

# Degrees of Freedom of the Interference Channel: a General Formula

Yihong Wu

Department of Electrical Engineering  
Princeton University  
Princeton, NJ 08544, USA  
Email: yihongwu@princeton.edu

Shlomo Shamai(Shitz)

Department of Electrical Engineering  
Technion, IIT  
Technion City, Haifa 32000, Israel  
Email: sshlomo@ee.technion.ac.il

Sergio Verdú

Department of Electrical Engineering  
Princeton University  
Princeton, NJ 08544, USA  
Email: verdu@princeton.edu

**Abstract**—THIS PAPER IS ELIGIBLE FOR THE STUDENT PAPER AWARD. We give a general formula for the degrees of freedom of the  $K$ -user real additive-noise interference channel involving maximization of information dimension. Previous results are recovered and improved with simplified proofs. Connections to fractal geometry are drawn.

## I. INTRODUCTION

Consider a  $K$ -user real memoryless Gaussian Interference Channel (IC) with a fixed deterministic channel matrix  $\mathbf{H} = [h_{ij}]$ , where the  $i^{\text{th}}$  user transmits  $X_i$  and receives

$$Y_i = \sum_{j=1}^K \sqrt{\text{snr}} h_{ij} X_j + N_i, \quad (1)$$

where  $\{X_i, N_i\}_{i=1}^K$  are independent with  $\mathbb{E}[X_i^2] \leq 1$  and  $N_i \sim \mathcal{N}(0, 1)$ .

Denote the the capacity region of (1) by  $\mathcal{C}(\mathbf{H}, \text{snr})$  with sum-rate capacity

$$\bar{C}(\mathbf{H}, \text{snr}) \triangleq \max \left\{ \sum_{i=1}^K R_i : R^K \in \mathcal{C}(\mathbf{H}, \text{snr}) \right\}. \quad (2)$$

The *degrees of freedom* (DoF) or the *multiplexing gain* is the pre-log of the sum-rate capacity in the high-SNR regime, defined by<sup>1</sup>

$$\text{DoF}(\mathbf{H}) = \lim_{\text{snr} \rightarrow \infty} \frac{\bar{C}(\mathbf{H}, \text{snr})}{\frac{1}{2} \log \text{snr}}. \quad (3)$$

Determining the degrees of freedom has been an active research subject. In [2] it is shown that  $\text{DoF}(\mathbf{H}) \leq \frac{K}{2}$  for *fully-connected*  $\mathbf{H}$ , i.e.,  $\mathbf{H}$  has no zero entries. Using Diophantine approximation, this upper bound is shown to be achievable for Lebesgue-almost every  $\mathbf{H}$  [3]. The almost sure achievability of  $\frac{K}{2}$  for vector interference channel with varying channel gains has been shown in [4]. Sufficient conditions on individual  $\mathbf{H}$  that guarantee  $\text{DoF}(\mathbf{H}) = \frac{K}{2}$  are also given in [1], [5]. On the converse side, based on additive-combinatorial results and deterministic channel approximation, it is shown that  $\text{DoF}(\mathbf{H}) < \frac{K}{2}$  if  $\mathbf{H}$  consists of all rational entries [1].

<sup>1</sup>In [1] the degrees of freedom is defined as a lim sup. In Theorem 1 we show that the limit in (3) always exists.

The goal of this paper is to give a general single-letter formula for  $\text{DoF}(\mathbf{H})$  via the maximization of a functional involving Rényi's *information dimension* [6]. This unified approach allows us to recover previously known results that are obtained using different methods, as well as uncover new results.

Our results apply to non-Gaussian noise as well. This is because the degrees of freedom are insensitive to the noise statistics as long as it has finite non-Gaussianness:  $D(N || \Phi_N) < \infty$ , where  $\Phi_N$  is a Gaussian random variable with the same mean and variance as  $N$ . In fact in our derivations we shall often assume that the noise is uniformly distributed on the unit interval.

Due to space limitations, omitted proofs are referred to [7].

## II. RÉNYI INFORMATION DIMENSION

A key concept in fractal geometry, Rényi [6] defined the *information dimension* (also known as the *entropy dimension* [8]) of a probability distribution. It measures the rate of growth of the entropy of successively finer discretizations.

**Definition 1.** Let  $X$  be a real-valued random variable. For  $m \in \mathbb{N}$ , let  $\langle X \rangle_m \triangleq \frac{\lfloor mX \rfloor}{m}$ . The information dimension of  $X$  is defined as

$$d(X) = \lim_{m \rightarrow \infty} \frac{H(\langle X \rangle_m)}{\log m}. \quad (4)$$

If the limit in (4) does not exist, the lim inf and lim sup are called lower and upper information dimensions of  $X$  respectively, denoted by  $\underline{d}(X)$  and  $\bar{d}(X)$ .

Definition 1 can be readily extended to random vectors, where the floor function  $\lfloor \cdot \rfloor$  is taken componentwise.

The information dimension of  $X$  is finite if and only if the mild condition

$$H(\lfloor X \rfloor) < \infty \quad (5)$$

is satisfied [9]. One sufficient condition for finite information dimension is  $\mathbb{E}[\log(1 + |X|)] < \infty$ , which is milder than finite variance. Therefore (5) is satisfied for all random variables considered in this paper.

Equivalent definitions of information dimension include:

- For an integer  $M \geq 2$ , write the  $M$ -ary expansion of  $X$  as

$$X = \lfloor X \rfloor + \sum_{i \in \mathbb{N}} (X)_i M^{-i}. \quad (6)$$

Then  $d(X)$  is the entropy rate of the digits  $\{(X)_i\}$  normalized by  $\log M$ .

- Denote by  $B(x, \epsilon)$  the  $\ell_\infty$ -ball of radius  $\epsilon$  centered at  $x$ . Then (see [10, Definition 4.2] and [9, Appendix A])

$$d(X) = \lim_{\epsilon \downarrow 0} \frac{\mathbb{E} \log P_X(B(X, \epsilon))}{\log \epsilon}. \quad (7)$$

The following are basic properties of information dimension [6], [9], the last three of which are inherited from Shannon entropy. For simplicity we assume all information dimensions involved exist.

**Lemma 1.**

- 

$$0 \leq \underline{d}(X^n) \leq \bar{d}(X^n) \leq n. \quad (8)$$

- Assume the distribution of  $X$  can be represented as

$$\nu = (1 - \rho)\nu_d + \rho\nu_c, \quad (9)$$

where  $\nu_d$  is a discrete measure,  $\nu_c$  is an absolutely continuous measure and  $0 \leq \rho \leq 1$ . Then

$$d(X) = \rho. \quad (10)$$

In particular, if  $X$  has a density with respect to the Lebesgue measure, then  $d(X) = 1$ ; if  $X$  is discrete, then  $d(X) = 0$ .

- Scale-invariance: for all  $\alpha \neq 0$ ,

$$d(\alpha X^n) = d(X^n). \quad (11)$$

- If  $X^n$  and  $Y^n$  are independent, then

$$\max\{d(X^n), d(Y^n)\} \leq d(X^n + Y^n) \quad (12)$$

$$\leq d(X^n) + d(Y^n). \quad (13)$$

- If  $\{X_i\}$  are independent and  $d(X_i)$  exists for all  $i$ , then

$$d(X^n) = \sum_{i=1}^n d(X_i). \quad (14)$$

- If  $X, Y, Z$  are independent, then

$$d(X + Y + Z) - d(Y + Z) \leq d(X + Z) - d(Z) \quad (15)$$

The high-SNR asymptotics of mutual information with additive noise is governed by the *input information dimension*. For notational convenience, denote

$$I(X, \text{snr}) \triangleq I(X; \sqrt{\text{snr}}X + N), \quad (16)$$

which is finite if and only if (5) holds [11]. Then [12]

$$\lim_{\text{snr} \rightarrow \infty} \frac{I(X, \text{snr})}{\frac{1}{2} \log \text{snr}} = d(X). \quad (17)$$

Therefore  $d(X)$  represents the single-user degrees of freedom when the input distribution is  $P_X$ . Naturally it also appears in

the characterization of degrees of freedom in the multi-user case. In fact, (17) also holds for random vectors [12].

Next we consider the behavior of information dimension under projections. Let  $\mathbf{A} \in \mathbb{R}^{m \times n}$  with  $m \leq n$ . Then for any  $X^n$ ,

$$d(\mathbf{A}X^n) \leq \min\{d(X^n), m\}. \quad (18)$$

Understanding how the dimension of a measure behaves under projections is a basic problem in fractal geometry. It is well-known that *almost every* projection preserves the dimension, be it Hausdorff dimension (Marstrand's projection theorem [13, Chapter 9]) or information dimension [10, Theorems 1.1 and 4.1]. However, computing the dimension for individual projections is in general difficult.

A problem closely related to determining  $\text{DoF}(\mathbf{H})$  is to determine the dimension difference of a product measure under two projections. Let  $p, q, p', q'$  be non-zero real numbers. By (12) and (13),

$$\begin{aligned} & d(pX + qY) - d(p'X + q'Y) \\ & \leq \frac{1}{2}(d(pX + qY) + d(p'X + q'Y) - d(X) - d(Y)) \quad (19) \end{aligned}$$

$$\leq \frac{1}{2}. \quad (20)$$

Therefore the dimension of a two-dimensional product measure under two projections can differ by at most *one half*. However, if the coefficients are all rational numbers, the dimension difference is *strictly* less than one half:

$$d(pX + qY) - d(p'X + q'Y) \leq \frac{1}{2} - \epsilon(p'q, pq'). \quad (21)$$

where  $\epsilon(a, b) = \frac{1}{12(a+b)+50}$ .

### III. MAIN RESULTS

#### A. A general formula

By the limiting characterization of interference channel capacity region [14], the sum-rate capacity is given by

$$\bar{C}(\mathbf{H}, \text{snr}) = \lim_{n \rightarrow \infty} \frac{1}{n} \sup_{X_1^n, \dots, X_K^n} \sum_{i=1}^K I(X_i^n; Y_i^n), \quad (22)$$

where  $X_i^n = [X_{i,1}, \dots, X_{i,n}]$  is the input of the  $i^{\text{th}}$  user, and the supremum is over independent  $X_1^n, \dots, X_K^n$ . Then

$$\begin{aligned} & I(X_i^n; Y_i^n) \\ & = I(X_1^n, \dots, X_K^n; Y_i^n) - I(X_1^n, \dots, X_K^n; Y_i^n | X_i^n) \quad (23) \end{aligned}$$

$$= I\left(\sum_j h_{ij} X_j^n, \text{snr}\right) - I\left(\sum_{j \neq i} h_{ij} X_j^n, \text{snr}\right). \quad (24)$$

Therefore the degrees of freedom admit the following limiting characterization:

$$\begin{aligned} \text{DoF}(\mathbf{H}) & = \lim_{\text{snr} \rightarrow \infty} \lim_{n \rightarrow \infty} \sup_{X_1^n, \dots, X_K^n} \frac{2}{n \log \text{snr}} \\ & \sum_{i=1}^K I\left\{\left(\sum_{j=1}^K h_{ij} X_j^n, \text{snr}\right) - I\left(\sum_{j \neq i} h_{ij} X_j^n, \text{snr}\right)\right\}, \quad (25) \end{aligned}$$

which can be single-letterized as follows:

**Theorem 1.** *Let*

$$\text{dof}(X^K, \mathbf{H}) \triangleq \sum_{i=1}^K d\left(\sum_{j=1}^K h_{ij} X_j\right) - d\left(\sum_{j \neq i} h_{ij} X_j\right). \quad (26)$$

Then

$$\text{DoF}(\mathbf{H}) = \sup_{X^K} \text{dof}(X^K, \mathbf{H}), \quad (27)$$

where the supremum is over independent  $X_1, \dots, X_K$  such that all information dimensions appearing in (26) exist.

The main difficulty in the converse proof lies in exchanging the supremum with the limits in (25), which amounts to proving that varying the input distribution with increasing SNR does not improve the degrees of freedom. To this end, we need the non-asymptotic version of (17) [7]: for any  $X^n$  and any  $\text{snr} > 0$ ,

$$-\frac{1}{2} \log \frac{e\pi}{6} \leq \frac{1}{n} I(X^n, \text{snr}) - \frac{1}{n} \mathbb{E} \left[ \log \frac{1}{P_{X^n}(B(X^n, \text{snr}^{-\frac{1}{2}}))} \right] \leq \log 2. \quad (28)$$

The basic idea to single-letterize (25) is as follows: Given any  $X^n$ , construct a single input  $X$  whose the first  $nM$  bits are formed by concatenating the first  $M$  bits from each  $X_i$ ; the remaining bits are independent copies of these  $nM$  bits. Then, the information dimension of  $X$  can be made close to  $\frac{1}{n} d(X^n)$  by choosing  $M$  sufficiently large. The same conclusion holds for information dimension of linear combinations.

### B. Corollaries

The following are immediate consequences of Theorem 1 combined with elementary properties of information dimension in Lemma 1:

- $\text{DoF}(\mathbf{H})$  is invariant under row or column scaling [1, Lemma 1], in view of (11).
- $\text{DoF}(\mathbf{H}) \leq K$ , with equality if and only if  $\mathbf{H}$  is a diagonal matrix with all diagonal entries nonzero.
- $\text{DoF}(\mathbf{H}) \geq 0$ , with equality if and only if  $\text{diag}(\mathbf{H}) = 0$ .
- Removing cross-links increases degrees of freedom: let  $\mathbf{H}'$  be obtained from  $\mathbf{H}$  by setting some of the off-diagonal entries to zero. By (15), for any independent  $X^K$ ,  $\text{dof}(X^K, \mathbf{H}) \leq \text{dof}(X^K, \mathbf{H}')$ . Therefore  $\text{DoF}(\mathbf{H}) \leq \text{DoF}(\mathbf{H}')$ .

Degrees of freedom of the following channels can be obtained by specializing Theorem 1.

- *Two-user IC:*

$$\text{DoF} \left( \begin{bmatrix} a & b \\ c & d \end{bmatrix} \right) = \begin{cases} 0 & a = d = 0 \\ 2 & a, d \neq 0, b = c = 0 \\ 1 & \text{otherwise} \end{cases} \quad (29)$$

- *Many-to-one IC:*  $\text{DoF}(\mathbf{H}) = K - 1$  where  $\mathbf{H}$  are all zero except for all diagonal entries and at least one off-diagonal entry in the first row. To see this, assuming

$h_{12} \neq 0$ , we have

$$\begin{aligned} & \text{dof}(X^K, \mathbf{H}) \\ &= d\left(\sum_j h_{1j} X_j\right) - d\left(\sum_{j \neq 1} h_{1j} X_j\right) + \sum_{j \neq 1} d(X_j) \end{aligned} \quad (30)$$

$$\leq d\left(\sum_j h_{1j} X_j\right) + \sum_{j=3}^K d(X_j) \quad (31)$$

$$\leq K - 1, \quad (32)$$

where (31) is due to (12) and  $K - 1$  is attained by choosing  $X_1$  discrete and the rest absolutely continuous.

- *One-to-many IC:* Using similar arguments, we obtain that  $\text{DoF}(\mathbf{H}) = K - 1$  where  $\mathbf{H}$  are all zero except for all entries on the diagonal and in the first column.
- *Multiple-access channel (MAC):* If  $\mathbf{H}$  is an all-one matrix, then  $\text{DoF}(\mathbf{H}) = 1$ , because

$$\text{dof}(X^K, \mathbf{H}) = K d\left(\sum X_j\right) - \sum_{i=1}^K d\left(\sum_{j \neq i} X_j\right) \quad (33)$$

$$\leq d\left(\sum X_j\right) \leq 1, \quad (34)$$

where we have used the following additive-combinatorial result [15, p. 3]:

$$(K - 1)H \left( \sum_{i=1}^K U_j \right) \leq \sum_{i=1}^K H \left( \sum_{j \neq i} U_j \right) \quad (35)$$

with  $\{U_i\}$  taking values on an arbitrary group. More generally,  $\text{DoF}(\mathbf{H}) \leq 1$  if all rows of  $\mathbf{H}$  are identical.

### C. Suboptimality of discrete-continuous mixture

In general discrete-continuous mixed input distributions are *strictly suboptimal*. In fact, discrete-continuous mixtures achieve *at most one* degree of freedom in the fully-connected case. To see this, let  $d(X_i) = \rho_i$ . If  $\mathbf{H}$  is fully-connected, then

$$\begin{aligned} & \text{dof}(X^K, \mathbf{H}) \\ &= \sum_{i=1}^K \left( 1 - \prod_j (1 - \rho_j) \right) - \left( 1 - \prod_{j \neq i} (1 - \rho_j) \right) \end{aligned} \quad (36)$$

$$= \sum_{i=1}^K \rho_i \prod_{j \neq i} (1 - \rho_j) \leq 1. \quad (37)$$

Therefore to obtain more than one degree of freedom, it is necessary to employ input distribution with *singular* components.

As we will show later, *singular* distributions of dimension *one half* are crucial in achieving the maximal degrees of freedom. Next we give a family of such distributions  $\{\mu_\lambda\}_{\lambda > 0}$  which are homogeneously self-similar [16]. For integer  $\lambda \geq 2$ ,  $\mu_\lambda$  is the distribution of a random variable whose  $\lambda$ -ary expansion has equiprobable even digits and zero odd digits. Then  $d(\mu_\lambda) = \frac{1}{2}$  in view of (6). For non-integer valued  $\lambda$ ,  $\mu_\lambda$  is defined as the invariant measure of an iterative function system [17].

#### D. Bounds and exact expressions

Next we prove that the degrees of freedom is upper bounded by  $\frac{K}{2}$  under more general conditions than fully-connectedness assumed in [2].

**Theorem 2.** *Let  $\pi$  be a fixed-point-free permutation on  $\{1, \dots, K\}$ , i.e.,  $\pi(i) \neq i$  for all  $i$ . If  $h_{\pi(i),i} \neq 0$  for each  $i$ , then*

$$\text{DoF}(\mathbf{H}) \leq \frac{K}{2}. \quad (38)$$

Moreover, if  $\pi$  is cyclic (i.e., consisting of one cycle), then  $\text{dof}(X^K, \mathbf{H}) = \frac{K}{2}$  if and only if for each  $i$ ,

$$d(X_i) = \frac{1}{2} \quad (39)$$

$$d\left(\sum_{j \neq i} h_{ij} X_j\right) = \frac{1}{2} \quad (40)$$

$$d\left(\sum_j h_{ij} X_j\right) = 1. \quad (41)$$

*Proof of (38):* Counting in different ways, we have

$$\begin{aligned} & 2 \text{dof}(X^K, \mathbf{H}) \\ &= \sum_{i=1}^K d\left(\sum_j h_{ij} X_j\right) + d\left(\sum_j h_{ij} X_j\right) \\ &\quad - d\left(\sum_{j \neq i} h_{ij} X_j\right) - d\left(\sum_{j \neq \pi(i)} h_{\pi(i)j} X_j\right) \quad (42) \end{aligned}$$

$$\leq K + \sum_{i=1}^K d\left(\sum_j h_{ij} X_j\right) - d\left(\sum_{j \neq i} h_{ij} X_j\right) - d(X_i) \quad (43)$$

$$\leq K, \quad (44)$$

where (43) follows from (8), (11) and (12), and (44) is due to (13). ■

Next we give various sufficient conditions on  $\mathbf{H}$  that guarantee  $\text{DoF}(\mathbf{H}) = \frac{K}{2}$ .

**Theorem 3.** *If the off-diagonal entries of  $\mathbf{H}$  are rational and the diagonal entries are irrational, then  $\text{DoF}(\mathbf{H}) = \frac{K}{2}$ .*

Theorem 3 implies the following previously known results, obtained using different methods:

- [1, Theorem 1], which relies on the Thue-Siegel-Roth theorem and requires the diagonal entries to be irrational algebraic numbers. Note that the set of real algebraic numbers is dense but countable. Therefore almost every real number is transcendental.
- [5, Theorem 1(2)], which assumes that the diagonal and off-diagonal entries are equal to one and  $h$  respectively, with  $h$  irrational. Upon scaling, this is equivalent to a channel matrix with unit off-diagonal entries and irrational diagonal entries  $h^{-1}$ .

**Theorem 4.**  $\text{DoF}(\mathbf{H}) = \frac{K}{2}$  for Lebesgue-a.e.  $\mathbf{H}$ .

Proof sketch of Theorem 4: construct input distributions depending only on the off-diagonal entries of  $\mathbf{H}$  such that (39)

and (40) are satisfied. Using the projection theorem [10] [10, Theorem 1] with respect to the diagonal entries, (41) holds for almost all  $h_{ii}$ .

The degrees of freedom of channel matrices with rational coefficients are strictly less than  $\frac{K}{2}$ . Next we give a sufficient condition for the three-user case.

**Theorem 5.** *Let  $K = 3$ . If there exists distinct  $i, j, k$ , such that  $h_{ij}, h_{ii}, h_{kj}$  and  $h_{ki}$  are non-zero rationals, then  $\text{DoF}(\mathbf{H}) < \frac{3}{2}$ .*

The following example illustrates the tightness of the condition in Theorem 5:

$$\text{DoF}\left(\begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 2 & 0 & 1 \end{bmatrix}\right) = \frac{3}{2}, \quad (45)$$

attained by choosing  $P_{X_1} = P_{X_2} = P_{X_3} = \mu_2$ , since  $d(X_1) = \frac{1}{2}$  and  $d(X_1 + 2X_2) = 1$ .

#### IV. DEGREES OF FREEDOM REGION

In addition to the sum-rate degrees of freedom, the degrees of freedom region  $\mathcal{DoF}(\mathbf{H})$  (see [4, Section II] for definition) is characterized by the following theorem:

**Theorem 6.**  $\mathcal{DoF}(\mathbf{H})$  is the collection of all  $r^K \in [0, 1]^K$ , such that for any probability vector  $w^K$ ,

$$\langle r^K, w^K \rangle \leq \sup_{X^K} \sum_{i=1}^K w_i d\left(\sum_j h_{ij} X_j\right) - w_i d\left(\sum_{j \neq i} h_{ij} X_j\right), \quad (46)$$

where the supremum is over independent  $X_1, \dots, X_K$  such that all information dimensions appearing in (46) exist.

The following results are analogous to Theorems 2 and 4:

**Theorem 7.** For any fully connected  $\mathbf{H}$ ,

$$\mathcal{DoF}(\mathbf{H}) \subset \text{co}\left\{\mathbf{e}_1, \dots, \mathbf{e}_K, \frac{1}{2}\mathbf{1}\right\}, \quad (47)$$

where  $\mathbf{e}_1, \dots, \mathbf{e}_K$  are standard basis. Moreover, (47) holds with equality for Lebesgue-a.e.  $\mathbf{H}$ .

#### V. AN EXAMPLE OF LOWER-TRIANGULAR CHANNEL MATRIX

Consider the following lower-triangular matrix [1, Section V]:

$$\mathbf{H}_\lambda \triangleq \begin{bmatrix} 1 & 0 & 0 \\ 1 & \lambda & 0 \\ 1 & 1 & 1 \end{bmatrix}. \quad (48)$$

**Theorem 8.**

$$\text{DoF}(\mathbf{H}_\lambda) = 1 + \sup_{X_1, X_2} d(X_1 + \lambda X_2) - d(X_1 + X_2). \quad (49)$$

Moreover,

- 1)  $\text{DoF}(\mathbf{H}_\lambda) \geq 1$  with equality if and only if  $\lambda = 0$  or 1;  
 $\text{DoF}(\mathbf{H}_\lambda) \leq \frac{3}{2}$  with equality if and only if  $\lambda$  is irrational.
- 2) For all  $\lambda \neq 0$ ,

$$\text{DoF}(\mathbf{H}_\lambda) = \text{DoF}(\mathbf{H}_{\lambda^{-1}}). \quad (50)$$

3) For integer  $\lambda \geq 2$ ,

$$\frac{3}{2} - \frac{1}{\lambda \log \lambda} \sum_{i=1}^{\lambda-1} i \log \frac{\lambda}{i} \leq \text{DoF}(\mathbf{H}_\lambda) \quad (51)$$

$$\leq \frac{3}{2} - \frac{1}{12 \log \lambda + 50} \quad (52)$$

Therefore as  $\lambda \rightarrow \infty$ ,  $\text{DoF}(\mathbf{H}_\lambda) = \frac{3}{2} + \Theta\left(\frac{1}{\log \lambda}\right)$ .

4) Let  $\phi = \frac{1+\sqrt{5}}{2}$  denote the golden ratio. Then

$$\text{DoF}(\mathbf{H}_2) \geq 1 + \log_6 \phi \approx 1.27. \quad (53)$$

Since  $\lambda$  is the channel gain of the direct link for the second user, it seems that  $\text{DoF}(\mathbf{H}_\lambda)$  should be increasing in  $|\lambda|$ . However, (50) shows that this is not the case.

*Proof of (49) and (53):*

$$\begin{aligned} & \text{dof}(X^3, \mathbf{H}_\lambda) \\ & \geq d(X_1) + d(X_1 + \lambda X_2) - d(X_1) \\ & \quad + d(X_1 + X_2 + X_3) - d(X_1 + X_2) \end{aligned} \quad (54)$$

$$= d(X_1 + X_2 + X_3) + d(X_1 + \lambda X_2) - d(X_1 + X_2). \quad (55)$$

To maximize (55), choosing an absolutely continuous  $P_{X_3}$  yields  $d(X_1 + X_2 + X_3) = 1$ , regardless of  $P_{X_1}$  or  $P_{X_2}$ . This proves (49). To achieve (53) for  $\lambda = 2$ , consider the following singular input distributions:

$$X_1 = \sum_{i \geq 1} U_i 6^{-i}, \quad X_2 = \sum_{i \geq 1} V_i 6^{-i}, \quad (56)$$

where  $\{(U_i, V_i)\}$  are i.i.d. copies of  $(U, V)$ , with  $U$  and  $V$  independently valued on  $\{0, 1\}$  and  $\{0, 1, 2\}$  respectively. Then  $X_1 + X_2 = \sum_{i \geq 1} (U_i + V_i) 6^{-i}$  and  $X_1 + 2X_2 = \sum_{i \geq 1} (U_i + 2V_i) 6^{-i}$ , where  $U + V$  and  $U + 2V$  are valued on  $\{0, \dots, 3\}$  and  $\{0, \dots, 5\}$  respectively. By the entropy-rate definition of information dimension in (6), we have

$$d(X_1 + X_2) = \frac{H(U + V)}{\log 6} \quad (57)$$

$$d(X_1 + 2X_2) = \frac{H(U + 2V)}{\log 6} = \frac{H(U) + H(V)}{\log 6}. \quad (58)$$

Next we maximize  $H(U) + H(V) - H(U + V) = H(U|U + V)$ . It can be shown that  $H(U|U + V)$  is concave in  $P_U$  and  $P_V$  individually. Moreover,  $H(U|U + V)$  is invariant if  $U$  is replaced by  $1 - U$  or  $V$  replaced by  $2 - V$ . Therefore the optimal  $U$  and  $V$  are symmetric. In particular,  $U$  is equiprobable Bernoulli. Let  $\mathbb{P}\{V = 0\} = q$ . Maximizing  $H(U|U + V)$  over  $0 \leq q \leq \frac{1}{2}$ , we obtain the optimal  $q = \frac{\phi}{\sqrt{5}}$  and  $H(U|U + V) = \log \phi$ , which, in view of (57)–(58), gives (53). ■

## VI. CONCLUSIONS

This paper gives a general formula for the degrees of freedom of the  $K$ -user interference channel in terms of a single-letter optimization (over  $K$ -dimensional input distributions) of a linear combination of information dimensions. Benefits of this method include:

- It recovers and improves known results with unified and simplified proofs, many of which are consequences of the calculus of information dimension. For instance, (21) and (52) are pure additive-combinatorial results, whereas the counterpart in [1] relies on additional techniques of deterministic channel approximation and the assumption of no repeated codewords.
- The power constraint becomes immaterial. In fact the same degrees of freedom holds even if  $\mathbb{E}[X^2] = \infty$ , as long as (5) is satisfied.
- It provides achievable rates (such as (53) which improves the lower bound  $\frac{2+\log_2 3}{3} \approx 1.19$  in [1, p. 4945]) that are not easily obtained from constructing explicit coding or communication schemes.

## REFERENCES

- [1] R. H. Etkin and E. Ordentlich, "The Degrees-of-Freedom of the  $K$ -User Gaussian Interference Channel Is Discontinuous at Rational Channel Coefficients," *IEEE Transactions on Information Theory*, vol. 55, no. 11, pp. 4932–4946, 2009.
- [2] A. Host-Madsen and A. Nosratinia, "The multiplexing gain of wireless networks," in *Proceedings of the 2005 International Symposium on Information Theory*, Adelaide, Australia, Sep. 2005, pp. 2065–2069.
- [3] A. Motahari, S. Gharan, M. Maddah-Ali, and A. Khandani, "Real interference alignment: Exploiting the potential of single antenna systems," submitted to *IEEE Transactions on Information Theory*, Nov. 2009.
- [4] V. R. Cadambe and S. A. Jafar, "Interference Alignment and the Degrees of Freedom for the  $K$ -User Interference Channel," *IEEE Transactions on Information Theory*, vol. 54, no. 8, pp. 3425 – 3441, Aug. 2008.
- [5] A. S. Motahari, A. K. Khandani, and S. O. Gharan, "On the Degrees of Freedom of the 3-user Gaussian interference channel: The symmetric case," in *Proceedings of the 2009 International Symposium on Information Theory*, Seoul, South Korea, Jun. 2009, pp. 1914–1918.
- [6] A. Rényi, "On the dimension and entropy of probability distributions," *Acta Mathematica Hungarica*, vol. 10, no. 1 – 2, Mar. 1959.
- [7] Y. Wu, S. Shamai(Shitz), and S. Verdú, "Degrees of freedom of the interference channel: a general formula," 2011, draft.
- [8] Y. Peres and B. Solomyak, "Existence of  $L_q$  dimensions and entropy dimension for self-conformal measures," *Indiana University Mathematics Journal*, no. 49, pp. 1603–1621, 2000.
- [9] Y. Wu and S. Verdú, "Rényi Information Dimension: Fundamental Limits of Almost Lossless Analog Compression," *IEEE Transactions on Information Theory*, vol. 56, no. 8, pp. 3721 – 3748, August 2010.
- [10] B. R. Hunt and V. Y. Kaloshin, "How projections affect the dimension spectrum of fractal measures," *Nonlinearity*, vol. 10, pp. 1031–1046, 1997.
- [11] Y. Wu and S. Verdú, "Functional Properties of MMSE and Mutual Information," submitted to *IEEE Transactions on Information Theory*, 2010.
- [12] A. Guionnet and D. Shlyakhtenko, "On classical analogues of free entropy dimension," *Journal of Functional Analysis*, vol. 251, no. 2, pp. 738 – 771, Oct. 2007.
- [13] P. Mattila, *Geometry of Sets and Measures in Euclidean Spaces: Fractals and Rectifiability*. Cambridge, UK: Cambridge University Press, 1999.
- [14] R. Ahlswede, "Multi-way communication channels," in *Proceedings of the Second International Symposium on Information Theory*, Tsahkadzor, Armenia, USSR, Sep. 1971, pp. 23–52.
- [15] M. Madiman, A. Marcus, and P. Tetali, "Entropy and set cardinality inequalities for partition-determined functions, with applications to sunsets," in revision for *Random Structures and Algorithms*, 2010.
- [16] Y. Peres and B. Solomyak, "Self-similar measures and intersections of Cantor sets," *Transactions of the American Mathematical Society*, vol. 350, no. 10, pp. 4065–4087.
- [17] K. Falconer, *Fractal Geometry: Mathematical Foundations and Applications*, 2nd ed. New York: Wiley, 2003.