Explicit and Almost Sure Conditions for $K/2$ Degrees of Freedom

David Stotz and Helmut Bölcskei

Dept. IT & EE, ETH Zurich, Switzerland
Email: {dstotz, boelcskei}@nari.ee.ethz.ch

Abstract—It is well known that in K-user constant single-antenna interference channels $K/2$ degrees of freedom (DoF) can be achieved for almost all channel matrices. Explicit conditions on the channel matrix to admit $K/2$ DoF are, however, not available. The purpose of this paper is to identify such explicit conditions, which are satisfied for almost all channel matrices. We also provide a construction of corresponding asymptotically DoF-optimal input distributions. The main technical tool used is a recent breakthrough result by Hochman in fractal geometry [1].

I. INTRODUCTION

Characterizing the degrees of freedom (DoF) in interference channels (ICs) under various assumptions on the channel matrix has become a heavily researched topic in recent years [2]. A particularly surprising result states that $K/2$ DoF can be achieved in single-antenna K-user ICs with constant channel matrix [3]. This statement was shown to hold for (Lebesgue) almost all channel matrices [3, Thm. 1]. The technical arguments—from Diophantine approximation theory—used in the proof of [3, Thm. 1] do not seem to allow an explicit characterization of the set of these channel matrices. What is known, though, is that channel matrices with all entries rational admit strictly less than $K/2$ DoF [4] and hence belong to the set of exceptions relative to the "almost-all result" in [3].

Recently, Wu et al. [5] developed a general framework, based on (Rényi's) information dimension, for characterizing the DoF in constant single-antenna ICs. While this general and elegant theory allows to recover, inter alia, the "almost-all result" from [3], it does not provide insights into the structure of the set of channel matrices admitting $K/2$ DoF.

Contributions: The main contribution of this paper is to complement the results in [3]–[5] by providing *explicit and almost surely satisfied* conditions on the channel matrix to admit $K/2$ DoF. The conditions we find essentially require that the set of all monomial expressions in the channel coefficients be linearly independent over the rational numbers. The proof of our main theorem employs a recent breakthrough result from fractal geometry [1], which allows us to compute the information dimension of self-similar distributions under much milder conditions than the previously required open set condition [6]. For the channels satisfying our explicit and almost sure conditions, we furthermore provide a construction of asymptotically DoF-optimal input distributions.

Finally, we show that the explicit sufficient conditions for $K/2$ DoF we identify are not necessary. This is accomplished by constructing examples of channel matrices that admit $K/2$ DoF but do not satisfy our conditions. The set of all such channel matrices, however, necessarily has Lebesgue measure zero.

Notation: Random variables are represented by uppercase letters from the end of the alphabet. Lowercase letters are used exclusively for deterministic quantities. Boldface uppercase letters indicate matrices. Sets are denoted by uppercase calligraphic letters. For $x \in \mathbb{R}$, we write $\lfloor x \rfloor$ for the largest integer not exceeding x. All logarithms are taken to the base 2. $\mathbb{E}[\cdot]$ denotes the expectation operator. $H(\cdot)$ stands for entropy and $h(\cdot)$ for differential entropy.

II. SYSTEM MODEL

We consider a constant single-antenna K-user IC with channel matrix $\mathbf{H} = (h_{ij})_{1 \leqslant i,j \leqslant K} \in \mathbb{R}^{K \times K}$, and input-output relation

$$Y_i = \sqrt{\mathsf{snr}} \sum_{j=1}^{K} h_{ij} X_j + Z_i, \quad i = 1, ..., K, \tag{1}$$

where $X_i \in \mathbb{R}$ is the input at the i-th transmitter, $Y_i \in \mathbb{R}$ is the output at the i-th receiver, and $Z_i \in \mathbb{R}$ is noise of absolutely continuous distribution with $h(Z_i) > -\infty$ and $H(\lfloor Z_i \rfloor) < \infty$. The input signals at different transmitters are independent and noise is i.i.d. across users and channel uses.

The channel matrix $\mathbf{H}$ is assumed to be known perfectly at all transmitters and receivers. We impose the average power constraint

$$\frac{1}{n} \sum_{k=1}^{n} \left(x_i^{(k)} \right)^2 \leqslant 1$$

on codewords $\left(x_i^{(1)} \, ... \, x_i^{(n)} \right)$ of block-length n transmitted by user $i = 1, ..., K$. The DoF are defined as

$$\mathsf{DoF}(\mathbf{H}) := \limsup_{\mathsf{snr} \to \infty} \frac{\overline{C}(\mathbf{H}; \mathsf{snr})}{\frac{1}{2} \log \mathsf{snr}}, \tag{2}$$

where $\overline{C}(\mathbf{H}; \mathsf{snr})$ is the sum-capacity of the channel (1).

III. MAIN RESULT

We denote the vector consisting of the off-diagonal coefficients of $\mathbf{H}$ by $\check{\mathbf{h}} \in \mathbb{R}^{K(K-1)}$, and let $f_1, f_2, ...$ be the

monomials of all degrees[1] in $K(K-1)$ variables enumerated as follows: $f_1, ..., f_{\varphi(d)}$ are the monomials of degree not larger than d, where

$$\varphi(d) := \binom{K(K-1)+d}{d}.$$

The main result of this paper is the following theorem.

Theorem 1: Suppose that the channel matrix $\mathbf{H}$ satisfies the following condition:

For each $i = 1, ..., K$, the set

$$\{f_j(\check{\mathbf{h}}) : j \geqslant 1\} \cup \{h_{ii} f_j(\check{\mathbf{h}}) : j \geqslant 1\} \qquad (*)$$

is linearly independent over $\mathbb{Q}$.

Then, we have

$$\mathsf{DoF}(\mathbf{H}) = K/2.$$

Proof: See Section VII. ∎

The proof of Theorem 1 is constructive in the sense of providing an explicit construction of a sequence of input distributions that asymptotically achieves $\mathsf{DoF}(\mathbf{H}) = K/2$ for all $\mathbf{H}$ satisfying Condition $(*)$.

Note that the prominent example from [4] with all entries of $\mathbf{H}$ rational, shown in [4] to admit strictly less than $K/2$ DoF, does not satisfy Condition $(*)$, since two rationals are always linearly dependent over $\mathbb{Q}$.

To see that Condition $(*)$ is satisfied for (Lebesgue) almost all channel matrices, we first note that for fixed $d \in \mathbb{N}$, fixed $a_1, ..., a_{\varphi(d)}, b_1, ..., b_{\varphi(d)} \in \mathbb{Z}$, and fixed $i \in \{1, ..., K\}$,

$$\sum_{j=1}^{\varphi(d)} a_j f_j(\check{\mathbf{h}}) + \sum_{j=1}^{\varphi(d)} b_j h_{ii} f_j(\check{\mathbf{h}}) = 0 \qquad (3)$$

is satisfied only on a set of measure 0 with respect to the choice of $\mathbf{H}$. It suffices to consider linear combinations with coefficients in $\mathbb{Z}$ as (3) with rational coefficients can be multiplied by a common denominator. Since the set of equations (3) is countable with respect to $d \in \mathbb{N}$, $a_1, ..., a_{\varphi(d)}, b_1, ..., b_{\varphi(d)} \in \mathbb{Z}$, and $i \in \{1, ..., K\}$, it follows that Condition $(*)$ is satisfied for almost all channel matrices $\mathbf{H}$. Theorem 1 therefore provides (Lebesgue) almost surely satisfied explicit conditions for $\mathbf{H}$ to admit $K/2$ DoF.

We proceed by developing, in Sections IV and V, preparatory material needed for the proof of Theorem 1.

IV. INFORMATION DIMENSION AND DOF

Definition 1: Let X be a random variable with distribution μ. We define the *lower* and *upper information dimension* of X as

$$\underline{d}(X) := \liminf_{k \to \infty} \frac{H(\langle X \rangle_k)}{\log k} \quad \text{and} \quad \overline{d}(X) := \limsup_{k \to \infty} \frac{H(\langle X \rangle_k)}{\log k},$$

where $\langle X \rangle_k := \lfloor kX \rfloor / k$. If $\underline{d}(X) = \overline{d}(X)$, then we set $d(X) := \underline{d}(X) = \overline{d}(X)$ and call $d(X)$ *the information dimension* of X. Since $\underline{d}(X), \overline{d}(X)$, and $d(X)$ depend on μ only, we sometimes also write $\underline{d}(\mu), \overline{d}(\mu)$, and $d(\mu)$, respectively.

The relevance of information dimension in characterizing DoF stems from the following relation [5], [7], [8]

$$\limsup_{\mathsf{snr} \to \infty} \frac{h(\sqrt{\mathsf{snr}}X + Z)}{\frac{1}{2} \log \mathsf{snr}} = \overline{d}(X), \qquad (4)$$

which holds for arbitrary independent random variables X and Z, with the distribution of Z absolutely continuous and such that $h(Z) > -\infty$, $H(\lfloor Z \rfloor) < \infty$.

We can apply (4) to ICs as follows. By standard random coding arguments it follows for the IC (1) that the sum-rate

$$I(X_1; Y_1) + ... + I(X_K; Y_K) \qquad (5)$$

is achievable, where $X_1, ..., X_K$ are independent input distributions with $\mathbb{E}[X_i^2] \leqslant 1$, $i = 1, ..., K$. Using the chain rule, we get

$$I(X_i; Y_i) = h(Y_i) - h(Y_i \,|\, X_i) \qquad (6)$$

$$= h\left(\sqrt{\mathsf{snr}} \sum_{j=1}^{K} h_{ij} X_j + Z_i\right) - h\left(\sqrt{\mathsf{snr}} \sum_{j \neq i}^{K} h_{ij} X_j + Z_i\right) \qquad (7)$$

for $i = 1, ..., K$. Combining (4)-(7), we obtain

$$\mathrm{dof}(X_1, ..., X_K; \mathbf{H}) :=$$

$$\sum_{i=1}^{K} \left[d\left(\sum_{j=1}^{K} h_{ij} X_j\right) - d\left(\sum_{j \neq i}^{K} h_{ij} X_j\right) \right] \qquad (8)$$

$$\leqslant \mathsf{DoF}(\mathbf{H}), \qquad (9)$$

for all independent $X_1, ..., X_K$ with[2] $\mathbb{E}[X_i^2] < \infty$, $i = 1, ..., K$, and such that all information dimension terms appearing in (8) exist. Relation (9) was first reported in [5]. A striking result in [5] shows that input distributions of discrete, continuous, or mixed discrete-continuous nature can achieve no more than 1 DoF. For $K > 2$, DoF-optimal input distributions therefore necessarily have a singular component.

V. ITERATED FUNCTION SYSTEMS

A class of singular distributions with explicit expressions for their information dimension are self-similar distributions defined as follows. Consider a finite set $\Phi_r := \{\varphi_{i,r} : i = 1, ..., n\}$ consisting of affine contractions of $\mathbb{R}$, i.e.,

$$\varphi_{i,r}(x) = rx + w_i, \qquad (10)$$

where $r \in I \subseteq (0, 1)$ and the w_i are pairwise different real numbers. We furthermore let $\mathcal{W} := \{w_1, ..., w_n\}$. Φ_r is called an iterated function system (IFS) parametrized by the contraction parameter $r \in I$. By classical fractal geometry [9, Ch. 9] every IFS has an associated unique attractor, i.e., a non-empty compact set $\mathcal{A} \subseteq \mathbb{R}$ such that

$$\mathcal{A} = \bigcup_{i=1}^{n} \varphi_{i,r}(\mathcal{A}).$$

---

[1]The "degree" of a monomial is to be understood as the sum of all exponents of the variables involved (sometimes called the total degree).

[2]We only need the conditions $\mathbb{E}[X_i^2] < \infty$ as scaling of the inputs does not affect $\mathrm{dof}(X_1, ..., X_K; \mathbf{H})$.

Moreover, for each probability vector $(p_1, ..., p_n)$, there is a unique (Borel) probability distribution μ_r on $\mathbb{R}$ such that

$$\mu_r = \sum_{i=1}^{n} p_i (\varphi_{i,r})_* \mu_r, \tag{11}$$

where $(\varphi_{i,r})_* \mu_r$ denotes the push-forward of μ_r by $\varphi_{i,r}$. The distribution μ_r is supported on $\mathcal{A}$ and is called the self-similar distribution corresponding to the IFS Φ_r with underlying probability vector $(p_1, ..., p_n)$. We can give the following explicit expression for a random variable X with distribution μ_r in (11)

$$X = \sum_{k=0}^{\infty} r^k W_k, \tag{12}$$

where $\{W_k\}_{k \geqslant 0}$ is a set of i.i.d. copies of a random variable W drawn from the set $\mathcal{W}$ according to $(p_1, ..., p_n)$.

VI. THE MAIN IDEAS

Classical results in fractal geometry allow an analytical expression for the information dimension of a self-similar distribution under the so-called open set condition [10, Thm. 2]. This condition requires the existence of a non-empty open set whose images under the elements of Φ_r do not overlap and all lie in the open set (see, e.g., [6]). Wu et al. [5] ensure that the open set condition is satisfied by imposing an upper bound on the contraction parameter r according to

$$r \leqslant \frac{\mathsf{m}(\mathcal{W})}{\mathsf{m}(\mathcal{W}) + \mathsf{M}(\mathcal{W})}. \tag{13}$$

Here, $\mathsf{m}(\mathcal{W}) := \min_{i \neq j} |w_i - w_j|$ and $\mathsf{M}(\mathcal{W}) := \max_{i,j} |w_i - w_j|$. The authors of [5] construct $K/2$ DoF-achieving input distributions by building $\mathcal{W}$ from $\mathbb{Z}$-linear combinations of monomial expressions in the off-diagonal channel coefficients; this idea inspired our Condition $(*)$. However, to ensure that the construction in [5] meets (13) the minimum and maximum distance of the elements in $\mathcal{W}$ needs to be controlled. This results in the question of how fast a polynomial in real variables with integer coefficients approaches an integer [5], a problem studied in Diophantine approximation theory. The nature of the results applied in [5] to deal with this question does not seem to allow an explicit characterization of the channel matrices that admit $K/2$ DoF. Recent groundbreaking work by Hochman [1] replaces the open set condition by a much weaker condition, which essentially requires that the IFS must not allow "exact overlap". This improvement turns out to be instrumental in the proof of Theorem 1. Specifically, we use the following simple consequence of a key result by Hochman [1, Thm. 1.8].

Theorem 2: If $I \subseteq (0, 1)$ is a non-empty compact interval, and μ_r is the self-similar distribution with underlying contraction parameter $r \in I$ and probability vector $(p_1, ..., p_n)$, then[3]

$$d(\mu_r) = \min\left\{ \frac{\sum p_i \log p_i}{\log r}, 1 \right\}, \tag{14}$$

[3]The "1" in the minimum simply accounts for the fact that information dimension cannot exceed the dimension of the ambient space.

for all $r \in I \setminus E$, where E is a set of Hausdorff and packing dimension 0.

Proof: For $\mathbf{i} \in \{1, ..., n\}^k$, let $\varphi_{\mathbf{i},r} := \varphi_{i_1,r} \circ \ldots \circ \varphi_{i_k,r}$ and define

$$\Delta_{\mathbf{i},\mathbf{j}}(r) := \varphi_{\mathbf{i},r}(0) - \varphi_{\mathbf{j},r}(0),$$

for $\mathbf{i}, \mathbf{j} \in \{1, ..., n\}^k$. Extend this definition to infinite sequences $\mathbf{i}, \mathbf{j} \in \{1, ..., n\}^{\mathbb{N}}$ according to

$$\Delta_{\mathbf{i},\mathbf{j}}(r) := \lim_{k \to \infty} \Delta_{(i_1,...,i_k),(j_1,...,j_k)}(r).$$

Using (10) it follows that

$$\Delta_{\mathbf{i},\mathbf{j}}(r) = \sum_{k=1}^{\infty} r^{k-1}(w_{i_k} - w_{j_k}).$$

Since the w_i are pairwise different and a power series can vanish on a non-empty interval only if it is identically zero, we get that $\Delta_{\mathbf{i},\mathbf{j}} \equiv 0$ on a compact interval I if and only if $\mathbf{i} = \mathbf{j}$. The result now follows by application of [1, Thm. 1.8]. ∎

Remark 1: Note that (14) can be rewritten in terms of the entropy of the random variable W, defined implicitly in (12), which takes value w_i with probability p_i:

$$d(\mu_r) = \min\left\{ \frac{H(W)}{\log(1/r)}, 1 \right\}. \tag{15}$$

We wish to construct self-similar input distributions that yield $\mathrm{dof}(X_1, ..., X_K; \mathbf{H}) = K/2$ for all channel matrices satisfying Condition $(*)$. To this end, we first note that choosing all inputs X_i to have self-similar distributions as in (12) with identical contraction parameter r, the distributions of the random variables $\sum h_{ij} X_j$ appearing in (8) are again self-similar. This allows us to compute the information dimension terms in (8) using (15). The freedom we exploit in constructing full DoF-achieving X_i lies in the choice of the set $\mathcal{W}$, $(p_1, ..., p_n)$ is assumed uniform on $\mathcal{W}$, for simplicity of exposition. Specifically, we want to ensure that the first term inside the sum (8) is twice as big as the second term, for all i. This means that the sum of the desired signal $h_{ii} X_i$ and the interference $\sum_{j \neq i} h_{ij} X_j$ should be twice as "rich" as the interference term alone. From (15) it follows that this doubling is accomplished if

$$\left| \sum_{j=1}^{K} h_{ij} \mathcal{W} \right| \approx \left| \sum_{j \neq i}^{K} h_{ij} \mathcal{W} \right|^2. \tag{16}$$

Let us now turn to the main idea for realizing (16). We build $\mathcal{W}$ as a set of $\mathbb{Z}$-linear combinations of monomial expressions in the off-diagonal channel coefficients, an idea that was introduced in [3]. Multiplying the elements of the so-constructed set $\mathcal{W}$ by an off-diagonal channel coefficient simply increases the degrees of the involved monomials by 1 so that the algebraic structure of $h_{ij} \mathcal{W}$, for $i \neq j$, is the same as that of $\mathcal{W}$. In addition, we let the degrees of the involved monomials be large and say that multiplication of $\mathcal{W}$ by h_{ij}, $i \neq j$, results in a set that is "similar" to $\mathcal{W}$, denoted as $h_{ij} \mathcal{W} \approx \mathcal{W}$. Finally, we also take $|\mathcal{W}|$ to be large so that $\sum_{j \neq i} h_{ij} \mathcal{W} \approx \mathcal{W}$, see Fig. 1 (a).

(a) Sum of two sets with common algebraic structure.

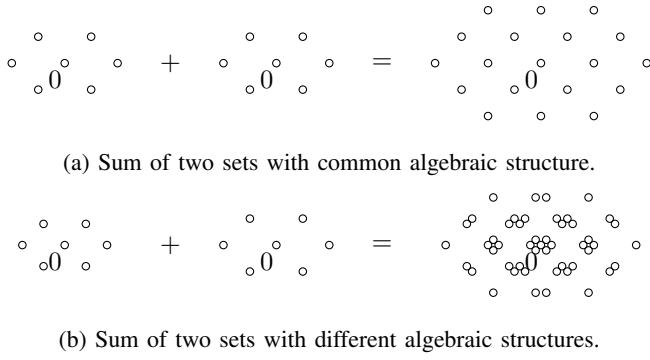(b) Sum of two sets with different algebraic structures.

Fig. 1: The cardinality of the sum in (a) is 19 and hence small compared to the $7^2 = 49$ pairs summed up, whereas the sum in (b) has cardinality 49.

As h_{ii} does not participate in the monomial expressions in $\mathcal{W}$, the set $h_{ii}\mathcal{W}$ is unlikely to have the same algebraic structure as $\mathcal{W}$. In fact, Condition $(*)$ guarantees that this does not happen. We can then conclude that the set $h_{ii}\mathcal{W} + \sum_{j \neq i} h_{ij}\mathcal{W} \approx h_{ii}\mathcal{W} + \mathcal{W}$ roughly has cardinality $|\mathcal{W}|^2$, as desired, see Fig. 1 (b). Moreover, Condition $(*)$ guarantees that every element of $\mathcal{W}$ has exactly one representation as a $\mathbb{Z}$-linear combination of the monomials in the off-diagonal channel coefficients. This lets us control the cardinality of $\mathcal{W}$ by controlling the number of possible $\mathbb{Z}$-linear combinations.

VII. PROOF OF THEOREM 1

We first note that Condition $(*)$ implies that $\mathbf{H}$ cannot contain zeros, i.e., $\mathbf{H}$ must be fully connected. The upper bound $\mathsf{DoF}(\mathbf{H}) \leqslant K/2$ therefore follows from [11, Prop. 1].

The remainder of the proof establishes the lower bound $\mathsf{DoF}(\mathbf{H}) \geqslant K/2$. Let N and d be positive integers. We begin by setting

$$\mathcal{W}_N := \left\{ \sum_{i=1}^{\varphi(d)} a_i f_i(\check{\mathbf{h}}) \; : \; a_1, ..., a_{\varphi(d)} \in \{1, ..., N\} \right\} \quad (17)$$

and $r_N := |\mathcal{W}_N|^{-2}$. Let $\{W_{i,k} \; : \; 1 \leqslant i \leqslant K, k \geqslant 0\}$ be i.i.d. uniform random variables on $\mathcal{W}_N$. For $\varepsilon > 0$, we set $I_\varepsilon := [r_N - \varepsilon, r_N]$ and consider the inputs

$$X_i = \sum_{k=0}^{\infty} r^k W_{i,k}, \quad i = 1, ..., K,$$

where $r \in I_\varepsilon$. Then, the signals

$$\sum_{j=1}^{K} h_{ij} X_j = \sum_{k=0}^{\infty} r^k \sum_{j=1}^{K} h_{ij} W_{j,k}$$

$$\text{and} \quad \sum_{j \neq i}^{K} h_{ij} X_j = \sum_{k=0}^{\infty} r^k \sum_{j \neq i}^{K} h_{ij} W_{j,k}$$

also have self-similar distributions with alphabets $\sum_{j=1}^{K} h_{ij}\mathcal{W}_N$ and $\sum_{j \neq i} h_{ij}\mathcal{W}_N$, respectively. Applying Theorem 2, we can therefore conclude the existence of an

$\tilde{r} \in I_\varepsilon$ such that for $X_i = \sum_{k=0}^{\infty} \tilde{r}^k W_{i,k}$, $i = 1, ..., K$, we have

$$d\left(\sum_{j=1}^{K} h_{ij} X_j \right) = \min\left\{ \frac{H\left(\sum_{j=1}^{K} h_{ij} W_{j,0} \right)}{\log(1/\tilde{r})}, 1 \right\}$$

$$\text{and} \quad d\left(\sum_{j \neq i}^{K} h_{ij} X_j \right) = \min\left\{ \frac{H\left(\sum_{j \neq i} h_{ij} W_{j,0} \right)}{\log(1/\tilde{r})}, 1 \right\},$$

for $i = 1, ..., K$. Taking $\varepsilon \to 0$, we get $\tilde{r} \to r_N$, and by the continuity of $\log(\cdot)$ and (9) it follows that

$$\sum_{i=1}^{K} \left[\min\left\{ \frac{H\left(\sum_{j=1}^{K} h_{ij} W_{j,0} \right)}{\log(1/r_N)}, 1 \right\} \right.$$
$$\left. - \min\left\{ \frac{H\left(\sum_{j \neq i}^{K} h_{ij} W_{j,0} \right)}{\log(1/r_N)}, 1 \right\} \right] \leqslant \mathsf{DoF}(\mathbf{H}). \quad (18)$$

Note that the random variable $\sum_{j \neq i} h_{ij} W_{j,0}$ takes value in

$$\left\{ \sum_{i=1}^{\varphi(d+1)} a_i f_i(\check{\mathbf{h}}) \; : \; a_1, ..., a_{\varphi(d+1)} \in \{1, ..., (K-1)N\} \right\}. \quad (19)$$

By Condition $(*)$ it follows that each element in the set (19) has exactly one representation as a $\mathbb{Z}$-linear combination with coefficients $a_1, ..., a_{\varphi(d+1)} \in \{1, ..., (K-1)N\}$. This allows us to conclude that the cardinality of the set (19) is given by $((K-1)N)^{\varphi(d+1)}$, which implies $H\left(\sum_{j \neq i} h_{ij} W_{j,0} \right) \leqslant \varphi(d+1) \log((K-1)N)$. With $\log(1/r_N) = 2\log|\mathcal{W}_N| = 2\varphi(d)\log N$, we therefore get

$$\frac{H\left(\sum_{j \neq i}^{K} h_{ij} W_{j,0} \right)}{\log(1/r_N)} \leqslant \frac{\varphi(d+1)\log((K-1)N)}{2\varphi(d)\log N} \xrightarrow{d,N\to\infty} \frac{1}{2}, \quad (20)$$

where we used

$$\frac{\varphi(d+1)}{\varphi(d)} = \frac{K(K-1)+d+1}{d+1} \xrightarrow{d\to\infty} 1.$$

Next, note that Condition $(*)$ implies that the sum $h_{ii}W_{i,0} + \sum_{j \neq i} h_{ij}W_{j,0}$ can be separated uniquely into the terms $h_{ii}W_{i,0}$ and $\sum_{j \neq i} h_{ij}W_{j,0}$ in the sense that the pair $(h_{ii}W_{i,0}, \sum_{j \neq i} h_{ij}W_{j,0})$ and the sum $h_{ii}W_{i,0} + \sum_{j \neq i} h_{ij}W_{j,0}$ are related through a bijection. It therefore follows that the pair and the sum have equal entropy [12, Ex. 2.4], and since the $W_{j,0}$, $1 \leqslant j \leqslant K$, are independent, we find that

$$H\left(\sum_{j=1}^{K} h_{ij} W_{j,0} \right) = H\left(h_{ii}W_{i,0}, \sum_{j \neq i} h_{ij} W_{j,0} \right) \quad (21)$$

$$= H(h_{ii}W_{i,0}) + H\left(\sum_{j \neq i} h_{ij} W_{j,0} \right). \quad (22)$$

Putting the pieces together, we obtain

$$\frac{H\left(\sum_{j=1}^{K} h_{ij}W_{j,0}\right) - H\left(\sum_{j\neq i}^{K} h_{ij}W_{j,0}\right)}{\log(1/r_N)} \qquad (23)$$

$$= \frac{H(h_{ii}W_{i,0})}{2\varphi(d)\log N} = \frac{\varphi(d)\log N}{2\varphi(d)\log N} = \frac{1}{2}, \qquad (24)$$

where we used the scaling invariance of entropy, the fact that $W_{i,0}$ is uniform on $\mathcal{W}$, and $|\mathcal{W}| = N^{\varphi(d)}$. Finally, (24) and (20) imply that the left-hand side of (18) tends to $\sum_{i=1}^{K}(1 - \frac{1}{2}) = K/2$ for $d, N \to \infty$. This completes the proof. ∎

VIII. CONDITION (∗) IS NOT NECESSARY

While Condition (∗) is sufficient for $\mathsf{DoF}(\mathbf{H}) = K/2$, we next show that it is not necessary. This will be accomplished by constructing a class of example channels which fail to satisfy Condition (∗) but still admit $K/2$ DoF. As almost every channel matrix satisfies Condition (∗) this example class necessarily has Lebesgue measure zero. Specifically, we consider channel matrices that have $h_{ii} \in \mathbb{R}\backslash\mathbb{Q}$, $i = 1, ..., K$, and $h_{ij} \in \mathbb{Q}\backslash\{0\}$, for $i, j = 1, ..., K$ with $i \neq j$. Since this means, in particular, that all channel coefficients are nonzero, $\mathbf{H}$ is fully connected. The upper bound $\mathsf{DoF}(\mathbf{H}) \leqslant K/2$ is therefore again implied by [11, Prop. 1]. Moreover, as two rational numbers are linearly dependent over $\mathbb{Q}$, these channel matrices do not satisfy Condition (∗). We next show that this example class nevertheless has $\mathsf{DoF}(\mathbf{H}) \geqslant K/2$ and hence $\mathsf{DoF}(\mathbf{H}) = K/2$. This will be accomplished by constructing corresponding asymptotically DoF-optimal input distributions. First, we argue that we may assume $h_{ij} \in \mathbb{Z}$, for $i \neq j$. This follows from the fact that multiplying the channel coefficients h_{ij}, $i, j = 1, ..., K$, by a common denominator of the h_{ij}, $i \neq j$, is equivalent to scaling of the inputs X_i, which does not impact DoF.

Let

$$\mathcal{W} := \{0, ..., N-1\},$$

for some $N > 0$, and let $\{W_{i,k} : 1 \leqslant i \leqslant K, k \geqslant 0\}$ be i.i.d. uniform on $\mathcal{W}$. We set the contraction parameter to

$$r = 2^{-2\log(2h_{\max}KN)}, \qquad (25)$$

where $h_{\max} := \max\{|h_{ij}| : i \neq j\}$. Since $\sum_{j\neq i} h_{ij}W_{j,0}$ is integer-valued and $h_{ii}W_{i,0}$ is irrational, the sum $\sum_{j=1}^{K} h_{ij}W_{j,0}$ can uniquely be separated into the terms $h_{ii}W_{i,0}$ and $\sum_{j\neq i} h_{ij}W_{j,0}$ in the sense of (22). As $h_{ij} \in \mathbb{Z}$ for $i \neq j$, we have

$$\sum_{j\neq i}^{K} h_{ij}W_{j,0} \in \{-h_{\max}(K-1)N, ..., 0, ..., h_{\max}(K-1)N\}$$

and hence $H\left(\sum_{j\neq i} h_{ij}W_{j,0}\right) \leqslant \log(2h_{\max}KN)$. Since the $W_{j,0}$, $1 \leqslant j \leqslant K$, are i.i.d., we furthermore have

$H(h_{ii}W_{i,0}) \leqslant H(\sum_{j\neq i} h_{ij}W_{j,0})$ [12, Ex. 2.14] and (22) implies that

$$H\left(\sum_{j=1}^{K} h_{ij}W_{j,0}\right) \leqslant 2H\left(\sum_{j\neq i}^{K} h_{ij}W_{j,0}\right) \leqslant 2\log(2h_{\max}KN).$$

With (25) we therefore obtain

$$\min\left\{\frac{H\left(\sum_{j=1}^{K} h_{ij}W_{j,0}\right)}{\log(1/r)}, 1\right\} = \frac{H\left(\sum_{j=1}^{K} h_{ij}W_{j,0}\right)}{\log(1/r)}.$$

Application of Theorem 2 now yields that for each $\varepsilon > 0$, there is an $\widetilde{r} \in I_\varepsilon := [r - \varepsilon, r]$ such that the inputs $X_i = \sum_{k=0}^{\infty} \widetilde{r}^k W_{i,k}$, for $i = 1, ..., K$, satisfy

$$d\left(\sum_{j=1}^{K} h_{ij}X_j\right) - d\left(\sum_{j\neq i}^{K} h_{ij}X_j\right) = \frac{H(h_{ii}W_{i,0})}{\log(1/\widetilde{r})} = \frac{\log N}{\log(1/\widetilde{r})},$$

where we used (22). As ε can be made arbitrarily small, we find that

$$\mathsf{DoF}(\mathbf{H}) \geqslant \frac{K\log N}{\log(1/r)} = \frac{K\log N}{2\log(2h_{\max}KN)}. \qquad (26)$$

Letting $N \to \infty$, the right-hand side of (26) approaches $K/2$, which completes the proof.

We conclude by noting that the example class studied here was investigated before in [4, Thm. 1] and [5, Thm. 6]. In contrast to [4], [5] our proof of DoF-optimality is, however, not based on arguments from Diophantine approximation theory.

REFERENCES

[1] M. Hochman, "On self-similar sets with overlaps and inverse theorems for entropy," *Annals of Mathematics*, to appear.

[2] S. A. Jafar, "Interference alignment — A new look at signal dimensions in a communication network," *Foundations and Trends in Communications and Information Theory*, vol. 7, no. 1, 2011.

[3] A. S. Motahari, S. O. Gharan, M.-A. Maddah-Ali, and A. K. Khandani, "Real interference alignment: Exploiting the potential of single antenna systems," *arXiv:0908.2282 [cs.IT]*, Aug. 2009.

[4] R. H. Etkin and E. Ordentlich, "The degrees-of-freedom of the K-user Gaussian interference channel is discontinuous at rational channel coefficients," *IEEE Trans. Inf. Theory*, vol. 55, no. 11, pp. 4932–4946, Nov. 2009.

[5] Y. Wu, S. Shamai (Shitz), and S. Verdú, "A formula for the degrees of freedom of the interference channel," *Submitted to IEEE Trans. Inf. Theory*, revised version, Jun. 2013.

[6] C. Brandt, N. Viet Hung, and H. Rao, "On the open set condition for self-similar fractals," *Proc. of the AMS*, vol. 134, no. 5, pp. 1369–1374, Oct. 2005.

[7] D. Stotz and H. Bölcskei, "Degrees of freedom in vector interference channels," *Proc. 50th Ann. Allerton Conf. on Communication, Control, and Computing*, pp. 1755–1760, Oct. 2012.

[8] A. Guionnet and D. Shlyakhtenko, "On classical analogues of free entropy dimension," *Journal of Functional Analysis*, vol. 251, pp. 738–771, Oct. 2007.

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

[10] J. S. Geronimo and D. P. Hardin, "An exact formula for the measure dimensions associated with a class of piecewise linear maps," *Constructive Approximation*, vol. 5, pp. 89–98, Dec. 1989.

[11] A. Høst-Madsen and A. Nosratinia, "The multiplexing gain of wireless networks," *Proc. IEEE Int. Symp. on Inf. Theory*, pp. 2065–2069, Sep. 2005.

[12] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New York, NY: Wiley-Interscience, 2006.