^m} \tag{437}$$

$$t \mapsto \{\boldsymbol{x} \in \mathbb{R}^m : f(t, \boldsymbol{x}) \leq \alpha\} \tag{438}$$

is $\mathscr{S}(\mathcal{T})$-measurable and closed-valued.

We can now state the result on set-valued mappings needed to prove the existence of a measurable decoder in Section V-A.

Lemma I.5. Let $(\mathcal{T}, \mathscr{S}(\mathcal{T}))$ be a measurable space, $\alpha \in \mathbb{R}$, and $f\colon \mathcal{T} \times \mathbb{R}^m \to \overline{\mathbb{R}}$. Suppose that f is a normal integrand with respect to $\mathscr{S}(\mathcal{T})$ and $\mathcal{K} \subseteq \mathbb{R}^m$ is compact and nonempty. Then, the following properties hold.

i) The set-valued mapping

$$P_{\mathcal{K}}\colon \mathcal{T} \to 2^{\mathbb{R}^m} \tag{439}$$

$$t \mapsto \{\boldsymbol{x} \in \mathcal{K} : f(t, \boldsymbol{x}) \leq \alpha\} \tag{440}$$

is $\mathscr{S}(\mathcal{T})$-measurable and closed-valued.

ii) $P_{\mathcal{K}}^{-1}(\mathbb{R}^m) = \{t \in \mathcal{T} : \exists\, \boldsymbol{x} \in \mathcal{K} \text{ with } f(t, \boldsymbol{x}) \leq \alpha\} \in \mathscr{S}(\mathcal{T})$.

iii) There exists a $\mathscr{S}(\mathcal{T})$-measurable mapping

$$p_{\mathcal{K}}\colon P_{\mathcal{K}}^{-1}(\mathbb{R}^m) \to \mathbb{R}^m \tag{441}$$

$$t \mapsto p_{\mathcal{K}}(t) \in P_{\mathcal{K}}(t). \tag{442}$$

Proof. We start with the proof of i). For each $t \in \mathcal{T}$, we can write $P_{\mathcal{K}}(t) = L_\alpha(t) \cap \mathcal{K}$, where L_α is the $\mathscr{S}(\mathcal{T})$-measurable closed-valued level-set mapping from Lemma I.4. Since L_α is closed-valued and the intersection of a closed set with a compact set is closed, $P_{\mathcal{K}}$ is closed-valued. To prove that $P_{\mathcal{K}}$ is $\mathscr{S}(\mathcal{T})$-measurable it suffices, thanks to Lemma I.1, to show that $P_{\mathcal{K}}^{-1}(\mathcal{A}) \in \mathscr{S}(\mathcal{T})$ for all compact sets $\mathcal{A} \subseteq \mathbb{R}^m$. To this end, let $\mathcal{A} \subseteq \mathbb{R}^m$ be an arbitrary but fixed compact set. Since the intersection of two compact sets is compact it follows that $\mathcal{K} \cap \mathcal{A}$ is compact. As L_α is $\mathscr{S}(\mathcal{T})$-measurable and $\mathcal{K} \cap \mathcal{A}$ is compact, $L_\alpha^{-1}(\mathcal{K} \cap \mathcal{A}) \in \mathscr{S}(\mathcal{T})$ by Lemma I.1. Therefore, as $L_\alpha^{-1}(\mathcal{K} \cap \mathcal{A}) = P_{\mathcal{K}}^{-1}(\mathcal{A})$, we can conclude that $P_{\mathcal{K}}^{-1}(\mathcal{A}) \in \mathscr{S}(\mathcal{T})$. Since $\mathcal{A}$ was arbitrary, this proves i). Now, $\mathscr{S}(\mathcal{T})$-measurability of $P_{\mathcal{K}}$ implies $P_{\mathcal{K}}^{-1}(\mathbb{R}^m) \in \mathscr{S}(\mathcal{T})$, and thereby ii). Finally, the existence of the $\mathscr{S}(\mathcal{T})$-measurable mapping $p_{\mathcal{K}}$ in iii) follows from i) and Lemma I.2. □

Appendix J
Properties of Sequences of Functions in Several Variables

In this appendix, we summarize properties of sequences of functions in several variables needed in the proof of Theorem III.2. We start with a result that establishes a sufficient condition for uniform convergence.

Theorem J.1. (Weierstrass M-test) [42, Theorem 7.10] Consider a sequence $\{f_n\}_{n\in\mathbb{N}}$ of functions $f_n\colon \mathbb{R}^k \to \mathbb{R}$ and suppose that there exists a sequence $\{M_n\}_{n\in\mathbb{N}}$ of nonnegative real numbers such that

$$|f_n(\boldsymbol{x})| \leq M_n \quad \text{for all } \boldsymbol{x} \in \mathbb{R}^k \text{ and } n \in \mathbb{N} \tag{443}$$

and $\sum_{n\in\mathbb{N}} M_n < \infty$. Then, the sequence $\{s_n\}_{n\in\mathbb{N}}$ of partial sums

$$s_n = \sum_{i=1}^{n} f_i \tag{444}$$

converges uniformly to a function $f\colon \mathbb{R}^k \to \mathbb{R}$.

Next, we state a result that allows us to interchange the order of summation and differentiation for certain sequences of differentiable functions. We start with the corresponding statement for differentiable functions in one variable.

Theorem J.2. [42, Theorem 7.17] Consider a sequence $\{f_n\}_{n\in\mathbb{N}}$ of functions $f_n\colon \mathbb{R} \to \mathbb{R}$, each of which is differentiable on the closed interval $[a, b] \subseteq \mathbb{R}$, such that $\lim_{n\to\infty} f_n(x_0)$ exists and is finite for at least one $x_0 \in [a, b]$. Let f_n' denote the derivative of f_n, $n \in \mathbb{N}$. If $\{f_n'\}_{n\in\mathbb{N}}$ converges uniformly on $[a, b]$, then $\{f_n\}_{n\in\mathbb{N}}$ converges uniformly on $[a, b]$ to a function $f\colon \mathbb{R} \to \mathbb{R}$ satisfying

$$\frac{\mathrm{d}f(x)}{\mathrm{d}x} = \lim_{n\to\infty} f_n'(x) \quad \text{for all } x \in [a, b]. \tag{445}$$

Corollary J.1. Consider a sequence $\{f_n\}_{n\in\mathbb{N}}$ of functions $f_n\colon \mathbb{R}^k \to \mathbb{R}$ converging uniformly to $f\colon \mathbb{R}^k \to \mathbb{R}$. Suppose that there exists an $i \in \{1, \dots, k\}$ such that the partial derivatives $\partial f_n(\boldsymbol{x})/\partial x_i$ exist and are finite for all $\boldsymbol{x} \in \mathbb{R}^k$ and $n \in \mathbb{N}$ and that the sequence $\{\partial f_n/\partial x_i\}_{n\in\mathbb{N}}$ converges uniformly. Then,

$$\frac{\partial f(\boldsymbol{x})}{\partial x_i} = \lim_{n\to\infty} \frac{\partial f_n(\boldsymbol{x})}{\partial x_i} \quad \text{for all } \boldsymbol{x} \in \mathbb{R}^k. \tag{446}$$

Proof. Let $\boldsymbol{x} = (x_1 \dots x_k)^\mathsf{T} \in \mathbb{R}^k$ be arbitrary but fixed and denote by $\{g_n\}_{n\in\mathbb{N}}$ the sequence of functions defined according to $g_n(t) = f_n((x_1 \dots x_{i-1}\ t\ x_{i+1} \dots x_k)^\mathsf{T})$. Since $\{f_n\}_{n\in\mathbb{N}}$ converges uniformly to f by assumption, and $g_n(x_i) = f_n(\boldsymbol{x})$ for all $n \in \mathbb{N}$, it follows that

$$\lim_{n\to\infty} g_n(x_i) = \lim_{n\to\infty} f_n(\boldsymbol{x}) \tag{447}$$

$$= f(\boldsymbol{x}). \tag{448}$$

Since $\{\partial f_n/\partial x_i\}_{n\in\mathbb{N}}$ converges uniformly, so does $\{\mathrm{d}g_n/\mathrm{d}t\}_{n\in\mathbb{N}}$. In particular, there exists a closed interval $[a, b] \subseteq \mathbb{R}$ containing x_i such that $\{\mathrm{d}g_n/\mathrm{d}t\}_{n\in\mathbb{N}}$ converges uniformly on that interval $[a, b]$. Theorem J.2 therefore implies that $\{g_n\}_{n\in\mathbb{N}}$ converges uniformly on $[a,b]$ to a function $g\colon \mathbb{R}\to\mathbb{R}$ satisfying

$$\frac{\mathrm{d}g(t)}{\mathrm{d}t} = \lim_{n\to\infty} \frac{\mathrm{d}g_n(t)}{\mathrm{d}t} \quad \text{for all } t\in[a,b]. \tag{449}$$

But $g(x_i) = f(\boldsymbol{x})$ and $\mathrm{d}g_n(x_i)/\mathrm{d}x_i = \partial f_n(\boldsymbol{x})/\partial x_i$ for all $n\in\mathbb{N}$, which implies

$$\partial f(\boldsymbol{x})/\partial x_i = \lim_{n\to\infty} \partial f_n(\boldsymbol{x})/\partial x_i. \tag{450}$$

As $\boldsymbol{x}$ was arbitrary, this finishes the proof. □

Appendix K
Properties of Real Analytic Mappings and Consequences Thereof

In this appendix, we review material on real analytic mappings, on which our strong converse result in Sections IV and VII relies. We start with the definition of a real analytic mapping.

Definition K.1. (Real analytic mapping) [58, Definition 2.2.1] Let $\mathcal{U}$ be an open set in $\mathbb{R}^m$.

i) A function $f\colon \mathcal{U}\to\mathbb{R}$ is real analytic on $\mathcal{U}$ if, for each $\boldsymbol{x}\in\mathcal{U}$, f may be represented by a convergent power series (see Definition [58, Definition 2.1.4]) in some open neighborhood of $\boldsymbol{x}$; if $\mathcal{U}=\mathbb{R}^m$, then f is real analytic.
ii) A mapping $f\colon \mathcal{U}\to\mathbb{R}^n$, $\boldsymbol{x}\mapsto (f_1(\boldsymbol{x})\dots f_n(\boldsymbol{x}))^{\mathsf{T}}$ is real analytic on $\mathcal{U}$ if every component f_i, $i=1,\dots,n$, is real analytic on $\mathcal{U}$; if $\mathcal{U}=\mathbb{R}^m$, then f is real analytic.

Lemma K.1. [58, Corollary 1.2.4] If a power series

$$f(x) = \sum_{j=0}^{\infty} a_j (x-\alpha)^j \tag{451}$$

converges on an open interval $\mathcal{I}\subseteq\mathbb{R}$, then f is real analytic on $\mathcal{I}$.

Lemma K.2. [58, Proposition 2.2.2] Let $\mathcal{U}$ and $\mathcal{V}$ be open sets in $\mathbb{R}^m$. If $f\colon\mathcal{U}\to\mathbb{R}$ is real analytic on $\mathcal{U}$ and $g\colon\mathcal{V}\to\mathbb{R}$ is real analytic on $\mathcal{V}$, then $f+g$ and $f\cdot g$ are both real analytic on $\mathcal{U}\cap\mathcal{V}$. Furthermore, if $g(\boldsymbol{x})\neq 0$ for all $\boldsymbol{x}\in\mathcal{U}\cap\mathcal{V}$, then f/g is real analytic on $\mathcal{U}\cap\mathcal{V}$.

Corollary K.1. All polynomials on $\mathbb{R}^m$ are real analytic.

Lemma K.3. [58, Proposition 2.2.3] Let $\mathcal{U}$ be an open set in $\mathbb{R}^m$ and suppose that $f\colon\mathcal{U}\to\mathbb{R}$ is real analytic on $\mathcal{U}$. Then, the partial derivatives—of arbitrary order—of f are real analytic on $\mathcal{U}$.

Lemma K.4. [58, Proposition 2.2.8] Let $\mathcal{U}$ be an open set in $\mathbb{R}^m$ and $\mathcal{V}$ an open set in $\mathbb{R}^n$. Suppose that $f\colon\mathcal{U}\to\mathbb{R}^n$ is real analytic on $\mathcal{U}$ and $g\colon\mathcal{V}\to\mathbb{R}^k$ is real analytic on $\mathcal{V}$. Then, for every $\boldsymbol{x}\in\mathcal{U}$ with $f(\boldsymbol{x})\in\mathcal{V}$, there exists an open set $\mathcal{W}\subseteq\mathbb{R}^m$ such that $\boldsymbol{x}\in\mathcal{W}$ and $(g\circ f)|_{\mathcal{W}}\colon\mathcal{W}\to\mathbb{R}^k$ is real analytic on $\mathcal{W}$.

Corollary K.2. If $f\colon\mathbb{R}^m\to\mathbb{R}^n$ and $g\colon\mathbb{R}^n\to\mathbb{R}^k$ are both real analytic, then $g\circ f\colon\mathbb{R}^m\to\mathbb{R}^k$ is real analytic.

Lemma K.5. [58, p. 83] A real analytic function $f\colon\mathbb{R}^m\to\mathbb{R}$ vanishes either identically or on a set of Lebesgue measure zero.

The following definition extends the notion of a diffeomorphism of class C^r [40, Definition 3.1.18] to a real analytic diffeomorphism.

Definition K.2. (Real analytic diffeomorphism) Let $\mathcal{U}\subseteq\mathbb{R}^n$ be an open set and $\rho\colon\mathcal{U}\to\mathbb{R}^n$ real analytic. The mapping ρ is a real analytic diffeomorphism if $\mathcal{V}=\rho(\mathcal{U})$ is an open set and the inverse ρ^{-1} exists on $\mathcal{V}$ and is real analytic on $\mathcal{V}$.

Theorem K.1. [58, Theorem 2.5.1] Let $\mathcal{U}\subseteq\mathbb{R}^m$ be open and $f\colon\mathcal{U}\to\mathbb{R}^m$ real analytic. If $Jf(\boldsymbol{x}_0)>0$ for some $\boldsymbol{x}_0\in\mathcal{U}$, then f^{-1} exists on an open set $\mathcal{V}\subseteq\mathbb{R}^m$ containing $f(\boldsymbol{x}_0)$ and is real analytic on $\mathcal{V}$.

Corollary K.3. Let $\mathcal{U}$ be an open set in $\mathbb{R}^m$ and suppose that $f\colon\mathcal{U}\to\mathbb{R}^m$ is real analytic on $\mathcal{U}$. If there exists an $\boldsymbol{x}_0\in\mathcal{U}$ such that $Jf(\boldsymbol{x}_0)>0$, then there exists an $r>0$ such that $f|_{\mathcal{B}_m(\boldsymbol{x}_0,r)}$ is a real analytic diffeomorphism.

Proof. Suppose that there exists an $\boldsymbol{x}_0\in\mathcal{U}$ such that $Jf(\boldsymbol{x}_0)>0$. Theorem K.1 then implies that there exists an open set $\mathcal{V}\subseteq\mathbb{R}^m$ such that $f(\boldsymbol{x}_0)\in\mathcal{V}$, and f^{-1} exists on $\mathcal{V}$ and is real analytic on $\mathcal{V}$. Since $\mathcal{V}$ is open and $f(\boldsymbol{x}_0)\in\mathcal{V}$, there must exist an $\varepsilon>0$ such that $\mathcal{B}_m(f(\boldsymbol{x}_0),\varepsilon)\subseteq\mathcal{V}$. By continuity of f, which follows from real analyticity, there must exist an $r>0$ such that $f(\mathcal{B}_m(\boldsymbol{x}_0,r))\subseteq\mathcal{B}_m(f(\boldsymbol{x}_0),\varepsilon)\subseteq\mathcal{V}$. We set $\mathcal{W}=f(\mathcal{B}_m(\boldsymbol{x}_0,r))$. Summarizing, f is real analytic on $\mathcal{B}_m(\boldsymbol{x}_0,r)$ and f^{-1} exists on $\mathcal{W}=f(\mathcal{B}_m(\boldsymbol{x}_0,r))$ and is real analytic on $\mathcal{W}$. It remains to show that $\mathcal{W}$ is open. This follows by noting that $\mathcal{W}=(f^{-1})^{-1}(\mathcal{B}_m(\boldsymbol{x}_0,r))$ is the inverse image of an open set under a real analytic (and therefore continuous) mapping, and is hence open. We conclude that $f|_{\mathcal{B}_m(\boldsymbol{x}_0,r)}$ is a real analytic diffeomorphism. □

Corollary K.4. Let $\mathcal{U}$ be an open set in $\mathbb{R}^m$ and suppose that $f\colon\mathcal{U}\to\mathbb{R}^n$ with $n\geq m$ is real analytic on $\mathcal{U}$. If there exists an $\boldsymbol{x}_0\in\mathcal{U}$ such that $Jf(\boldsymbol{x}_0)>0$, then there exists an $r>0$ such that f is one-to-one on $\mathcal{B}_m(\boldsymbol{x}_0,r)$.

Proof. Suppose that there exists an $\boldsymbol{x}_0\in\mathcal{U}$ such that $Jf(\boldsymbol{x}_0)>0$. Then, $n\geq m$ implies $\operatorname{rank}(Df(\boldsymbol{x}_0))=m$. Thus, the $n\times m$ matrix $Df(\boldsymbol{x}_0)$ has m linearly independent row vectors. Denote the indices of m such linearly independent row vectors by $\{i_1,\dots,i_m\}$, and consider the mapping

$$g\colon\mathcal{U}\to\mathbb{R}^m \tag{452}$$
$$\boldsymbol{x}\mapsto (f_{i_1}(\boldsymbol{x})\dots f_{i_m}(\boldsymbol{x}))^{\mathsf{T}}. \tag{453}$$

Since f is real analytic on $\mathcal{U}$, so is g. Furthermore, $\operatorname{rank}(Dg(\boldsymbol{x}_0))=m$ and hence $Jg(\boldsymbol{x}_0)>0$. Corollary K.3 therefore implies the existence of an $r>0$ such that $g|_{\mathcal{B}_m(\boldsymbol{x}_0,r)}$ is a real analytic diffeomorphism. In particular, g is one-to-one on $\mathcal{B}_m(\boldsymbol{x}_0,r)$, which in turn implies that f is one-to-one on $\mathcal{B}_m(\boldsymbol{x}_0,r)$. □

Next, we show that the square root is a real analytic diffeomorphism on the set of positive real numbers.

Lemma K.6. The function $\sqrt{\cdot}\colon\mathbb{R}_+\to\mathbb{R}_+$, $x\mapsto\sqrt{x}$, where $\mathbb{R}_+=\{x\in\mathbb{R}: x>0\}$, is a real analytic diffeomorphism.

Proof. The function $(\cdot)^2\colon\mathbb{R}_+\to\mathbb{R}_+$, $x\mapsto x^2$ is real analytic by Corollary K.1. Let $y\in\mathbb{R}_+$ be arbitrary but fixed and set

$x = \sqrt{y}$. Since $\mathrm{d}y/\mathrm{d}x = 2x > 0$, Theorem K.1 implies that there exists an $r > 0$ such that the inverse of $(\cdot)^2$, given by $\sqrt{\cdot}$, exists on $(y-r, y+r)$ and is real analytic on $(y-r, y+r)$. As y was arbitrary, it follows that $\sqrt{\cdot}$ is real analytic on $\mathbb{R}_+$. Finally, since $(\cdot)^2$ is the inverse of $\sqrt{\cdot}$ and $\sqrt{\mathbb{R}_+} = \mathbb{R}_+$ is open, $\sqrt{\cdot}$ is a real analytic diffeomorphism. □

Lemma K.7. If $f\colon \mathbb{R}^m \to \mathbb{R}^n$ is real analytic, so is Jf. In particular, Jf vanishes either identically or on a set of Lebesgue measure zero.

Proof. Suppose that $f\colon \mathbb{R}^m \to \mathbb{R}^n$ is real analytic. Recall that $Jf(\boldsymbol{x}) = \sqrt{g(\boldsymbol{x})}$, where $g\colon \mathbb{R}^m \to \mathbb{R}$,

$$g(\boldsymbol{x}) = \begin{cases} \det(Df(\boldsymbol{x})(Df(\boldsymbol{x}))^{\mathsf{T}}) & \text{if } n < m \\ \det((Df(\boldsymbol{x}))^{\mathsf{T}} Df(\boldsymbol{x})) & \text{else.} \end{cases} \tag{454}$$

Lemmata K.2 and K.3 imply that g is real analytic. As $\sqrt{\cdot}$ is real analytic on $\mathbb{R}_+$ by Lemma K.6, real analyticity of Jf follows from Lemma K.4. Finally, as Jf is real analytic, it vanishes by Lemma K.5 either identically or on a set of Lebesgue measure zero. □

We have the following important properties of real analytic mappings.

Lemma K.8. Let $h\colon \mathbb{R}^s \to \mathbb{R}^m$ be a real analytic mapping of s-dimensional Jacobian $Jh \not\equiv 0$. Then, the following properties hold.

i) The set $\mathcal{O} = \{\boldsymbol{z} \in \mathbb{R}^s : Jh(\boldsymbol{z}) > 0\}$ is open and satisfies $\lambda^s(\mathbb{R}^s \backslash \mathcal{O}) = 0$.
ii) For every set $\mathcal{A} \in \mathscr{B}(\mathbb{R}^s)$ of positive Lebesgue measure, there exists a set $\mathcal{B} \in \mathscr{B}(\mathbb{R}^s)$ of positive Lebesgue measure such that $\mathcal{B} \subseteq \mathcal{A}$ and the mapping $h|_{\mathcal{B}}$ is an embedding, i.e., $Jh(\boldsymbol{z}) > 0$ for all $\boldsymbol{z} \in \mathcal{B}$ and h is one-to-one on $\mathcal{B}$.

Proof. Openness of $\mathcal{O}$ follows from continuity of Jh, and Lemma K.7 together with $Jh \not\equiv 0$ implies $\lambda^s(\mathbb{R}^s \backslash \mathcal{O}) = 0$. To prove ii), consider $\mathcal{A} \in \mathscr{B}(\mathbb{R}^s)$ of positive Lebesgue measure. As $\lambda^s(\mathbb{R}^s \backslash \mathcal{O}) = 0$, with $\mathcal{O}$ from i), it follows that $\lambda^s(\mathcal{A} \cap \mathcal{O}) > 0$. Thus, by Lemma H.9, there must exist a $\boldsymbol{z}_0 \in \mathcal{A} \cap \mathcal{O}$ such that

$$\lambda^s(\mathcal{B}_s(\boldsymbol{z}_0, r) \cap \mathcal{A} \cap \mathcal{O}) > 0 \quad \text{for all } r > 0. \tag{455}$$

Since $Jh(\boldsymbol{z}_0) > 0$, which follows from $\boldsymbol{z}_0 \in \mathcal{O}$, Corollary K.4 implies that there must exist an $r_0 > 0$ such that h is one-to-one on $\mathcal{B}_s(\boldsymbol{z}_0, r_0)$. Setting $\mathcal{B} = \mathcal{B}_s(\boldsymbol{z}_0, r_0) \cap \mathcal{A} \cap \mathcal{O}$ concludes the proof. □

Definition K.3. (Real analytic submanifold) [58, Definition 2.7.1] A subset $\mathcal{M} \subseteq \mathbb{R}^n$ is an m-dimensional real analytic submanifold of $\mathbb{R}^n$ if, for each $\boldsymbol{y} \in \mathcal{M}$, there exist an open set $\mathcal{U} \subseteq \mathbb{R}^m$ and a real analytic immersion $f\colon \mathcal{U} \to \mathbb{R}^n$ such that $\boldsymbol{y} \in f(\mathcal{U})$ and open subsets of $\mathcal{U}$ are mapped onto relatively open subsets in $\mathcal{M}$. Here, a subset $\mathcal{W} \subseteq \mathcal{M}$ is called relatively open in $\mathcal{M}$ if there exists an open set $\mathcal{V} \subseteq \mathbb{R}^n$ such that $\mathcal{W} = \mathcal{V} \cap \mathcal{M}$.

Equivalent definitions of real analytic submanifolds are listed in [58, Proposition 2.7.3].

Lemma K.9. Let $\mathcal{M} \subseteq \mathbb{R}^s$ be a t-dimensional real analytic submanifold of $\mathbb{R}^s$ and $\boldsymbol{z}_0 \in \mathcal{M}$. Then, there exist a real analytic embedding $\zeta\colon \mathbb{R}^t \to \mathcal{M} \subseteq \mathbb{R}^s$ and an $\eta > 0$ such that

$$\zeta(\mathbf{0}) = \boldsymbol{z}_0, \tag{456}$$
$$\mathcal{B}_s(\boldsymbol{z}_0, \eta) \cap \mathcal{M} \subseteq \zeta(\mathbb{R}^t), \tag{457}$$

and $\zeta(\mathbb{R}^t)$ is relatively open in $\mathcal{M}$.

Proof. Since $\mathcal{M} \subseteq \mathbb{R}^s$ is a t-dimensional real analytic submanifold of $\mathbb{R}^s$, Definition K.3 implies the existence of an open set $\mathcal{U} \subseteq \mathbb{R}^t$ and a real analytic immersion $\xi\colon \mathcal{U} \to \mathbb{R}^s$ that maps open subsets of $\mathcal{U}$ onto relatively open subsets in $\mathcal{M}$ and satisfies $\boldsymbol{z}_0 = \xi(\boldsymbol{u}_0)$ with $\boldsymbol{u}_0 \in \mathcal{U}$. As ξ is an immersion,

$$\operatorname{rank}(D\xi(\boldsymbol{v})) = t \quad \text{for all } \boldsymbol{v} \in \mathcal{U}. \tag{458}$$

Since $\mathcal{U}$ is open and $\boldsymbol{u}_0 \in \mathcal{U}$, there exists a $\rho > 0$ such that $\mathcal{B}_t(\boldsymbol{u}_0, \rho) \subseteq \mathcal{U}$, which implies in turn that

$$\xi\big(\mathcal{B}_t(\boldsymbol{u}_0, \rho)\big) \subseteq \xi(\mathcal{U}). \tag{459}$$

Using Corollary K.4 and $J\xi(\boldsymbol{u}_0) > 0$ (recall that ξ is an immersion) we may choose ρ sufficiently small for ξ to be one-to-one on $\mathcal{B}_t(\boldsymbol{u}_0, \rho)$. As $\xi(\mathcal{B}_t(\boldsymbol{u}_0, \rho))$ is relatively open in $\mathcal{M}$, there must exist an open set $\mathcal{V} \subseteq \mathbb{R}^s$ such that

$$\xi\big(\mathcal{B}_t(\boldsymbol{u}_0, \rho)\big) = \mathcal{V} \cap \mathcal{M}. \tag{460}$$

Now, as $\mathcal{V}$ is open and $\boldsymbol{z}_0 = \xi(\boldsymbol{u}_0) \in \mathcal{V}$, we can find an $\eta > 0$ such that $\mathcal{B}_s(\boldsymbol{z}_0, \eta) \subseteq \mathcal{V}$, which, together with (460), yields

$$\mathcal{B}_s(\boldsymbol{z}_0, \eta) \cap \mathcal{M} \subseteq \xi\big(\mathcal{B}_t(\boldsymbol{u}_0, \rho)\big). \tag{461}$$

Let $\kappa\colon \mathbb{R}^t \to \mathcal{B}_t(\mathbf{0}, \rho)$ be the real analytic diffeomorphism constructed in Lemma K.10 below and set

$$\zeta\colon \mathbb{R}^t \to \mathbb{R}^s \tag{462}$$
$$\boldsymbol{v} \mapsto \xi(\kappa(\boldsymbol{v}) + \boldsymbol{u}_0). \tag{463}$$

The mapping ζ is a real analytic mapping by Lemmata K.2 and K.4. Clearly, ζ is one-to-one on $\mathbb{R}^t$ as κ is a diffeomorphism and ξ is one-to-one on $\mathcal{B}_t(\boldsymbol{u}_0, \rho)$. Since $\zeta(\mathbb{R}^t) = \xi(\mathcal{B}_t(\boldsymbol{u}_0, \rho))$, (461) establishes (457) and (460) proves that $\zeta(\mathbb{R}^t)$ is relatively open in $\mathcal{M}$. Finally, (456) follows from $\zeta(\mathbf{0}) = \xi(\kappa(\mathbf{0}) + \boldsymbol{u}_0) = \xi(\boldsymbol{u}_0) = \boldsymbol{z}_0$, where in the second equality we used (467) in Lemma K.10 below. It remains to show that ζ is an immersion, which is effected by proving that $\operatorname{rank}(D\zeta(\boldsymbol{v})) = t$ for all $\boldsymbol{v} \in \mathbb{R}^t$. The chain rule implies

$$D\zeta(\boldsymbol{v}) = (D\xi)(\kappa(\boldsymbol{v}) + \boldsymbol{u}_0) D\kappa(\boldsymbol{v}) \quad \text{for all } \boldsymbol{v} \in \mathbb{R}^t. \tag{464}$$

It now follows

i) from (469) in Lemma K.10 below that $\operatorname{rank}(D\kappa(\boldsymbol{v})) = t$ for all $\boldsymbol{v} \in \mathbb{R}^t$, and
ii) from (458) and $\kappa(\boldsymbol{v}) + \boldsymbol{u}_0 \in \mathcal{B}_t(\boldsymbol{u}_0, \rho) \subseteq \mathcal{U}$ that $\operatorname{rank}((D\xi)(\kappa(\boldsymbol{v}) + \boldsymbol{u}_0)) = t$ for all $\boldsymbol{v} \in \mathbb{R}^t$.

Applying Lemma K.12 to $(D\xi)(\kappa(\boldsymbol{v}) + \boldsymbol{u}_0) \in \mathbb{R}^{s \times t}$ and $D\kappa(\boldsymbol{v}) \in \mathbb{R}^{t \times t}$, and using i) and ii) above yields $\operatorname{rank}(D\zeta(\boldsymbol{v})) \geq t$ for all $\boldsymbol{v} \in \mathbb{R}^t$, which in turn implies $J\zeta(\boldsymbol{v}) > 0$ for all $\boldsymbol{v} \in \mathbb{R}^t$, thereby concluding the proof. □

Lemma K.10. For $\rho > 0$, the mapping

$$\kappa\colon \mathbb{R}^k \to \mathcal{B}_k(\mathbf{0}, \rho) \tag{465}$$

$$\boldsymbol{x} \mapsto \frac{\rho \boldsymbol{x}}{\sqrt{1 + \|\boldsymbol{x}\|_2^2}} \tag{466}$$

is a real analytic diffeomorphism on $\mathbb{R}^k$ satisfying

$$\kappa(\mathbf{0}) = \mathbf{0}, \tag{467}$$

$$\kappa(\mathbb{R}^k) = \mathcal{B}_k(\mathbf{0}, \rho), \tag{468}$$

$$\operatorname{rank}(D\kappa(\boldsymbol{x})) = k \quad \text{for all } \boldsymbol{x} \in \mathbb{R}^k. \tag{469}$$

Proof. It follows from the definition of κ that $\kappa(\mathbf{0}) = \mathbf{0}$. The mapping κ is real analytic thanks to Lemmata K.2, K.4, and K.6. Now, consider the mapping

$$\sigma\colon \mathcal{B}_k(\mathbf{0}, \rho) \to \mathbb{R}^k \tag{470}$$

$$\boldsymbol{y} \mapsto \frac{\boldsymbol{y}}{\sqrt{\rho^2 - \|\boldsymbol{y}\|_2^2}}. \tag{471}$$

Again, σ is real analytic thanks to Lemmata K.2, K.4, and K.6. Since $(\kappa \circ \sigma)(\boldsymbol{y}) = \boldsymbol{y}$ for all $\boldsymbol{y} \in \mathcal{B}_k(\mathbf{0}, \rho)$, it follows that $\kappa(\mathbb{R}^k) = \mathcal{B}_k(\mathbf{0}, \rho)$. Moreover, as $(\sigma \circ \kappa)(\boldsymbol{x}) = \boldsymbol{x}$ for all $\boldsymbol{x} \in \mathbb{R}^k$, σ is the inverse of κ, which establishes that κ is a real analytic diffeomorphism on $\mathbb{R}^k$. Finally, since $(\sigma \circ \kappa)(\boldsymbol{x}) = \boldsymbol{x}$ for all $\boldsymbol{x} \in \mathbb{R}^k$, the chain rule implies $\boldsymbol{I}_k = D(\sigma \circ \kappa)(\boldsymbol{x}) = (D\sigma)(\kappa(\boldsymbol{x}))D\kappa(\boldsymbol{x})$ for all $\boldsymbol{x} \in \mathbb{R}^k$, which yields (469). □

Proposition K.1. [58, Proposition 2.7.3] Let $\mathcal{M} \subseteq \mathbb{R}^n$. The following statements are equivalent:

i) $\mathcal{M}$ is a m-dimensional real analytic submanifold of $\mathbb{R}^n$.
ii) For each $\boldsymbol{z} \in \mathcal{M}$ there exist an open set $\mathcal{U} \subseteq \mathbb{R}^n$ with $\boldsymbol{z} \in \mathcal{U}$, a real analytic diffeomorphism $\rho\colon \mathcal{U} \to \mathbb{R}^n$, and a m-dimensional linear subspace $\mathcal{L} \subseteq \mathbb{R}^n$, such that

$$\rho(\mathcal{M} \cap \mathcal{U}) = \rho(\mathcal{U}) \cap \mathcal{L}. \tag{472}$$

Proposition K.1 allows us to state the following sufficient condition on the inverse image of a point under a real analytic function to result in a real analytic submanifold.

Lemma K.11. Let $\psi\colon \mathbb{R}^s \to \mathbb{R}$ be a real analytic function, $y_0 \in \psi(\mathbb{R}^s)$, and set $\mathcal{M}_{y_0} = \psi^{-1}(\{y_0\})$. Suppose that $J\psi(\boldsymbol{z}) > 0$ for all $\boldsymbol{z} \in \mathcal{M}_{y_0}$. Then, $\mathcal{M}_{y_0}$ is a $(s-1)$-dimensional real analytic submanifold of $\mathbb{R}^s$.

Proof. We may assume, w.l.o.g., that $y_0 = 0$ (if $y_0 \neq 0$, we set $\tilde{\psi}(\boldsymbol{z}) = \psi(\boldsymbol{z}) - y_0$ and prove the lemma with $\tilde{\psi}$ in place of ψ, noting that $\mathcal{M}_{y_0} = \tilde{\psi}^{-1}(\{0\})$ and $J\tilde{\psi}(\boldsymbol{z}) = J\psi(\boldsymbol{z})$ for all $\boldsymbol{z} \in \mathbb{R}^s$.) The proof is effected by verifying that $\mathcal{M}_0$ satisfies Statement ii) of Proposition K.1. Let $\boldsymbol{z}_0 \in \mathcal{M}_0$ be arbitrary but fixed and set $\boldsymbol{a}_1^{\mathsf{T}} = D\psi(\boldsymbol{z}_0)$. Since $J\psi(\boldsymbol{z}_0) > 0$ by assumption, we must have $\boldsymbol{a}_1 \neq \mathbf{0}$. Choose $\boldsymbol{a}_2, \dots, \boldsymbol{a}_s \in \mathbb{R}^s$ such that $\boldsymbol{A} = (\boldsymbol{a}_1 \dots \boldsymbol{a}_s)^{\mathsf{T}}$ is a regular matrix and consider the mapping

$$\rho\colon \mathbb{R}^s \to \mathbb{R}^s \tag{473}$$

$$\boldsymbol{z} \mapsto \big(\psi(\boldsymbol{z})\ \boldsymbol{a}_2^{\mathsf{T}}\boldsymbol{z} \dots \boldsymbol{a}_s^{\mathsf{T}}\boldsymbol{z}\big)^{\mathsf{T}}. \tag{474}$$

Note that $D\rho(\boldsymbol{z}_0) = \boldsymbol{A}$. Since $\operatorname{rank}(\boldsymbol{A}) = s$ by construction, $J\rho(\boldsymbol{z}_0) > 0$. It therefore follows from Corollary K.3 that there exists an $r > 0$ such that $\rho|_{\mathcal{B}_s(\boldsymbol{z}_0, r)}$ is a real analytic diffeomorphism. Finally, we can write

$$\rho(\mathcal{M}_0 \cap \mathcal{B}_s(\boldsymbol{z}_0, r)) = \rho(\{\boldsymbol{z} \in \mathcal{B}_s(\boldsymbol{z}_0, r) : \psi(\boldsymbol{z}) = 0\}) \tag{475}$$

$$= \rho(\{\boldsymbol{z} \in \mathcal{B}_s(\boldsymbol{z}_0, r) : \boldsymbol{e}_1^{\mathsf{T}}\rho(\boldsymbol{z}) = 0\}) \tag{476}$$

$$= \rho(\mathcal{B}_s(\boldsymbol{z}_0, r)) \cap \mathcal{L}, \tag{477}$$

where (475) follows from $\mathcal{M}_0 = \psi^{-1}(\{0\})$, (476) is by (474), and in (477) we set $\mathcal{L} = \{(w_1 \dots w_s)^{\mathsf{T}} \in \mathbb{R}^s : w_1 = 0\}$. □

Lemma K.12. (Sylvester's inequality) [59, Chapter 0.4.5, Property (c)] If $\boldsymbol{A} \in \mathbb{R}^{m \times k}$ and $\boldsymbol{B} \in \mathbb{R}^{k \times n}$, then

$$\operatorname{rank}(\boldsymbol{A}) + \operatorname{rank}(\boldsymbol{B}) - k \leq \operatorname{rank}(\boldsymbol{AB}) \tag{478}$$

$$\leq \min\{\operatorname{rank}(\boldsymbol{A}), \operatorname{rank}(\boldsymbol{B})\}. \tag{479}$$

ACKNOWLEDGMENTS

The authors are thankful to the reviewers for very insightful comments, which helped us to improve our manuscript.

REFERENCES

[1] G. Alberti, H. Bölcskei, C. De Lellis, G. Koliander, and E. Riegler, "Lossless linear analog compression," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2016, pp. 2789–2793.
[2] D. L. Donoho and P. B. Stark, "Uncertainty principles and signal recovery," *SIAM J. Appl. Math.*, vol. 49, no. 3, pp. 906–931, Jun. 1989.
[3] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
[4] 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.
[5] E. J. Candès, Y. C. Eldar, D. Needell, and P. Randall, "Compressed sensing with coherent and redundant dictionaries," *Appl. Comp. Harm. Anal.*, vol. 31, no. 1, pp. 59–73, Jul. 2011.
[6] P. Kuppinger, G. Durisi, and H. Bölcskei, "Uncertainty relations and sparse signal recovery for pairs of general signal sets," *IEEE Trans. Inf. Theory*, vol. 58, no. 1, pp. 263–277, Jan. 2012.
[7] R. Gribonval and M. Nielsen, "Sparse representations in unions of bases," *IEEE Trans. Inf. Theory*, vol. 49, no. 12, pp. 3320–3325, Jan. 2003.
[8] M. Elad and A. M. Bruckstein, "A generalized uncertainty principle and sparse representation in pairs of bases," *IEEE Trans. Inf. Theory*, vol. 48, no. 9, pp. 2558–2567, Sep. 2002.
[9] D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ^1 minimization," in *Proc. Natl. Acad. Sci. USA*, vol. 100, no. 5, 2003, pp. 2197–2002.
[10] J. A. Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2231–2242, Oct. 2004.
[11] M. Elad, *Sparse and Redundant Representations – From Theory to Applications in Signal and Image Processing*. New York, NY: Springer, 2010.
[12] L. R. Welch, "Lower bounds on the maximum cross correlation of signals," *IEEE Trans. Inf. Theory*, vol. 20, no. 3, pp. 397–399, May 1974.
[13] E. J. Candès and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.
[14] J. A. Tropp, "On the conditioning of random subdictionaries," *Appl. Comp. Harm. Anal.*, vol. 25, pp. 1–24, 2008.
[15] G. Pope, A. Bracher, and C. Studer, "Probabilistic recovery guarantees for sparsely corrupted signals," *IEEE Trans. Inf. Theory*, vol. 59, no. 5, pp. 3104–3116, May 2013.
[16] R. G. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constructive Approximation*, vol. 28, no. 3, pp. 253–263, Dec. 2008.

[17] 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.

[18] ——, "Optimal phase transitions in compressed sensing," *IEEE Trans. Inf. Theory*, vol. 58, no. 10, pp. 6241–6263, Oct. 2012.

[19] E. Lindenstrauss and M. Tsukamoto, "From rate distortion theory to metric mean dimension: Variational principle," *IEEE Trans. Inf. Theory*, vol. 64, no. 5, pp. 3590–3609, May 2018.

[20] A. Velozo and R. Velozo, "Rate distortion theory, metric mean dimension and measure theoretic entropy," *arXiv:1707.05762*, 2018.

[21] Y. Gutman and A. Śpiewak, "Metric mean dimension and analog compression," *arXiv:1812.00458*, 2018.

[22] D. Stotz, E. Riegler, E. Agustsson, and H. Bölcskei, "Almost lossless analog signal separation and probabilistic uncertainty relations," *IEEE Trans. Inf. Theory*, vol. 63, no. 9, pp. 5445–5460, Sep. 2017.

[23] G. Koliander, G. Pichler, E. Riegler, and F. Hlawatsch, "Entropy and source coding for integer-dimensional singular random variables," *IEEE Trans. Inf. Theory*, vol. 62, no. 11, pp. 6124–6154, Nov. 2016.

[24] Y. C. Eldar and M. Mishali, "Robust recovery of signals from a structured union of subspaces," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 5302–5316, Nov. 2009.

[25] R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hegde, "Model-based compressive sensing," *IEEE Trans. Inf. Theory*, vol. 56, no. 4, pp. 1982–2001, Apr. 2010.

[26] P. Feng and Y. Bresler, "Spectrum-blind minimum-rate sampling and reconstruction of multiband signals," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, May 1996, pp. 1688–1691.

[27] M. Mishali and Y. C. Eldar, "Blind multiband signal reconstruction: Compressed sensing for analog signals," *IEEE Trans. Signal Process.*, vol. 57, no. 3, pp. 993–1009, Mar. 2009.

[28] R. G. Baraniuk and M. B. Wakin, "Random projections of smooth manifolds," *Found. Comput. Math.*, vol. 9, no. 1, pp. 51–77, Feb 2009.

[29] M. Stojnic, F. Parvaresh, and B. Hassibi, "On the reconstruction of block-sparse signals with an optimal number of measurements," *IEEE Trans. Signal Process.*, vol. 57, no. 8, pp. 3075–3085, Aug. 2009.

[30] Y. C. Eldar, P. Kuppinger, and H. Bölcskei, "Block-sparse signals: Uncertainty relations and efficient recovery," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3042–3054, Jun. 2010.

[31] J. Huang and T. Zhang, "The benefit of group sparsity," *Ann. Stat.*, vol. 38, no. 4, pp. 1978–2004, 2010.

[32] E. J. Candès and B. Recht, "Exact matrix completion via convex optimization," *Found. Comput. Math.*, vol. 9, no. 6, pp. 717–772, 2009.

[33] E. J. Candès and Y. Plan, "Tight oracle inequalities for low-rank matrix recovery from a minimal number of noisy random measurements," *IEEE Trans. Inf. Theory*, vol. 4, no. 57, pp. 2342–2359, Apr. 2011.

[34] E. Riegler, D. Stotz, and H. Bölcskei, "Information-theoretic limits of matrix completion," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2015, pp. 1836–1840.

[35] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1991.

[36] S. G. Krantz and H. R. Parks, *Geometric Integration Theory*. Princeton, NJ: Birkhäuser, 2008.

[37] J. C. Robinson, *Dimensions, Embeddings, and Attractors*. Cambridge, NY: Cambridge University Press, 2011.

[38] K. Falconer, *Fractal Geometry*, 3rd ed. New York, NY: Wiley, 2014.

[39] T. Kawabata and A. Dembo, "The rate-distortion dimension of sets and measures," *IEEE Trans. Inf. Theory*, vol. 40, no. 5, pp. 1564–1572, Sep. 1994.

[40] H. Federer, *Geometric Measure Theory*. New York, NY: Springer, 1969.

[41] K. Stromberg, "An elementary proof of Steinhaus's theorem," in *Proc. of the American Mathematical Society*, vol. 36, no. 1, Nov. 1972, p. 308.

[42] W. Rudin, *Principles of Mathematical Analysis*, 3rd ed. New York, NY: McGraw-Hill, 1989.

[43] J. M. Lee, *Introduction to Smooth Manifolds*, 2nd ed. New York, NY: Springer, 2013.

[44] G. Folland, *Real Analysis: Modern Techniques and Their Applications*, 2nd ed. New York, NY: Wiley, 2013.

[45] H. H. Sohrab, *Basic Real Analysis*, 2nd ed. New York, NY: Birkhäuser, 2014.

[46] P. Mattila, *Geometry of Sets and Measures in Euclidean Space: Fractals and Rectifiability*. Cambridge, UK: Cambridge Univ. Press, 1995.

[47] L. Ambrosio, N. Fusco, and D. Pallara, *Functions of Bounded Variation and Free Discontinuity Problems*. New York, NY: Oxford Univ. Press, 2000.

[48] L. C. Evans and R. F. Gariepy, *Measure Theory and Finite Properties of Functions*, revised ed. Boca Raton, FL: CRC Press, 2015.

[49] H. Federer, "Colloquium lectures on geometric measure theory," *Bull. Amer. Math. Soc.*, vol. 84, no. 3, pp. 291–338, May 1978.

[50] R. G. Bartle, *The Elements of Integration and Lebesgue Measure*. New York, NY: Wiley, 1995.

[51] T. Tao, *An Introduction to Measure Theory*. Providence RI: American Mathematical Society, 2011.

[52] L. Gasiński and N. S. Papageorgiou, *Exercises in Analysis-Part 2: Nonlinear Analysis*, ser. Problem Books in Mathematics. Cham, Switzerland: Springer, 2016.

[53] T.-W. Ma, *Classical Analysis on Normed Spaces*. Singapore: World Scientific Pub. Co., 1995.

[54] F. W. Warner, *Foundations of Differentiable Manifolds and Lie Groups*. New York, NY: Springer, 1971.

[55] C. H. Edwards, *Advanced Calculus of Several Variables*. New York, NY: Academic Press, Inc., 1973.

[56] A. Sard, "The measure of the critical values of differentiable maps," *Bull. Amer. Math. Soc.*, vol. 48, no. 12, pp. 883–890, Dec. 1942.

[57] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, 3rd ed. Heidelberg, Germany: Springer, 2009.

[58] S. G. Krantz and H. R. Parks, *A Primer of Real Analytic Functions*, 2nd ed. Basel, Switzerland: Birkhäuser, 2002.

[59] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2013.