ADDENDUM to "Almost Lossless Analog Signal Separation and Probabilistic Uncertainty Relations"

David Stotz, Erwin Riegler, Eirikur Agustsson, and Helmut Bölcskei
ETH Zurich, Switzerland

I. INTRODUCTION

This document is a supplement to [1]. It provides complete proofs of auxiliary results in [1], which are minor extensions of results available in the literature.

Notation: All notation conventions are adopted from [1]. For $\boldsymbol{x} \in \mathbb{R}^n$ and $\mathcal{T} \subseteq \{1, \ldots, n\}$, we let $\boldsymbol{x}_\mathcal{T}$ denote the $|\mathcal{T}|$-dimensional subvector that consists of the components of $\boldsymbol{x}$ corresponding to the indices in $\mathcal{T}$.

II. THE MINKOWSKI DIMENSION COMPRESSION RATE OF MIXED DISCRETE-CONTINUOUS SOURCES

We begin by restating the result to be proved, namely [1, Proposition 3].

Proposition 3: Suppose that $\mathbf{x}$ is distributed according to [1, Definition 6]. Then, we have

$$R_\mathrm{B}(\varepsilon) = (1-\lambda)\rho_1 + \lambda\rho_2, \tag{1}$$

for all $\varepsilon \in (0,1)$.

The proof of Proposition 3 provided below follows by adapting the arguments in the proof of [2, Thm. 15] to the signal separation setting. We start by stating an auxiliary lemma from [3, Thm. 4.16], whose short proof is included for completeness.

Lemma 1: Every non-empty bounded set $\mathcal{A} \subseteq \mathbb{R}^n$ with $\underline{\dim}_\mathrm{B}(\mathcal{A}) < n$ has Lebesgue measure zero.

Proof: Suppose that $\operatorname{Leb}^n(\mathcal{A}) > 0$. We then get the contradiction

$$\begin{aligned} 0 &= \liminf_{\delta\to 0} \frac{\log \operatorname{Leb}^n(\mathcal{A})}{\log \frac{1}{\delta}} && (2)\\ &\leqslant \liminf_{\delta\to 0} \frac{\log\big(N_{\mathcal{A}}(\delta)\alpha(n,\delta)\big)}{\log \frac{1}{\delta}} && (3)\\ &= \liminf_{\delta\to 0} \frac{\log\big(N_{\mathcal{A}}(\delta)C(n)\delta^n\big)}{\log \frac{1}{\delta}} && (4)\\ &= \underline{\dim}_{\mathrm{B}}(\mathcal{A}) - n && (5)\\ &< 0, && (6)\end{aligned}$$

where (2) is by $\operatorname{Leb}^n(\mathcal{A}) > 0$, (3) follows by covering $\mathcal{A}$ with $N_{\mathcal{A}}(\delta)$ balls of radius δ where each ball has volume $\alpha(n,\delta)$, (4) is by $\alpha(n,\delta) = \delta^n\alpha(n,1) = \delta^n C(n)$, (5) holds by definition of lower Minkowski dimension [1, (3)], and (6) is by assumption.

■

Proof of Proposition 3: We begin with preparatory steps. Recall the role of the parameter λ in [1, Definition 1] and the definition of concatenated source vectors x of mixed discrete-continuous distribution in [1, Definition 6]. The cases $\lambda = 0$ and $\lambda = 1$ are equivalent to the case $\lambda = 1/2$, $\rho_1 = \rho_2$, $\mu_{\mathrm{d}_1} = \mu_{\mathrm{d}_2}$, and $\mu_{\mathrm{c}_1} = \mu_{\mathrm{c}_2}$. We can therefore assume, without loss of generality, that $0 < \lambda < 1$, and take $0 \dot{<} \ell = \lfloor \lambda n \rfloor \dot{<} n$.

Let $\mathcal{A}_i \subseteq \mathbb{R}$ be the set of atoms of μ_{d_i}, i.e., values in $\mathcal{A}_i^c$ can only stem from the absolutely continuous part μ_{c_i}. Since $\mu_{\mathrm{d}_i}(\mathcal{A}_i^c) = 0$, we have

$$\begin{aligned}\mathbb{E}[\mathbb{1}_{\mathcal{A}_i^c}(\mathsf{X}_j)] &= \mu_{\mathsf{X}_j}(\mathcal{A}_i^c)\\ &= \begin{cases}\rho_1, & \text{for } i=1,\ j\in\{1,\dots,n-\ell\}\\ \rho_2, & \text{for } i=2,\ j\in\{n-\ell+1,\dots,n\}.\end{cases}\end{aligned}$$

By the weak law of large numbers, we get for $n\to\infty$

$$\frac{1}{n-\ell}\sum_{j=1}^{n-\ell} \mathbb{1}_{\mathcal{A}_1^c}(\mathsf{X}_j) \xrightarrow{\mathbb{P}} \rho_1 \tag{7}$$

$$\frac{1}{\ell}\sum_{j=n-\ell+1}^{n} \mathbb{1}_{\mathcal{A}_2^c}(\mathsf{X}_j) \xrightarrow{\mathbb{P}} \rho_2. \tag{8}$$

The assertion to be proved says that the Minkowski dimension compression rate is given by the average number of entries in x that are drawn according to the absolutely continuous parts μ_{c_i}. We next define the generalized support of a vector $\boldsymbol{x}\in\mathbb{R}^n$ as

$$\operatorname{spt}(\boldsymbol{x}) := \{i \in \{1,\dots,n-\ell\} \mid x_i \in \mathcal{A}_1^c\} \cup \{i\in\{n-\ell+1,\dots,n\} \mid x_i \in \mathcal{A}_2^c\},$$

i.e., as the set of indices with the corresponding entries drawn from the absolutely continuous parts μ_{c_i}. By (7), (8) we have

$$\frac{|\operatorname{spt}(\mathsf{x})|}{n} = \frac{n-\ell}{n}\frac{1}{n-\ell}\sum_{j=1}^{n-\ell}\mathbb{1}_{\mathcal{A}_1^c}(\mathsf{X}_j) + \frac{\ell}{n}\frac{1}{\ell}\sum_{j=n-\ell+1}^{n}\mathbb{1}_{\mathcal{A}_2^c}(\mathsf{X}_j) \xrightarrow{\mathbb{P}} (1-\lambda)\rho_1 + \lambda\rho_2, \tag{9}$$

where we used $\ell/n = \lfloor\lambda n\rfloor/n \xrightarrow{n\to\infty} \lambda$ as a consequence of $\lambda n - 1 < \lfloor\lambda n\rfloor \leqslant \lambda n$, and similarly $(n-\ell)/n = (n-\lfloor\lambda n\rfloor)/n \xrightarrow{n\to\infty} (1-\lambda)$ which follows by $(1-\lambda)n \leqslant n - \lfloor\lambda n\rfloor < (1-\lambda)n+1$.

The proof strategy is to establish that

$$\overline{R}_{\mathrm{B}}(\varepsilon) \leqslant (1-\lambda)\rho_1 + \lambda\rho_2 \leqslant \underline{R}_{\mathrm{B}}(\varepsilon), \tag{10}$$

for all $\varepsilon \in (0,1)$, which, owing to $\underline{R}_{\mathrm{B}}(\varepsilon) \leqslant \overline{R}_{\mathrm{B}}(\varepsilon)$, implies $\overline{R}_{\mathrm{B}}(\varepsilon) = \underline{R}_{\mathrm{B}}(\varepsilon) = R_{\mathrm{B}}(\varepsilon) = (1-\lambda)\rho_1 + \lambda\rho_2$ and hence finishes the proof. The main idea for establishing (10) is to consider sets of realizations that have certain entries fixed to atoms of the discrete parts μ_{d_i} and the remaining entries drawn from the absolutely continuous parts μ_{c_i}. We begin by establishing the left-hand inequality in (10). To this end, we construct an approximate support set $\mathcal{S}$ for x, i.e., we find an $\mathcal{S}$ such that $\mathbb{P}[\mathsf{x} \in \mathcal{S}] \geqslant 1-\varepsilon$, whose Minkowski dimension is smaller than $((1-\lambda)\rho_1 + \lambda\rho_2 + \kappa)n$, for $\kappa > 0$. First, note that by convergence in probability in (9), we get

$$\mathbb{P}\left[\left|\frac{|\operatorname{spt}(\mathsf{x})|}{n} - ((1-\lambda)\rho_1 + \lambda\rho_2)\right| < \kappa\right] \xrightarrow{n\to\infty} 1, \tag{11}$$

for all $\kappa > 0$. Setting

$$\mathcal{C} := \{\boldsymbol{x} \in \mathbb{R}^n \mid |\operatorname{spt}(\boldsymbol{x})| < ((1-\lambda)\rho_1 + \lambda\rho_2 + \kappa)n\}, \tag{12}$$

it follows that $\mathbb{P}[\mathsf{x} \in \mathcal{C}] \geqslant \mathbb{P}\left[\left|\frac{|\operatorname{spt}(\mathsf{x})|}{n} - ((1-\lambda)\rho_1 + \lambda\rho_2)\right| < \kappa\right]$, which, in turn, implies

$$\mathbb{P}[\mathsf{x} \in \mathcal{C}] \stackrel{\cdot}{\geqslant} 1 - \frac{\varepsilon}{2} \tag{13}$$

(the choice of the constant $\varepsilon/2$ in (13) is motivated by the desire to get $\mathbb{P}[\mathsf{x} \in \mathcal{S}] \stackrel{\cdot}{\geqslant} 1-\varepsilon$ in (16) below). In what follows $\mathcal{T}_1$ denotes a subset of $\{1,\dots,n-\ell\}$ and $\mathcal{T}_2$ a subset of $\{n-\ell+1,\dots,n\}$. Next, we formalize the idea of decomposing $\mathcal{C}$ into sets of possible realizations of the source that have certain entries fixed to atoms $\boldsymbol{z}_i \in \mathcal{A}_i$ of the μ_{d_i} and the remaining entries drawn from the absolutely continuous parts μ_{c_i}. Specifically, we decompose $\mathcal{C}$ according to

$$\mathcal{C} = \bigcup_{|\mathcal{T}_1|+|\mathcal{T}_2|<((1-\lambda)\rho_1+\lambda\rho_2+\kappa)n} \mathcal{U}_{\mathcal{T}_1,\mathcal{T}_2} \tag{14}$$

with

$$\mathcal{U}_{\mathcal{T}_1,\mathcal{T}_2} = \bigcup_{\boldsymbol{z}_1 \in \mathcal{A}_1^{n-\ell-|\mathcal{T}_1|}} \bigcup_{\boldsymbol{z}_2 \in \mathcal{A}_2^{\ell-|\mathcal{T}_2|}} \mathcal{U}_{\mathcal{T}_1,\mathcal{T}_2,\boldsymbol{z}_1,\boldsymbol{z}_2}$$

$$\mathcal{U}_{\mathcal{T}_1,\mathcal{T}_2,\boldsymbol{z}_1,\boldsymbol{z}_2} = \{\boldsymbol{x} \in \mathbb{R}^n \mid \operatorname{spt}(\boldsymbol{x}) = \mathcal{T}_1 \cup \mathcal{T}_2,\, \boldsymbol{x}_{\mathcal{T}_1^c} = \boldsymbol{z}_1, \boldsymbol{x}_{\mathcal{T}_2^c} = \boldsymbol{z}_2\}.$$

Here, the set $\mathcal{U}_{\mathcal{T}_1,\mathcal{T}_2,\boldsymbol{z}_1,\boldsymbol{z}_2}$ consists of vectors whose entries corresponding to indices in $\mathcal{T}_1 \cup \mathcal{T}_2$ result from the absolutely continuous parts μ_{c_i} and the remaining entries are fixed to atoms of the discrete parts μ_{d_i}. Note that $\mathcal{U}_{\mathcal{T}_1,\mathcal{T}_2,\boldsymbol{z}_1,\boldsymbol{z}_2}$ is contained in the $(|\mathcal{T}_1|+|\mathcal{T}_2|)$-dimensional affine space $\{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{x}_{\mathcal{T}_1^c} = \boldsymbol{z}_1,\ \boldsymbol{x}_{\mathcal{T}_2^c} = \boldsymbol{z}_2\}$, for $\boldsymbol{z}_1 \in \mathcal{A}_1^{n-\ell-|\mathcal{T}_1|}, \boldsymbol{z}_2 \in \mathcal{A}_2^{\ell-|\mathcal{T}_2|}$. Since the collection of all sets $\mathcal{U}_{\mathcal{T}_1,\mathcal{T}_2,\boldsymbol{z}_1,\boldsymbol{z}_2}$ is countable, we can relabel them as $\{\mathcal{U}_j \mid j \in \mathbb{N}\}$ and rewrite (14) according to

$$\mathcal{C} = \bigcup_{j\in\mathbb{N}} \mathcal{U}_j. \tag{15}$$

There exists a $J \in \mathbb{N}$ and an $r > 0$ such that for $\mathcal{S} := B^n(\mathbf{0}, r) \cap \bigcup_{j=1}^{J} \mathcal{U}_j$ we have

$$\mathbb{P}[\mathbf{x} \in \mathcal{S}] \overset{\cdot}{\geqslant} 1 - \varepsilon, \tag{16}$$

since

$$\mathbb{P}[\mathbf{x} \in \mathcal{S}] \xrightarrow{J,r\to\infty} \mathbb{P}[\mathbf{x} \in \mathcal{C}] \overset{\cdot}{\geqslant} 1 - \frac{\varepsilon}{2},$$

by (13). Now

$$\overline{\dim}_{\mathrm{B}}(\mathcal{S}) = \max_{j\in\{1,\dots,J\}} \overline{\dim}_{\mathrm{B}}(B^n(\mathbf{0}, r) \cap \mathcal{U}_j) \tag{17}$$

$$\leqslant ((1-\lambda)\rho_1 + \lambda\rho_2 + \kappa)n, \tag{18}$$

where (17) follows as upper Minkowski dimension is finitely stable [4, Sec. 3.2, (iii)], i.e., $\overline{\dim}_{\mathrm{B}}(\mathcal{A}\cup\mathcal{B}) = \max\{\overline{\dim}_{\mathrm{B}}(\mathcal{A}), \overline{\dim}_{\mathrm{B}}(\mathcal{B})\}$ and in (18) we use a fact established next, namely $\overline{\dim}_{\mathrm{B}}(B^n(\mathbf{0}, r)\cap\mathcal{U}_j) \leqslant |\mathcal{T}_1| + |\mathcal{T}_2|$, where $\mathcal{T}_1$ and $\mathcal{T}_2$ correspond to $\mathcal{U}_j$. First, note that $B^n(\mathbf{0}, r)\cap\mathcal{U}_j$ is a bounded subset of a $(|\mathcal{T}_1|+|\mathcal{T}_2|)$-dimensional affine subspace, which, in particular, is a $(|\mathcal{T}_1| + |\mathcal{T}_2|)$-dimensional smooth manifold [**?**, Ex. 1.24]. Then, apply [4, Sec. 3.2, (i)] to conclude that a bounded subset of a smooth m-dimensional manifold has Minkowski dimension smaller than m, and finally use $|\mathcal{T}_1|+|\mathcal{T}_2| < ((1-\lambda)\rho_1+\lambda\rho_2+\kappa)n$. Combining (16) with (18), we obtain $\overline{R}_{\mathrm{B}}(\varepsilon) \leqslant \lambda\rho_1 + (1-\lambda)\rho_2 + \kappa$ and since κ was arbitrary this proves that

$$\overline{R}_{\mathrm{B}}(\varepsilon) \leqslant \lambda\rho_1 + (1-\lambda)\rho_2. \tag{19}$$

We proceed to prove the right-hand inequality in (10). First, we set

$$\mathcal{D} := \{\boldsymbol{x} \in \mathbb{R}^n \mid |\operatorname{spt}(\boldsymbol{x})| > ((1-\lambda)\rho_1 + \lambda\rho_2 - \kappa)n\}. \tag{20}$$

Next, we note that $\mathbb{P}[\mathbf{x} \in \mathcal{D}] \geqslant \mathbb{P}\left[\left|\frac{|\operatorname{spt}(\mathbf{x})|}{n} - ((1-\lambda)\rho_1 + \lambda\rho_2)\right| < \kappa\right]$, and by (11), we have, for $\varepsilon \in (0, 1)$,

$$\mathbb{P}[\mathbf{x} \in \mathcal{D}] \overset{\cdot}{\geqslant} \frac{1+\varepsilon}{2}. \tag{21}$$

The choice of the constant $(1+\varepsilon)/2 < 1$ is motivated by the desire to get a positive lower bound in (25) below. Then, for a non-empty bounded set $\mathcal{K}$ with $\mathbb{P}[\mathbf{x} \in \mathcal{K}] \dot{\geqslant} 1-\varepsilon$, we have

$$\mathbb{P}[\mathbf{x} \in (\mathcal{K} \cap \mathcal{D})] = 1 - \mathbb{P}[\mathbf{x} \in (\mathcal{K}^c \cup \mathcal{D})^c] \tag{22}$$

$$\geqslant 1 - \mathbb{P}[\mathbf{x} \in \mathcal{K}^c] - \mathbb{P}[\mathbf{x} \in \mathcal{D}^c] \tag{23}$$

$$\dot{\geqslant} 1 - \varepsilon - \left(1 - \frac{1+\varepsilon}{2}\right) \tag{24}$$

$$= (1-\varepsilon)/2. \tag{25}$$

Moreover, we can write $\mathcal{K} \cap \mathcal{D}$ as the union

$$\mathcal{K} \cap \mathcal{D} = \bigcup_{|\mathcal{T}_1|+|\mathcal{T}_2|>((1-\lambda)\rho_1+\lambda\rho_2-\kappa)n} \mathcal{V}_{\mathcal{T}_1,\mathcal{T}_2} \tag{26}$$

with

$$\mathcal{V}_{\mathcal{T}_1,\mathcal{T}_2} = \bigcup_{\boldsymbol{z}_1 \in \mathcal{A}_1^{n-\ell-|\mathcal{T}_1|}} \bigcup_{\boldsymbol{z}_2 \in \mathcal{A}_2^{\ell-|\mathcal{T}_2|}} \mathcal{V}_{\mathcal{T}_1,\mathcal{T}_2,\boldsymbol{z}_1,\boldsymbol{z}_2}$$

$$\mathcal{V}_{\mathcal{T}_1,\mathcal{T}_2,\boldsymbol{z}_1,\boldsymbol{z}_2} = \{\boldsymbol{x} \in \mathcal{K} \mid \operatorname{spt}(\boldsymbol{x}) = \mathcal{T}_1 \cup \mathcal{T}_2,\ \boldsymbol{x}_{\mathcal{T}_1^c} = \boldsymbol{z}_1,\ \boldsymbol{x}_{\mathcal{T}_2^c} = \boldsymbol{z}_2\}.$$

Now, owing to $\mathbb{P}[\mathbf{x} \in (\mathcal{K} \cap \mathcal{D})] \dot{\geqslant} (1-\varepsilon)/2 > 0$, there exists at least one set $\mathcal{V}_{\mathcal{T}_1,\mathcal{T}_2,\boldsymbol{z}_1,\boldsymbol{z}_2}$ with $|\mathcal{T}_1|+|\mathcal{T}_2| > ((1-\lambda)\rho_1 + \lambda\rho_2 - \kappa)n$ and

$$\mu_{\mathbf{x}}(\mathcal{V}_{\mathcal{T}_1,\mathcal{T}_2,\boldsymbol{z}_1,\boldsymbol{z}_2}) \dot{>} 0, \tag{27}$$

which for this particular set $\mathcal{V}_{\mathcal{T}_1,\mathcal{T}_2,\boldsymbol{z}_1,\boldsymbol{z}_2}$, in turn, implies that

$$\underline{\dim}_{\mathrm{B}}(\mathcal{K}) \geqslant \underline{\dim}_{\mathrm{B}}(\mathcal{K} \cap \mathcal{D}) \tag{28}$$

$$\geqslant \underline{\dim}_{\mathrm{B}}(\mathcal{V}_{\mathcal{T}_1,\mathcal{T}_2,\boldsymbol{z}_1,\boldsymbol{z}_2}) \tag{29}$$

$$= \underline{\dim}_{\mathrm{B}}(\{\boldsymbol{x}_{\mathcal{T}_1 \cup \mathcal{T}_2} \mid \boldsymbol{x} \in \mathcal{V}_{\mathcal{T}_1,\mathcal{T}_2,\boldsymbol{z}_1,\boldsymbol{z}_2}\}) \tag{30}$$

$$\dot{>} ((1-\lambda)\rho_1 + \lambda\rho_2 - \kappa)n, \tag{31}$$

where (28) and (29) follow from the monotonicity of lower Minkowski dimension [4, Sec. 3.2, (ii)], i.e., $\underline{\dim}_{\mathrm{B}}(\mathcal{A}) \leqslant \underline{\dim}_{\mathrm{B}}(\mathcal{B})$ for $\mathcal{A} \subseteq \mathcal{B}$, (30) holds by the definition of $\mathcal{V}_{\mathcal{T}_1,\mathcal{T}_2,\boldsymbol{z}_1,\boldsymbol{z}_2}$, and (31) follows by application of Lemma 1, since $\mu_{\mathbf{x}}(\mathcal{V}_{\mathcal{T}_1,\mathcal{T}_2,\boldsymbol{z}_1,\boldsymbol{z}_2}) \dot{>} 0$ implies that the set $\{\boldsymbol{x}_{\mathcal{T}_1 \cup \mathcal{T}_2} \mid \boldsymbol{x} \in \mathcal{V}_{\mathcal{T}_1,\mathcal{T}_2,\boldsymbol{z}_1,\boldsymbol{z}_2}\} \subseteq \mathbb{R}^{|\mathcal{T}_1|+|\mathcal{T}_2|}$ has positive Lebesgue measure, and we used $|\mathcal{T}_1| + |\mathcal{T}_2| > ((1-\lambda)\rho_1 + \lambda\rho_2 - \kappa)n$. Letting $\kappa \to 0$ we therefore established that $\underline{\dim}_{\mathrm{B}}(\mathcal{K})/n \dot{\geqslant} (1-\lambda)\rho_1 + \lambda\rho_2$ for all bounded sets $\mathcal{K}$ satisfying $\mathbb{P}[\mathbf{x} \in \mathcal{K}] \dot{\geqslant} 1-\varepsilon$, which implies

$$\underline{R}_{\mathrm{B}}(\varepsilon) \geqslant (1-\lambda)\rho_1 + \lambda\rho_2, \tag{32}$$

and thereby finishes the proof. ■

III. Proof of Converse for mixed discrete-continuous sources

We begin by restating the result to be proved, namely [1, Proposition 4]. The proof follows by adapting the converse part of [2, Thm. 6] to our setting.

Proposition 4: Suppose that x is distributed according to [1, Definition 6] and let $\varepsilon \in (0,1)$. Then, the existence of a measurement matrix $\boldsymbol{H} = [\boldsymbol{A}\ \boldsymbol{B}] : \mathbb{R}^{n-\ell} \times \mathbb{R}^{\ell} \to \mathbb{R}^k$ and a corresponding measurable separator $g\colon \mathbb{R}^k \to \mathbb{R}^{n-\ell} \times \mathbb{R}^{\ell}$, with $k = \lfloor Rn \rfloor$, such that

$$\mathbb{P}[g([\boldsymbol{A}\ \boldsymbol{B}]\mathsf{x}) \neq \mathsf{x}] \mathrel{\dot{\leqslant}} \varepsilon, \tag{33}$$

imply $R \geqslant R_{\mathrm{B}}(\varepsilon)$.

Proof: We have to show that for all $\varepsilon \in (0,1)$, all $\boldsymbol{H} = [\boldsymbol{A}\ \boldsymbol{B}]$ and corresponding measurable g, (33) can hold only if

$$R \geqslant R_{\mathrm{B}}(\varepsilon). \tag{34}$$

We can assume, without loss of generality, that $(\rho_1, \rho_2) \neq (0,0)$, as for $(\rho_1, \rho_2) = (0,0)$ we have $R_{\mathrm{B}}(\varepsilon) = 0$ by Proposition 3 and $R \geqslant 0$ by definition. Fix $\kappa > 0$ such that $\kappa < R_{\mathrm{B}}(\varepsilon)/2$. Suppose, by way of contradiction, that the rate $R = R_{\mathrm{B}}(\varepsilon) - 2\kappa$ is achievable, i.e., there exists a measurement matrix $[\boldsymbol{A}\ \boldsymbol{B}]$ and a corresponding separator g achieving rate R with error probability ε for some $0 < \varepsilon < 1$. With $k = \lfloor Rn \rfloor$ and setting $k' = \lfloor (R_{\mathrm{B}}(\varepsilon) - \kappa)n \rfloor$ we have

$$k \mathrel{\dot{<}} k'. \tag{35}$$

Since R is achievable with error probability ε, it follows from [1, Definition 2] that for sufficiently large n there exists a Borel set $\mathcal{K}$, namely the set of realizations of $\mathsf{x} = [\mathsf{y}^{\mathrm{T}}\ \mathsf{z}^{\mathrm{T}}]^{\mathrm{T}}$ that is successfully separated, on which $[\boldsymbol{A}\ \boldsymbol{B}]$ is one-to-one and which moreover satisfies

$$\mathbb{P}[\mathsf{x} \in \mathcal{K}] \mathrel{\dot{\geqslant}} 1 - \varepsilon. \tag{36}$$

The mapping $[\boldsymbol{A}\ \boldsymbol{B}]$ being one-to-one on $\mathcal{K}$ for sufficiently large n is equivalent to

$$\ker([\boldsymbol{A}\ \boldsymbol{B}]) \cap (\mathcal{K} \ominus \mathcal{K}) \doteq \{\mathbf{0}\}. \tag{37}$$

The proof will be effected by showing that (37) leads to a contradiction. Repeating steps (22)–(27) in the proof of Proposition 3, and using $R_{\mathrm{B}}(\varepsilon) = (1-\lambda)\rho_1 + \lambda\rho_2$ by Proposition 3, we see that there exist support sets $\mathcal{T}_1 \subseteq \{1, \dots, n-\ell\}$ and $\mathcal{T}_2 \subseteq \{n-\ell+1, \dots, n\}$ and corresponding vectors $\boldsymbol{z}_1 \in \mathcal{A}_1^{n-\ell-|\mathcal{T}_1|}$ and $\boldsymbol{z}_2 \in \mathcal{A}_2^{\ell-|\mathcal{T}_2|}$, such that $|\mathcal{T}_1| + |\mathcal{T}_2| > (R_{\mathrm{B}}(\varepsilon) - \kappa)n \geqslant k'$ and the corresponding set

$$\mathcal{V}_{\mathcal{T}_1,\mathcal{T}_2,\boldsymbol{z}_1,\boldsymbol{z}_2} = \{\boldsymbol{x} \in \mathcal{K} \mid \operatorname{spt}(\boldsymbol{x}) = \mathcal{T}_1 \cup \mathcal{T}_2,\ \boldsymbol{x}_{\mathcal{T}_1^c} = \boldsymbol{z}_1,\ \boldsymbol{x}_{\mathcal{T}_2^c} = \boldsymbol{z}_2\} \tag{38}$$

satisfies $\mu_{\mathsf{x}}(\mathcal{V}_{\mathcal{T}_1,\mathcal{T}_2,\boldsymbol{z}_1,\boldsymbol{z}_2}) \dot{>} 0$. Now, let $\mathcal{F} := \{\boldsymbol{x}_{\mathcal{T}_1\cup\mathcal{T}_2} \mid \boldsymbol{x} \in \mathcal{V}_{\mathcal{T}_1,\mathcal{T}_2,\boldsymbol{z}_1,\boldsymbol{z}_2}\}$. Since the discrete parts μ_{d_i} do not contribute to $\mathcal{F}$, it follows that $\mu_{\mathsf{x}}(\mathcal{V}_{\mathcal{T}_1,\mathcal{T}_2,\boldsymbol{z}_1,\boldsymbol{z}_2}) \dot{>} 0$ is possible only if $\mathrm{Leb}^{|\mathcal{T}_1|+|\mathcal{T}_2|}(\mathcal{F}) \dot{>} 0$, and, therefore, by the Steinhaus Theorem [5], there exists a ball $B^{|\mathcal{T}_1|+|\mathcal{T}_2|}(\mathbf{0}, r)$ with radius $r > 0$ such that $B^{|\mathcal{T}_1|+|\mathcal{T}_2|}(\mathbf{0}, r) \dot{\subseteq} \mathcal{F} \ominus \mathcal{F}$. Hence, with $\mathcal{B} := \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{x}_{\mathcal{T}_1\cup\mathcal{T}_2} \in B^{|\mathcal{T}_1|+|\mathcal{T}_2|}(\mathbf{0}, r), \boldsymbol{x}_{(\mathcal{T}_1\cup\mathcal{T}_2)^c} = \mathbf{0}\}$, we have $\mathcal{B} \dot{\subseteq} \mathcal{V}_{\mathcal{T}_1,\mathcal{T}_2,\boldsymbol{z}_1,\boldsymbol{z}_2} \ominus \mathcal{V}_{\mathcal{T}_1,\mathcal{T}_2,\boldsymbol{z}_1,\boldsymbol{z}_2}$ and, as $\mathcal{V}_{\mathcal{T}_1,\mathcal{T}_2,\boldsymbol{z}_1,\boldsymbol{z}_2} \subseteq \mathcal{K}$ by (38), it follows that $\mathcal{B} \dot{\subseteq} \mathcal{K} \ominus \mathcal{K}$. The span of the ball $B^{|\mathcal{T}_1|+|\mathcal{T}_2|}(\mathbf{0}, r)$ is a $(|\mathcal{T}_1|+|\mathcal{T}_2|)$-dimensional vector space, which together with $|\mathcal{T}_1|+|\mathcal{T}_2| > k'$ implies that there exists a set of linearly independent vectors

$$\{\boldsymbol{f}_1, \dots, \boldsymbol{f}_{k'}\} \subseteq \mathcal{B} \dot{\subseteq} \mathcal{K} \ominus \mathcal{K}. \tag{39}$$

Moreover, since $\dim(\ker([\boldsymbol{A}\ \ \boldsymbol{B}])) \geqslant n - k = n - \lfloor Rn \rfloor$, there exists a set of linearly independent vectors

$$\{\boldsymbol{g}_1, \dots, \boldsymbol{g}_{n-k}\} \subseteq \ker([\boldsymbol{A}\ \ \boldsymbol{B}]). \tag{40}$$

Since $k \dot{<} k'$ by (35), it follows that the union of the two sets on the left-hand sides in (39) and (40) is linearly dependent when n is sufficiently large. This, in turn, implies that there exists a non-zero vector $\boldsymbol{v}$ such that

$$\boldsymbol{v} = \alpha_1 \boldsymbol{f}_1 + \dots \alpha_{k'} \boldsymbol{f}_{k'} = \beta_1 \boldsymbol{g}_1 + \dots + \beta_{n-k} \boldsymbol{g}_{n-k}, \tag{41}$$

with $\alpha_1, ..., \alpha_{k'}, \beta_1, ..., \beta_{n-k} \in \mathbb{R}$. Noting that we can multiply (41) by an arbitrary constant, we may assume that $\boldsymbol{v} \in \mathcal{B}$. Furthermore, by (40) we have $\boldsymbol{v} \in \ker([\boldsymbol{A}\ \ \boldsymbol{B}])$. We can therefore conclude that $\mathcal{B} \cap \ker([\boldsymbol{A}\ \ \boldsymbol{B}]) \dot{\neq} \{\mathbf{0}\}$, which contradicts (37) as $\mathcal{B} \dot{\subseteq} \mathcal{K} \ominus \mathcal{K}$. Since κ can be chosen arbitrarily small, this proves that the existence of a measurement matrix $[\boldsymbol{A}\ \ \boldsymbol{B}]$ and a corresponding separator g achieving error probability ε for some $0 < \varepsilon < 1$ necessarily implies $R \geqslant R_{\mathrm{B}}(\varepsilon)$ and thus completes the proof. ■

References

[1] D. Stotz, E. Riegler, E. Agustsson, and H. Bölcskei, "Almost lossless analog signal separation and probabilistic uncertainty relations," *to appear in IEEE Trans. Inf. Theory*, 2017.

[2] 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.

[3] K. T. Alligood, T. D. Sauer, and J. A. Yorke, *Chaos: An introduction to dynamical systems*, ser. Textbooks in Mathematical Sciences. Springer New York, 1996.

[4] K. Falconer, *Fractal Geometry: Mathematical Foundations and Applications*, 2nd ed. John Wiley & Sons, 2004.

[5] K. Stromberg, "An elementary proof of Steinhaus's theorem," *Proc. of the American Mathematical Society*, vol. 36, no. 1, p. 308, Nov. 1976.