ADDENDUM to "Degrees of Freedom in Vector Interference Channels"

David Stotz and Helmut Bölcskei

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 or restatements of results that appear in the literature without proof.

Notation: All notation conventions are adopted from [1].

II. Proofs of auxiliary results in [1, Appendix B]

The following lemma, which is a straightforward extension of [2, Lem. 4] to the vector case, provides a sufficient condition for the iterated function system used in the construction of the random vector in [1, (22)] to satisfy the open set condition.

Lemma 1. *Consider the iterated function system* $\{F_1, ..., F_m\}$ *with* $F_i(x) = rx + w_i$*, for* $x \in \mathbb{R}^n$, $r \in (0,1)$*, and pairwise different vectors* $w_1, ..., w_m \in \mathbb{R}^n$*. Let furthermore* $\mathcal{W} := \{w_1, ..., w_m\}$*. Then, the open set condition (see [1, Definition 2]) is satisfied if*

$$r \leqslant \frac{\mathsf{m}(\mathcal{W})}{\mathsf{m}(\mathcal{W}) + \mathsf{M}(\mathcal{W})}. \tag{1}$$

Proof: The idea is to construct a bounded open set $\mathcal{U}$ such that under (1) the images of every point in this set under different contractions lie sufficiently far apart for $F_i(\mathcal{U}) \cap F_j(\mathcal{U}) = \emptyset$ to hold,

The authors are with the Department of Information Technology and Electrical Engineering, ETH Zurich, CH-8092 Zurich, Switzerland (e-mail: {dstotz,boelcskei}@nari.ee.ethz.ch).

for all $i \neq j$, and to moreover ensure that $\bigcup_{i=1}^{m} F_i(\mathcal{U}) \subseteq \mathcal{U}$. Let $U := (a_1, b_1) \times ... \times (a_n, b_n)$, where $a_k := \frac{\min_i w_{i,k}}{1-r}$ and $b_k := \frac{\max_i w_{i,k}}{1-r}$. Then, for every i, we have $F_i(\mathcal{U}) \subseteq \mathcal{U}$, since both $ra_k + w_{i,k} \geqslant a_k$ and $rb_k + w_{i,k} \leqslant b_k$ hold for $k = 1, ..., n$. We therefore get $\bigcup_{i=1}^{n} F_i(\mathcal{U}) \subseteq \mathcal{U}$. It remains to prove that $F_i(\mathcal{U}) \cap F_j(\mathcal{U}) = \emptyset$ for all $i \neq j$. Let i_0, j_0 with $i_0 \neq j_0$ be given. We need to show that there exists at least one $k \in \{1, ..., n\}$ such that

$$r(b_k - a_k) \leqslant |w_{i_0,k} - w_{j_0,k}| .$$

For every ℓ we have

$$\begin{aligned} r(b_\ell - a_\ell) &= r\frac{\max_i w_{i,\ell} - \min_i w_{i,\ell}}{1-r} \\ &\leqslant \frac{\mathsf{m}(\mathcal{W})(\max_i w_{i,\ell} - \min_i w_{i,\ell})}{\mathsf{M}(\mathcal{W})} \\ &\leqslant \mathsf{m}(\mathcal{W}). \end{aligned}$$

In particular, we can choose k as the coordinate for which $\|w_{i_0} - w_{j_0}\|_\infty$ is attained. ■

Next, we bound the error in the entropy of the quantized output signals which results from replacing the input distributions by their fractional parts, a crucial step in the proof of [1, Theorem 2].

Lemma 2. *Consider the deterministic matrices* $\mathbf{H}_1, ..., \mathbf{H}_K \in \mathbb{R}^{M \times M}$ *and let* $\mathbf{X}_1^N, ..., \mathbf{X}_K^N$ *be random matrices in* $\mathbb{R}^{M \times N}$. *For every* $k \in \mathbb{N}$, *we have*

$$\left| H\left(\left[\sum_{j=1}^{K} \mathbf{H}_j \mathbf{X}_j^N\right]_k\right) - H\left(\left[\sum_{j=1}^{K} \mathbf{H}_j \big(\mathbf{X}_j^N\big)\right]_k\right)\right| \leqslant \sum_{j=1}^{K} H\big(\lfloor \mathbf{X}_j^N \rfloor\big) + MN \log 2,$$

where $(\mathbf{A}) := \mathbf{A} - \lfloor \mathbf{A} \rfloor$ *denotes the fractional part of the real matrix* $\mathbf{A}$.

Proof: We set

$$\mathbf{V} := 2^k \sum_{j=1}^{K} \mathbf{H}_j \big(\mathbf{X}_j^N\big), \quad \mathbf{W} := 2^k \sum_{j=1}^{K} \mathbf{H}_j \lfloor \mathbf{X}_j^N \rfloor,$$

and show that

$$H(\lfloor \mathbf{V} + \mathbf{W} \rfloor) - H(\lfloor \mathbf{V} \rfloor) \leqslant \sum_{j=1}^{K} H\big(\lfloor \mathbf{X}_j^N \rfloor\big) + MN \log 2 \tag{2}$$

$$H(\lfloor \mathbf{V} \rfloor) - H(\lfloor \mathbf{V} + \mathbf{W} \rfloor) \leqslant \sum_{j=1}^{K} H\big(\lfloor \mathbf{X}_j^N \rfloor\big) + MN \log 2, \tag{3}$$

which yields the claim since entropy is invariant to scaling. We first bound

$$H(\lfloor \mathbf{V}+\mathbf{W}\rfloor) - H(\lfloor \mathbf{V}\rfloor) \leqslant H(\lfloor \mathbf{V}+\mathbf{W}\rfloor \,|\, \lfloor \mathbf{V}\rfloor) \tag{4}$$

$$\leqslant H(\lfloor \lfloor \mathbf{V}\rfloor + \mathbf{W}\rfloor \,|\, \lfloor \mathbf{V}\rfloor) + MN\log 2 \tag{5}$$

$$\leqslant H(\lfloor \mathbf{W}\rfloor) + MN\log 2 \tag{6}$$

$$\leqslant \sum_{j=1}^{K} H\big(\lfloor \mathbf{X}_j^N \rfloor\big) + MN\log 2, \tag{7}$$

where (4) follows from the chain rule, (5) holds since $\lfloor \lfloor \mathbf{V}\rfloor + \mathbf{W}\rfloor = \lfloor \mathbf{V}\rfloor + \lfloor \mathbf{W}\rfloor$ and thus

$$H(\lfloor \mathbf{V}+\mathbf{W}\rfloor \,|\, \lfloor \mathbf{V}\rfloor) - H(\lfloor \lfloor \mathbf{V}\rfloor + \mathbf{W}\rfloor \,|\, \lfloor \mathbf{V}\rfloor) \leqslant H(\lfloor \mathbf{V}+\mathbf{W}\rfloor \,|\, \lfloor \mathbf{V}\rfloor, \lfloor \lfloor \mathbf{V}\rfloor + \mathbf{W}\rfloor) \tag{8}$$

$$= H(\lfloor \mathbf{V}+\mathbf{W}\rfloor \,|\, \lfloor \mathbf{V}\rfloor, \lfloor \mathbf{W}\rfloor) \tag{9}$$

$$\leqslant MN\log 2. \tag{10}$$

Furthermore, (6) is again due to $\lfloor \lfloor \mathbf{V}\rfloor + \mathbf{W}\rfloor = \lfloor \mathbf{V}\rfloor + \lfloor \mathbf{W}\rfloor$, and (7) follows from $H(\lfloor \mathbf{W}\rfloor) \leqslant H(\mathbf{W}) \leqslant H\big(\lfloor \mathbf{X}_1^N \rfloor, ..., \lfloor \mathbf{X}_K^N \rfloor\big) \leqslant \sum_{j=1}^{K} H\big(\lfloor \mathbf{X}_j^N \rfloor\big)$ where in the first two inequalities we used that $H(f(\mathbf{U})) \leqslant H(\mathbf{U})$ for discrete random matrices $\mathbf{U}$ and deterministic functions f. This proves (2).

The argument leading to (3) goes as follows:

$$H(\lfloor \mathbf{V}\rfloor) - H(\lfloor \mathbf{V}+\mathbf{W}\rfloor) \leqslant H(\lfloor \mathbf{V}\rfloor \,|\, \lfloor \mathbf{V}+\mathbf{W}\rfloor) \tag{11}$$

$$\leqslant H(\lfloor \mathbf{W}\rfloor) + H(\lfloor \mathbf{V}\rfloor \,|\, \lfloor \mathbf{V}+\mathbf{W}\rfloor, \lfloor \mathbf{W}\rfloor) \tag{12}$$

$$\leqslant H(\lfloor \mathbf{W}\rfloor) + MN\log 2 \tag{13}$$

$$\leqslant \sum_{j=1}^{K} H\big(\lfloor \mathbf{X}_j^N \rfloor\big) + MN\log 2, \tag{14}$$

where (11) and (12) follow from the chain rule, (13) holds since given $\lfloor \mathbf{V}+\mathbf{W}\rfloor$ and $\lfloor \mathbf{W}\rfloor$, each entry of $\mathbf{V}$ is determined up to 1 bit uncertainty, and (14) follows in the same way as (7) above. ∎

The following lemma is a straightforward extension of [2, Lem. 14] to the vector case.

Lemma 3. *Let $\mathcal{V} \subseteq \mathbb{R}^n$ be a set such that $0 < \mathsf{m}(\mathcal{V}), \mathsf{M}(\mathcal{V}) < \infty$ and let $r > 0$ be such that*

$$r \leqslant \frac{\mathsf{m}(\mathcal{V})}{\mathsf{m}(\mathcal{V}) + \mathsf{M}(\mathcal{V})}. \tag{15}$$

Then, for every $\ell \in \mathbb{N}$ with $\ell \geqslant 1$, we have

$$\mathsf{m}(\mathcal{V} + r\mathcal{V} + \ldots + r^{\ell-1}\mathcal{V}) \geqslant r^{\ell-1}\mathsf{m}(\mathcal{V}). \tag{16}$$

Moreover, the mapping $\mathcal{V}^\ell \to \mathcal{V} + r\mathcal{V} + \ldots + r^{\ell-1}\mathcal{V}$, $(v_1, ..., v_\ell) \mapsto v_1 + rv_2 + \ldots + r^{\ell-1}v_\ell$ is a one-to-one correspondence.

Proof: We begin by proving (16). Let $(v_1, ..., v_\ell)$ and $(w_1, ..., w_\ell)$ be distinct elements of $\mathcal{V}^\ell$, and let $k := \min\{i \mid v_i \neq w_i\}$. Using the reverse triangle inequality, we get

$$\begin{aligned}\left\|\sum_{i=1}^{\ell} r^{i-1}(v_i - w_i)\right\|_\infty &= \left\|\sum_{i=k}^{\ell} r^{i-1}(v_i - w_i)\right\|_\infty \\ &\geqslant \left\|r^{k-1}(v_k - w_k)\right\|_\infty - \left\|\sum_{i=k+1}^{\ell} r^{i-1}(v_i - w_i)\right\|_\infty \\ &\geqslant r^{k-1}\left\|v_k - w_k\right\|_\infty - \sum_{i=k+1}^{\ell} r^{i-1}\left\|v_i - w_i\right\|_\infty .\end{aligned}$$

It therefore follows that

$$\mathsf{m}(\mathcal{V} + r\mathcal{V} + \ldots + r^{\ell-1}\mathcal{V}) \geqslant \min_{1\leqslant k\leqslant \ell}\left\{r^{k-1}\mathsf{m}(\mathcal{V}) - \mathsf{M}(\mathcal{V})\sum_{i=k+1}^{\ell} r^{i-1}\right\}. \tag{17}$$

Here, for $k = \ell$, the sum over an empty index set on the RHS of (17) is to be understood as being equal to 0. The minimum in (17) is attained for $k = \ell$ as by (15) we have $r\mathsf{m}(\mathcal{V}) \leqslant \mathsf{m}(\mathcal{V}) - r\mathsf{M}(\mathcal{V})$ and therefore

$$\begin{aligned}r^{\ell-1}\mathsf{m}(\mathcal{V}) &= r^{\ell-2}r\mathsf{m}(\mathcal{V}) \\ &\leqslant r^{\ell-2}\left(\mathsf{m}(\mathcal{V}) - r\mathsf{M}(\mathcal{V})\right) \\ &\ldots \\ &\leqslant r^{k-1}\mathsf{m}(\mathcal{V}) - \mathsf{M}(\mathcal{V})\sum_{i=k+1}^{\ell} r^{i-1},\end{aligned}$$

for $k = 1, ..., \ell$. In particular, this shows that the mapping $\mathcal{V}^\ell \to \mathcal{V}+r\mathcal{V}+\ldots+r^{\ell-1}\mathcal{V}$, $(v_1, ..., v_\ell) \mapsto v_1 + rv_2 + \ldots + r^{\ell-1}v_\ell$ is injective. Since it is clearly also surjective, the proof is completed. ■

III. Proofs of various supplementary statements

A. Proof of [1, Lemma 1]

We first show, in the next two lemmata, that restricting to an exponential subsequence of k in the computation of information dimension [1, (7)] does not change the limit. This extends corresponding results for random variables in [3] to the vector case.

Lemma 4 ([3, Lem. 16])**.** *Let X be a random vector in $\mathbb{R}^n$. For $p, q \in \mathbb{N}\backslash\{0\}$ we have*

$$H(\langle X\rangle_p) \leqslant H(\langle X\rangle_q) + n\log\left(\left\lceil\frac{p}{q}\right\rceil + 1\right).$$

Proof: By the chain rule we find

$$H(\langle X\rangle_p) = H(\langle X\rangle_p, \langle X\rangle_q) - H(\langle X\rangle_q \,|\, \langle X\rangle_p) \tag{18}$$

$$\leqslant H(\langle X\rangle_p, \langle X\rangle_q) \tag{19}$$

$$= H(\langle X\rangle_q) + H(\langle X\rangle_p \,|\, \langle X\rangle_q). \tag{20}$$

Further $H(\langle X\rangle_p \,|\, \langle X\rangle_q) = \sum_{\ell=(\ell_1,\ldots,\ell_n)\in\mathbb{Z}^n} P_{\langle X\rangle_q}\left(\frac{\ell}{q}\right) H\left(\langle X\rangle_p \,|\, \langle X\rangle_q = \frac{\ell}{q}\right)$ and for fixed $\ell = (\ell_1, \ldots, \ell_n)$, we have

$$H\left(\langle X\rangle_p \,|\, \langle X\rangle_q = \frac{\ell}{q}\right) = H\left(\langle X\rangle_p \,|\, X \in \left[\frac{\ell_1}{q}, \frac{\ell_1+1}{q}\right) \times \ldots \times \left[\frac{\ell_n}{q}, \frac{\ell_n+1}{q}\right)\right) \tag{21}$$

$$= H\left(\lfloor pX\rfloor \,|\, pX \in \left[\frac{p\ell_1}{q}, \frac{p(\ell_1+1)}{q}\right) \times \ldots \times \left[\frac{p\ell_n}{q}, \frac{p(\ell_n+1)}{q}\right)\right) \tag{22}$$

$$\leqslant n\log\left(\left\lceil\frac{p}{q}\right\rceil + 1\right) \tag{23}$$

since there are at most $\left(\left\lceil\frac{p}{q}\right\rceil + 1\right)^n$ possible values of $\lfloor pX\rfloor$ given that $pX \in \left[\frac{p\ell_1}{q}, \frac{p(\ell_1+1)}{q}\right) \times \ldots \times \left[\frac{p\ell_n}{q}, \frac{p(\ell_n+1)}{q}\right)$. Therefore, we get $H(\langle X\rangle_p \,|\, \langle X\rangle_q) \leqslant n\log\left(\left\lceil\frac{p}{q}\right\rceil + 1\right)$, which, upon inserting into (20) establishes the result. ∎

Lemma 5 ([3, Prop. 2])**.** *Let X be a random vector in $\mathbb{R}^n$. For $a > 1$ we have*

$$\underline{d}(X) = \liminf_{\ell\to\infty} \frac{H(\langle X\rangle_{a^\ell})}{\log(a^\ell)} \quad \textit{and} \quad \overline{d}(X) = \limsup_{\ell\to\infty} \frac{H(\langle X\rangle_{a^\ell})}{\log(a^\ell)}.$$

Proof: For a fixed $k \in \mathbb{N}$ we find $\ell \in \mathbb{N}$ such that $a^{\ell-1} \leqslant k < a^{\ell}$. By Lemma 4 we then have

$$H(\langle X \rangle_{a^{\ell-1}}) \leqslant H(\langle X \rangle_k) + n \log 2$$

and

$$H(\langle X \rangle_k) \leqslant H(\langle X \rangle_{a^{\ell}}) + n \log 2.$$

Therefore, we find

$$\frac{H(\langle X \rangle_{a^{\ell-1}}) - n \log 2}{\log(a^{\ell})} \leqslant \frac{H(\langle X \rangle_k)}{\log k} \leqslant \frac{H(\langle X \rangle_{a^{\ell}}) + n \log 2}{\log(a^{\ell-1})}$$

and the lemma follows by taking k (and hence also ℓ) to infinity. ■

We are now ready to prove [1, Lemma 1]. For brevity we will only prove the statement involving $\overline{d}$. The proof for $\underline{d}$ is obtained by replacing each "$\limsup$" by "$\liminf$" in the steps below.

Proof of [1, Lemma 1]: We first argue that it suffices to prove the statement for the ℓ^{∞}-ball. Then, for the specific case of the ℓ^{∞}-ball, we relate $\mathbb{E}[\log \mu(B(X;\varepsilon))]$ to the entropy of $[X]_k$ employing an idea already used in the proof of [1, Lemma 5]. By Lemma 5 this then leads to $\overline{d}(X)$ according to

$$\overline{d}(X) = \limsup_{k\to\infty} \frac{H([X]_k)}{k}. \tag{24}$$

Suppose we have two norms, $\|\cdot\|_{\mathrm{A}}$ and $\|\cdot\|_{\mathrm{B}}$, on $\mathbb{R}^n$. Since $\mathbb{R}^n$ is a finite-dimensional vector space, these norms are equivalent [4, p. 273] in the following sense: there exist constants $c, C > 0$ such that for all $x \in \mathbb{R}^n$

$$c\|x\|_{\mathrm{B}} \leqslant \|x\|_{\mathrm{A}} \leqslant C\|x\|_{\mathrm{B}}.$$

Therefore, for $\varepsilon > 0$ we have

$$B_{\|\cdot\|_{\mathrm{B}}}\Big(x; \frac{\varepsilon}{C}\Big) \subseteq B_{\|\cdot\|_{\mathrm{A}}}(x;\varepsilon) \subseteq B_{\|\cdot\|_{\mathrm{B}}}\Big(x; \frac{\varepsilon}{c}\Big),$$

where $B_{\|\cdot\|}(x;\varepsilon)$ denotes the ball with center x and radius ε with respect to the norm $\|\cdot\|$. It now follows that

$$\limsup_{\varepsilon\to 0} \frac{\mathbb{E}\big[\log \mu\big(B_{\|\cdot\|_{\mathrm{B}}}(X; \frac{\varepsilon}{C})\big)\big]}{\log \varepsilon} \leqslant \limsup_{\varepsilon\to 0} \frac{\mathbb{E}\big[\log \mu\big(B_{\|\cdot\|_{\mathrm{A}}}(X;\varepsilon)\big)\big]}{\log \varepsilon} \tag{25}$$

$$\leqslant \limsup_{\varepsilon\to 0} \frac{\mathbb{E}\big[\log \mu\big(B_{\|\cdot\|_{\mathrm{B}}}(X; \frac{\varepsilon}{c})\big)\big]}{\log \varepsilon}. \tag{26}$$

We furthermore have

$$\limsup_{\varepsilon\to 0}\frac{\mathbb{E}\big[\log\mu\big(B_{\|\cdot\|_{\mathrm{B}}}\big(X;\frac{\varepsilon}{C}\big)\big)\big]}{\log\varepsilon}=\limsup_{\varepsilon\to 0}\frac{\mathbb{E}\big[\log\mu\big(B_{\|\cdot\|_{\mathrm{B}}}\big(X;\frac{\varepsilon}{C}\big)\big)\big]}{\log\varepsilon}\frac{\log\varepsilon}{\log\frac{\varepsilon}{C}}$$

$$=\limsup_{\varepsilon'\to 0}\frac{\mathbb{E}\big[\log\mu\big(B_{\|\cdot\|_{\mathrm{B}}}(X;\varepsilon')\big)\big]}{\log\varepsilon'},$$

and similarly for C replaced by c. In summary, it thus follows from (25) and (26) that

$$\limsup_{\varepsilon\to 0}\frac{\mathbb{E}\big[\log\mu\big(B_{\|\cdot\|_{\mathrm{A}}}(X;\varepsilon)\big)\big]}{\log\varepsilon}=\limsup_{\varepsilon\to 0}\frac{\mathbb{E}\big[\log\mu\big(B_{\|\cdot\|_{\mathrm{B}}}(X;\varepsilon)\big)\big]}{\log\varepsilon},$$

which shows that it suffices to prove the statement for $\|\cdot\|_\infty$.

Next, we establish the relationship between $\mathbb{E}\big[\log\mu\big(B_{\|\cdot\|_\infty}(X;\varepsilon)\big)\big]$ and $H([X]_k)$. Let $\varepsilon>0$ be fixed and determine $k\in\mathbb{N}$ such that

$$2^{-k}\leqslant\varepsilon<2^{-k+1}. \tag{27}$$

We then decompose $\mathbb{R}^n$ into cubes of sidelength 2^{-k} according to

$$\mathbb{R}^n=\bigcup_{\mathcal{Q}\in\{\mathcal{Q}_k(x)\,|\,x\in\mathbb{R}^n\}}\mathcal{Q},$$

where

$$\mathcal{Q}_k(x):=\big[\,[x_1]_k,[x_1]_k+2^{-k}\big)\times\ldots\times\big[\,[x_n]_k,[x_n]_k+2^{-k}\big) \tag{28}$$

is the (unique) cube containing x, cf. [1, (130)]. Note that $H([X]_k)=\mathbb{E}\Big[\log\frac{1}{\mu(\mathcal{Q}_k(X))}\Big]$. Since $\mathcal{Q}_k(x)\subseteq B_{\|\cdot\|_\infty}(x;2^{-k})\subseteq B_{\|\cdot\|_\infty}(x;\varepsilon)$, it follows that

$$\mathbb{E}\left[\log\frac{1}{\mu(B_{\|\cdot\|_\infty}(X;\varepsilon))}\right]\leqslant H([X]_k). \tag{29}$$

Next, we construct cubes which are just large enough to contain $B_{\|\cdot\|_\infty}(x;\varepsilon)$ by setting

$$\widehat{\mathcal{Q}}_k(x):=\big[\,[x_1]_k-2^{-k+1},[x_1]_k+3\cdot 2^{-k}\big)\times\ldots\times\big[\,[x_n]_k-2^{-k+1},[x_n]_k+3\cdot 2^{-k}\big)\,.$$

By (27) we get $B_{\|\cdot\|_\infty}(x;\varepsilon)\subseteq\widehat{\mathcal{Q}}_k(x)$, which implies

$$H([X]_k)-\mathbb{E}\left[\log\frac{1}{\mu(B_{\|\cdot\|_\infty}(X;\varepsilon))}\right]=\mathbb{E}\left[\log\frac{\mu(B_{\|\cdot\|_\infty}(X;\varepsilon))}{\mu(\mathcal{Q}_k(X))}\right] \tag{30}$$

$$\leqslant\mathbb{E}\left[\log\frac{\mu(\widehat{\mathcal{Q}}_k(X))}{\mu(\mathcal{Q}_k(X))}\right] \tag{31}$$

$$\overset{\text{Jensen}}{\leqslant}\log\mathbb{E}\left[\frac{\mu(\widehat{\mathcal{Q}}_k(X))}{\mu(\mathcal{Q}_k(X))}\right] \tag{32}$$

$$=n\log 5, \tag{33}$$

where in (33) we used the fact that for each $x \in \mathbb{R}^n$ the function $\mu(\mathcal{Q}_k(X))$ is constant on the event $X \in \mathcal{Q}_k(x)$ and that each $\widehat{\mathcal{Q}}_k(x)$ is the union of 5^n cubes of the form (28). Putting (29) and (33) together and using (27), we get

$$\frac{H([X]_k) - n\log 5}{k} \leqslant \frac{\mathbb{E}\big[\log \mu(B_{\|\cdot\|_\infty}(X;\varepsilon))\big]}{\log \varepsilon} \leqslant \frac{H([X]_k)}{k-1}.$$

Sending $\varepsilon \to 0$ and thereby $k \to \infty$, together with (24), completes the proof. ■

B. Proof of [1, Proposition 2]

We first consider the extremal cases $\alpha = 0$ and $\alpha = 1$, which will then allow us to deduce the general result.

Lemma 6. *Suppose X is a random vector in $\mathbb{R}^n$ with discrete distribution and $H(\lfloor X \rfloor) < \infty$. Then, we have*

$$d(X) = 0.$$

Proof: Follows directly from the proof for the scalar case [5, pp. 196–197], since entropy does not depend on the values the discrete random variable takes on, but only on the underlying probabilities. ■

Lemma 7. *Suppose X is a random vector in $\mathbb{R}^n$ with absolutely continuous distribution and $H(\lfloor X \rfloor) < \infty$. Then, we have*

$$d(X) = n.$$

Proof: The arguments follow the lines of the proof for the scalar case [5, Thm. 1]. For completeness we detail the proof for the vector setting. By [1, (8)] it suffices to show that $\underline{d}(X) \geqslant n$. Since the distribution of X is absolutely continuous it has a density $f\colon \mathbb{R}^n \to \mathbb{R}$. The main idea of the proof is to first show the statement for bounded approximations of f and then consider the limit of the relevant terms to get the statement for f. Specifically, for $A > 0$, we define

$$f_A(x) := \begin{cases} f(x), & \text{if } f(x) < A \\ 0, & \text{otherwise.} \end{cases}$$

Since $f_A \to f$ for $A \to \infty$ pointwise and monotonically, by Lebesgue's monotone convergence theorem, we have

$$S(A) := \int_{\mathbb{R}^n} f_A(x)\mathrm{d}x \xrightarrow{A\to\infty} \int_{\mathbb{R}^n} f(x)\mathrm{d}x = 1.$$

For $\ell = (\ell_1, ..., \ell_n) \in \mathbb{Z}^n$ and $k \in \mathbb{N}\backslash\{0\}$, we introduce the notation

$$\mathcal{Q}_k(\ell) := \left[\frac{\ell_1}{k}, \frac{\ell_1+1}{k}\right) \times \ldots \times \left[\frac{\ell_n}{k}, \frac{\ell_n+1}{k}\right)$$
$$p_\ell^{(k)} := \int_{\mathcal{Q}_k(\ell)} f(x)\mathrm{d}x$$
$$p_\ell^{(k)}(A) := \int_{\mathcal{Q}_k(\ell)} f_A(x)\mathrm{d}x.$$

Since f_A is bounded, the integral $\int_{\mathbb{R}^n} f_A(x)\log\left(\frac{1}{f_A(x)}\right)\mathrm{d}x$ exists and we get

$$\int_{\mathbb{R}^n} f_A(x)\log\left(\frac{1}{f_A(x)}\right)\mathrm{d}x = \sum_{\ell\in\mathbb{Z}^n}\int_{\mathcal{Q}_k(\ell)} f_A(x)\log\left(\frac{1}{f_A(x)}\right)\mathrm{d}x \tag{34}$$
$$\leqslant \sum_{\ell\in\mathbb{Z}^n} p_\ell^{(k)}(A)\log\left(\frac{1}{k^n p_\ell^{(k)}(A)}\right) \tag{35}$$
$$= \sum_{\ell\in\mathbb{Z}^n} p_\ell^{(k)}(A)\log\left(\frac{1}{p_\ell^{(k)}(A)}\right) - S(A)n\log k \tag{36}$$
$$\leqslant H(\langle X\rangle_k) - S(A)n\log k, \tag{37}$$

where (35) follows from Jensen's inequality and (36) holds since $\sum_{\ell\in\mathbb{Z}^n} p_\ell^{(k)}(A) = \sum_{\ell\in\mathbb{Z}^n}\int_{\mathcal{Q}_k(\ell)} f_A(x) = S(A)$. For (37) we choose k large enough for $p_\ell^{(k)} \leqslant \frac{1}{e}$ to hold for all $\ell \in \mathbb{Z}^n$; this is possible by absolute continuity of the distribution of X and the fact that the function $z\log\frac{1}{z}$ is monotonically increasing on $[0, \frac{1}{e}]$. Note that by Lemma 4 with $p = k$, $q = 1$, and thanks to $H(\lfloor X\rfloor) < \infty$ by assumption, it follows that $H(\langle X\rangle_k) < \infty$ for $k \in \mathbb{N}\backslash\{0\}$. Moreover, we have

$$S(A)\log\left(\frac{1}{A}\right) \leqslant \int_{\mathbb{R}^n} f_A(x)\log\left(\frac{1}{f_A(x)}\right)\mathrm{d}x,$$

which, when combined with (37) and $H(\langle X\rangle_k) < \infty$, yields that $\int_{\mathbb{R}^n} f_A(x)\log\left(\frac{1}{f_A(x)}\right)\mathrm{d}x$ is finite. From (37) we therefore obtain the finite lower bound

$$\frac{H(\langle X\rangle_k)}{\log k} \geqslant S(A)n + \frac{\int_{\mathbb{R}^n} f_A(x)\log\left(\frac{1}{f_A(x)}\right)\mathrm{d}x}{\log k}.$$

For $\varepsilon > 0$, we now choose A large enough for $n(1 - S(A)) < \varepsilon$ to hold, which then allows us to conclude that

$$\underline{d}(X) = \liminf_{k\to\infty} \frac{H(\langle X\rangle_k)}{\log k} \geqslant n - \varepsilon.$$

Since ε can be arbitrarily small, it follows that $\underline{d}(X) \geqslant n$, which was to be proven. ∎

We are now ready to prove [1, Proposition 2] for general α. As in the proof of Lemma 7, we set

$$\mathcal{Q}_k(\ell) := \left[\frac{\ell_1}{k}, \frac{\ell_1 + 1}{k}\right) \times \ldots \times \left[\frac{\ell_n}{k}, \frac{\ell_n + 1}{k}\right),$$

and define further

$$\begin{aligned}
p_\ell^{(k)} &:= \mu(\mathcal{Q}_k(\ell)) \\
q_\ell^{(k)} &:= \mu_{\text{ac}}(\mathcal{Q}_k(\ell)) \\
r_\ell^{(k)} &:= \mu_{\text{d}}(\mathcal{Q}_k(\ell)) \\
\lambda_\ell^{(k)} &:= \begin{cases} \frac{\alpha q_\ell^{(k)}}{p_\ell^{(k)}}, & \text{for } p_\ell^{(k)} \neq 0 \\ 0, & \text{otherwise,} \end{cases}
\end{aligned}$$

for $k \in \mathbb{N}\backslash\{0\}$ and $\ell \in \mathbb{Z}^n$. Note that $\mu = \alpha\mu_{\text{ac}} + (1-\alpha)\mu_{\text{d}}$ implies $p_\ell^{(k)} = \alpha q_\ell^{(k)} + (1-\alpha) r_\ell^{(k)}$. Denoting the binary entropy function by $H_b(\cdot)$, we find that

$$\begin{aligned}
p_\ell^{(k)} H_b(\lambda_\ell^{(k)}) &= \alpha q_\ell^{(k)} \log\left(\frac{p_\ell^{(k)}}{\alpha q_\ell^{(k)}}\right) + (1-\alpha) r_\ell^{(k)} \log\left(\frac{p_\ell^{(k)}}{(1-\alpha) r_\ell^{(k)}}\right) \\
&= \alpha q_\ell^{(k)} \log\left(\frac{1}{q_\ell^{(k)}}\right) + \alpha q_\ell^{(k)} \log\left(\frac{1}{\alpha}\right) \\
&\quad - \alpha q_\ell^{(k)} \log\left(\frac{1}{p_\ell^{(k)}}\right) + (1-\alpha) r_\ell^{(k)} \log\left(\frac{1}{r_\ell^{(k)}}\right) \\
&\quad + (1-\alpha) r_\ell^{(k)} \log\left(\frac{1}{1-\alpha}\right) - (1-\alpha) r_\ell^{(k)} \log\left(\frac{1}{p_\ell^{(k)}}\right).
\end{aligned}$$

Summing over ℓ then yields

$$0 \leqslant \alpha H(\langle X_{\text{ac}}\rangle_k) + (1-\alpha) H(\langle X_{\text{d}}\rangle_k) + H_b(\alpha) - H(\langle X\rangle_k) = \sum_{\ell\in\mathbb{Z}^n} p_\ell^{(k)} H_b(\lambda_\ell^{(k)}), \tag{38}$$

where X_{ac} and X_{d} are distributed according to μ_{ac} and μ_{d}, respectively. Note that the series on the RHS of (38) converges because $H_b(\lambda_\ell^{(k)}) \leqslant 1$ and $\sum_{\ell\in\mathbb{Z}^n} p_\ell^{(k)} = 1$. By [1, Lemma 4] the assumption $H(\lfloor X \rceil) < \infty$ implies $H(\langle X \rangle_k) < \infty$ for all $k \in \mathbb{N}\backslash\{0\}$, and therefore (38) yields a finite upper bound on $\alpha H(\langle X_{\text{ac}} \rangle_k) + (1-\alpha) H(\langle X_{\text{d}} \rangle_k)$ for all $k \in \mathbb{N}\backslash\{0\}$. In particular, $H(\lfloor X_{\text{ac}} \rceil) < \infty$ and $H(\lfloor X_{\text{d}} \rceil) < \infty$, which allows us to apply Lemmata 6 and 7. Dividing (38) by $\log k$ and taking $k \to \infty$ yields

$$d(X) = \lim_{k\to\infty} \frac{H(\langle X \rangle_k)}{\log k} = \alpha d(X_{\text{ac}}) + (1-\alpha) d(X_{\text{d}}) = n\alpha,$$

as desired. ∎

C. *Proof of [1, Proposition 5]*

By assumption the distribution of each X_j decomposes into a mixture of an absolutely continuous and a discrete part according to $\mu^{(j)} = \alpha_j \mu_{\text{ac}}^{(j)} + (1-\alpha_j)\mu_{\text{d}}^{(j)}$, where $\alpha_j \in [0,1]$. Thanks to $\det \mathbf{H}_{i,j} \neq 0$, for all i, j, again by assumption, all $\mathbf{H}_{i,j}$ are isomorphisms. Since absolutely continuous and discrete distributions are preserved under isomorphisms, the distribution of $\sum_{j=1}^{K} \mathbf{H}_{i,j} X_j$ is given by a convolution of discrete-continuous mixtures with mixture coefficients α_j. Expanding this convolution yields a sum of distributions which are all absolutely continuous except for the term that arises as the convolution of all discrete parts (note that a convolution of distributions is absolutely continuous if one of the factors is absolutely continuous (cf. [1, (104)]), whereas the convolution of discrete distributions is again discrete). Thus, the distribution of $\sum_{j=1}^{K} \mathbf{H}_{i,j} X_j$ is again a discrete-continuous mixture with mixture coefficient $1 - \prod_{j=1}^{K}(1-\alpha_j)$. By [1, Proposition 2] and [1, (24)], it therefore follows that

$$\mathsf{dof}(X_1, ..., X_K; \mathbf{H}) = M \left(\sum_{i=1}^{K} \left[(1 - \prod_{j=1}^{K}(1-\alpha_j)) - (1 - \prod_{j\neq i}^{K}(1-\alpha_j)) \right] \right) \tag{39}$$

$$= M \sum_{i=1}^{K} \alpha_i \prod_{j\neq i}^{K} (1-\alpha_j) \tag{40}$$

$$\leqslant M, \tag{41}$$

where (41) is a consequence of

$$c(K) := \sum_{i=1}^{K} \alpha_i \prod_{j\neq i}^{K} (1-\alpha_j) \leqslant 1. \tag{42}$$

The inequality in (42) can be shown by induction: For $K = 1$, the statement is immediate (using the convention that a product over an empty index set equals 1). For the induction step, we get

$$
\begin{aligned}
c(K+1) &= c(K)(1-\alpha_{K+1}) + \alpha_{K+1}\prod_{j=1}^{K}(1-\alpha_j) && (43)\\
&\leqslant (1-\alpha_{K+1}+\alpha_{K+1})\max\{c(K), \prod_{j=1}^{K}(1-\alpha_j)\} && (44)\\
&= \max\{c(K), \prod_{j=1}^{K}(1-\alpha_j)\} && (45)\\
&\leqslant 1, && (46)
\end{aligned}
$$

where in (46) we used the induction hypothesis. ■

D. $\mathsf{dof}(X_1, X_2, X_3; \mathbf{H}) = 3$ *for* $K = 3$ *and* $M = 2$ *proving [1,* (64)*]*

We set

$$X_1 := \widetilde{X}_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad X_2 := \widetilde{X}_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad X_3 := \widetilde{X}_3 \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \tag{47}$$

where $\widetilde{X}_1, \widetilde{X}_2, \widetilde{X}_3$ are independent random variables with absolutely continuous distributions, and hence $d(\widetilde{X}_i) = 1$, for $i = 1, 2, 3$, by [1, Proposition 2]. From [1, (24)] we get

$$
\begin{aligned}
\mathsf{dof}(X_1, X_2, X_3; \mathbf{H}) &= d\left(\widetilde{X}_1 \begin{pmatrix} a[1] \\ a[2] \end{pmatrix} + (\widetilde{X}_2 + \widetilde{X}_3)\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) - d\left((\widetilde{X}_2 + \widetilde{X}_3)\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right)\\
&\quad + d\left(\widetilde{X}_2 \begin{pmatrix} b[1] \\ b[2] \end{pmatrix} + (\widetilde{X}_1 + \widetilde{X}_3)\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) - d\left((\widetilde{X}_1 + \widetilde{X}_3)\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right)\\
&\quad + d\left(\widetilde{X}_3 \begin{pmatrix} c[1] \\ c[2] \end{pmatrix} + (\widetilde{X}_1 + d[1]\widetilde{X}_2)\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) - d\left((\widetilde{X}_1 + d[1]\widetilde{X}_2)\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right)\\
&= 2 - 1 + 2 - 1 + 2 - 1 = 3,
\end{aligned}
$$

where we used [1, (18)], [1, (19)], and the fact that each pair

$$\left\{\begin{pmatrix} a[1] \\ a[2] \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}, \quad \left\{\begin{pmatrix} b[1] \\ b[2] \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}, \quad \left\{\begin{pmatrix} c[1] \\ c[2] \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\}$$

is linearly independent as $a[1] \neq a[2]$, $b[1] \neq b[2]$, and $c[1] \neq c[2]$, all by assumption.

E. Proof of the extension of [2, Thm. 8] to the vector case

We begin with a lemma showing that "more" interference always leads to a decrease in the terms inside the sum in [1, (24)].

Lemma 8 ([2, Lem. 1])**.** *Let X, Y, and Z be independent random vectors in $\mathbb{R}^n$. Then,*

$$d(X+Y+Z) - d(Y+Z) \leqslant d(X+Y) - d(Y), \tag{48}$$

provided that all appearing information dimension terms exist.

Proof: First, we note that for independent discrete random vectors U, V, W we have by the data processing inequality $I(U; U+V+W) \leqslant I(U; U+V)$ and hence

$$H(U+V+W) - H(V+W) \leqslant H(U+V) - H(V). \tag{49}$$

Applying (49) to $\lfloor kX \rfloor, \lfloor kY \rfloor, \lfloor kZ \rfloor$ and using [1, Lemma 9] thrice with $\delta = 1$ and ε obtained from

$$0 \leqslant \lfloor k(X+Y+Z) \rfloor - (\lfloor kX \rfloor + \lfloor kY \rfloor + \lfloor kZ \rfloor) \leqslant 2,$$

$$0 \leqslant \lfloor k(Y+Z) \rfloor - (\lfloor kY \rfloor + \lfloor kZ \rfloor) \leqslant 1,$$

$$0 \leqslant \lfloor k(X+Y) \rfloor - (\lfloor kX \rfloor + \lfloor kY \rfloor) \leqslant 1,$$

we find that

$$H(\langle X+Y+Z \rangle_k) - H(\langle Y+Z \rangle_k) \leqslant H(\langle X+Y \rangle_k) - H(\langle Y \rangle_k) + n \log 3 + 2n \log 2.$$

Dividing this inequality by $\log k$ and taking $k \to \infty$ yields the claim. ■

We are now able to show that for rational entries in the subchannel matrices, the normalized DoF of the parallel IC are strictly less than $K/2$.

Proposition 1. *Consider a parallel IC with fully connected channel matrices in standard form*

$$\mathbf{H}[m] = \begin{pmatrix} a[m] & 1 & 1 \\ 1 & b[m] & 1 \\ 1 & d[m] & c[m] \end{pmatrix}, \quad m = 1, ..., M,$$

where $a[m], b[m], c[m], d[m] \neq 0$*,* $m = 1, ..., M$*, are rational and* $a[1] = ... = a[M] =: a$*. Then, we have*

$$\frac{\mathsf{DoF}(\mathbf{H})}{M} < \frac{3}{2}.$$

Proof: We define

$$\mathbf{H}'[m] := \begin{pmatrix} a & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad m = 1, ..., M,$$

and denote the overall channel matrix of the corresponding parallel IC by $\mathbf{H}'$. Since all entries of $\mathbf{H}$ and $\mathbf{H}'$ are rational, the DoF-formula [1, (26)] holds by [1, Theorem 4] for both $\mathbf{H}$ and $\mathbf{H}'$. Note that each $\mathbf{H}'[m]$ is obtained by setting three entries of the corresponding $\mathbf{H}[m]$ matrix to zero and then rescaling the second column and the third row. It follows from [1, (26)] and Lemma 8 applied to [1, (53)] that the DoF can only increase if we delete a given interference link over all m, i.e., if we replace $h_{i,j}[m]$, for a fixed pair (i, j) with $i \neq j$, by 0 for all $m = 1, ..., M$. By [1, Lemma 3] scaling of rows and columns with nonzero constants does not change the DoF. In summary, we find that $\mathsf{DoF}(\mathbf{H}) \leqslant \mathsf{DoF}(\mathbf{H}')$. For general transmit vectors X_1, X_2, X_3, we get

$$\mathsf{dof}(X_1, X_2, X_3; \mathbf{H}') = \underbrace{d(aX_1 + X_3) - d(X_1 + X_3)}_{<M/2} + \underbrace{d(X_1 + X_2 + X_3)}_{\leqslant M} < \frac{3M}{2},$$

where $d(aX_1 + X_3) - d(X_1 + X_3) < M/2$ for $a \in \mathbb{Q}$ follows by extending [2, Thm. 3] to the vector case. (The proof of this extension follows along the lines of the arguments used in [2, App. B], except for inequality [2, Eq. (327)], where the “2” has to be replaced by “$2M$”, which then results in the upper bound $M/2 - \varepsilon$ instead of $1/2 - \varepsilon$ as in [2, Thm. 3].) ■

References

[1] D. Stotz and H. Bölcskei, “Degrees of freedom in vector interference channels,” *to appear in IEEE Trans. Inf. Theory, arXiv:1210.2259v2*, vol. cs.IT, 2016.

[2] Y. Wu, S. Shamai (Shitz), and S. Verdú, “A formula for the degrees of freedom of the interference channel,” *IEEE Trans. Inf. Theory*, vol. 61, no. 1, pp. 256–279, Jan. 2015.

[3] 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.

[4] R. A. Horn and C. R. Johnson, *Matrix analysis*. Cambridge, UK: Cambridge University Press, 1990.

[5] A. Rényi, “On the dimension and entropy of probability distributions,” *Acta Mathematica Hungarica*, vol. 10, pp. 1–23, Mar. 1959.