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*—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 even generalized in certain cases with simplified proofs. Connections to fractal geometry are drawn.**

***Index Terms*—Shannon theory, interference channel, degrees of freedom, additive noise, Rényi information dimension.**

I. Introduction

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

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

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

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

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

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[1]

$$\mathsf{DoF}(\mathbf{H}) = \lim_{\mathsf{snr}\to\infty} \frac{\bar{C}(\mathbf{H}, \mathsf{snr})}{\frac{1}{2}\log \mathsf{snr}}. \tag{3}$$

Determining the degrees of freedom has been an active research subject. In [2] it is shown that $\mathsf{DoF}(\mathbf{H}) \leq \frac{K}{2}$ 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 $\mathsf{DoF}(\mathbf{H}) = \frac{K}{2}$ are also given in [1], [5]. On the converse side, based on additive-combinatorial results and deterministic channel approximation, [1] showed that $\mathsf{DoF}(\mathbf{H}) < \frac{K}{2}$ if $\mathbf{H}$ consists of all rational entries.

The goal of this paper is to give a general single-letter formula for $\mathsf{DoF}(\mathbf{H})$ 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. We also give results for the more general case in which the rates are, unlike (2), not equally weighted.

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 Φ_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.

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

$$d(X) = \lim_{m\to\infty} \frac{H\left(\langle X\rangle_m\right)}{\log m}. \tag{4}$$

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

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

$$H(\lfloor X\rfloor) < \infty \tag{5}$$

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

This work was partially supported by NSF under grant CCF-1016625 and by the Center for Science of Information, an NSF Science and Technology Center, under grant agreement CCF-0939370.

[1] In [1] the degrees of freedom is defined as a $\limsup$. In Theorem 1 we show that the limit in (3) always exists.

Equivalent definitions of information dimension include:

- For an integer $M \geq 2$, write the M-ary expansion of X as
$$X = \lfloor X \rfloor + \sum_{i \in \mathbb{N}} (X)_i M^{-i}. \tag{6}$$
Then $d(X)$ is the entropy rate of the digits $\{(X)_i\}$ normalized by $\log M$.
- Denote by $B(x, \epsilon)$ the ℓ_∞-ball of radius ϵ centered at x. Then (see [10, Definition 4.2] and [9, Appendix A])
$$d(X) = \lim_{\epsilon \downarrow 0} \frac{\mathbb{E} \log P_X(B(X, \epsilon))}{\log \epsilon}. \tag{7}$$

The following are basic properties of information dimension [6], [9], the last three of which are inherited from Shannon entropy.

Lemma 1.

- $$0 \leq \underline{d}(X^n) \leq \overline{d}(X^n) \leq n. \tag{8}$$
- *Assume the distribution of X can be represented as*
$$\nu = (1-\rho)\nu_{\rm d} + \rho \nu_{\rm c}, \tag{9}$$
where $\nu_{\rm d}$ is a discrete probability measure, $\nu_{\rm c}$ is an absolutely continuous probability measure and $0 \leq \rho \leq 1$. Then
$$d(X) = \rho. \tag{10}$$
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 $\alpha \neq 0$,*
$$d(\alpha X^n) = d(X^n). \tag{11}$$
- *If X^n and Y^n are independent, then*
$$\begin{aligned} \max\{d(X^n), d(Y^n)\} &\leq d(X^n + Y^n) && (12)\\ &\leq d(X^n) + d(Y^n). && (13) \end{aligned}$$
- *If $\{X_i\}$ are independent and $d(X_i)$ exists for all i, then*
$$d(X^n) = \sum_{i=1}^n d(X_i). \tag{14}$$
- *If X, Y, Z are independent, then*
$$d(X+Y+Z) + d(Z) \leq d(X+Z) + d(Y+Z). \tag{15}$$

The high-SNR asymptotics of mutual information with additive noise is governed by the *input information dimension*. For convenience, denote
$$I(X, \mathsf{snr}) \triangleq I(X; \sqrt{\mathsf{snr}} X + N), \tag{16}$$
which is finite if and only if (5) holds [11]. Then [12]
$$\lim_{\mathsf{snr} \to \infty} \frac{I(X, \mathsf{snr})}{\frac{1}{2} \log \mathsf{snr}} = d(X). \tag{17}$$
Therefore $d(X)$ represents the single-user degrees of freedom when the input distribution is constrained to be P_X. Naturally information dimension, as we will see, 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,
$$d(\mathbf{A} X^n) \leq \min\{d(X^n), \mathsf{rank}(\mathbf{A})\}. \tag{18}$$
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 $\mathsf{DoF}(\mathbf{H})$ 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),
$$\begin{aligned} & d(pX + qY) - d(p'X + q'Y) \\ \leq & \frac{1}{2}(d(pX + qY) + d(pX + qY) - d(X) - d(Y)) && (19)\\ \leq & \frac{1}{2}. && (20) \end{aligned}$$
Therefore the dimension of a two-dimensional product measure under two projections can differ by at most *one half*. However, if the coefficients are rational, then the dimension difference is *strictly* less than one half [7]:
$$d(pX + qY) - d(p'X + q'Y) \leq \frac{1}{2} - \epsilon(p'q, pq'). \tag{21}$$
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
$$\bar{C}(\mathbf{H}, \mathsf{snr}) = \lim_{n \to \infty} \frac{1}{n} \sup_{X_1^n, \ldots, X_K^n} \sum_{i=1}^K I(X_i^n; Y_i^n), \tag{22}$$
where $X_i^n = [X_{i,1}, \ldots, X_{i,n}]$ is the input of the $i^{\rm th}$ user, and the supremum is over independent $X_1^n, \ldots, X_K^n$. Then
$$\begin{aligned} & I(X_i^n; Y_i^n) \\ &= I(X_1^n, \ldots, X_K^n; Y_i^n) - I(X_1^n, \ldots, X_K^n; Y_i^n | X_i^n) && (23)\\ &= I\left(\sum\nolimits_j h_{ij} X_j^n, \mathsf{snr}\right) - I\left(\sum\nolimits_{j \neq i} h_{ij} X_j^n, \mathsf{snr}\right). && (24) \end{aligned}$$
Therefore the degrees of freedom admit the following limiting characterization:[2]
$$\begin{aligned} \mathsf{DoF}(\mathbf{H}) = & \lim_{\mathsf{snr} \to \infty} \lim_{n \to \infty} \sup_{X_1^n, \ldots, X_K^n} \frac{2}{n \log \mathsf{snr}} \\ & \sum_{i=1}^K I\left\{\left(\sum_{j=1}^K h_{ij} X_j^n, \mathsf{snr}\right) - I\left(\sum_{j \neq i} h_{ij} X_j^n, \mathsf{snr}\right)\right\}, \end{aligned} \tag{25}$$

[2]The second limit in (25) can be replaced by supremum over $n \in \mathbb{N}$.

Our main result is the single-letterization of (25):

Theorem 1. *Let*

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

Then

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

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 invoke the following non-asymptotic version of (17) whose proof can be found in [7]: for any X^n and any $\mathsf{snr} > 0$,

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

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 the 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 the information dimensions of linear combinations.

B. Corollaries

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

- $\mathsf{DoF}(\mathbf{H})$ is invariant under row or column scaling [1, Lemma 1], in view of (11).
- $\mathsf{DoF}(\mathbf{H}) \leq K$, with equality if and only if $\mathbf{H}$ is a diagonal matrix with all diagonal entries nonzero.
- $\mathsf{DoF}(\mathbf{H}) \geq 0$, with equality if and only if $\mathsf{diag}(\mathbf{H}) = 0$.
- 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, $\mathsf{dof}(X^K, \mathbf{H}) \leq \mathsf{dof}(X^K, \mathbf{H}')$. Therefore $\mathsf{DoF}(\mathbf{H}) \leq \mathsf{DoF}(\mathbf{H}')$.

As we illustrate next, the degrees of freedom of various channels can be obtained by specializing Theorem 1.

- *Two-user IC*:

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

- *Many-to-one IC*: $\mathsf{DoF}(\mathbf{H}) = K - 1$ 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

$$\begin{aligned} &\mathsf{dof}(X^K, \mathbf{H}) \\ &= d\left(\sum\nolimits_j h_{ij} X_j\right) - d\left(\sum\nolimits_{j \neq 1} h_{ij} X_j\right) + \sum_{j \neq 1} d(X_j) \quad (30) \\ &\leq d\left(\sum\nolimits_j h_{ij} X_j\right) + \sum_{j=3}^{K} d(X_j) \quad (31) \\ &\leq K - 1, \quad (32) \end{aligned}$$

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

- *One-to-many IC*: Using similar arguments, we obtain that $\mathsf{DoF}(\mathbf{H}) = K - 1$ 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 $\mathsf{DoF}(\mathbf{H}) = 1$, because

$$\begin{aligned} \mathsf{dof}(X^K, \mathbf{H}) &= K d\Big(\sum_{j=1}^{K} X_j\Big) - \sum_{i=1}^{K} d\Big(\sum_{j \neq i} X_j\Big) \quad (33) \\ &\leq d\Big(\sum\nolimits_{j=1}^{K} X_j\Big) \leq 1, \quad (34) \end{aligned}$$

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

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

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

C. Suboptimality of discrete-continuous mixture

Discrete-continuous mixed input distributions are usually *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

$$\mathsf{dof}(X^K, \mathbf{H}) = \sum_{i=1}^{K} \Big(1 - \prod_{j=1}^{K} (1 - \rho_j)\Big) - \Big(1 - \prod_{j \neq i} (1 - \rho_j)\Big) \quad (36)$$

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

Therefore to obtain more than one degree of freedom, it is necessary to employ input distributions 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 $\lambda \geq 2$, μ_λ is the distribution of a random variable whose λ-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 λ, μ_λ is defined as the invariant measure of an iterative function system [17].

D. Bounds and exact expressions

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

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

$$\mathsf{DoF}(\mathbf{H}) \leq \frac{K}{2}. \tag{38}$$

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

$$d(X_i) = d\left(\sum\nolimits_{j\neq i} h_{ij}X_j\right) = \frac{1}{2} \tag{39}$$

$$d\left(\sum\nolimits_{j=1}^{K} h_{ij}X_j\right) = 1. \tag{40}$$

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

$$\begin{aligned}
&2\,\mathsf{dof}(X^K,\mathbf{H})\\
&= \sum_{i=1}^{K} d\left(\sum\nolimits_{j} h_{ij}X_j\right) + d\left(\sum\nolimits_{j} h_{ij}X_j\right)\\
&\qquad - d\left(\sum\nolimits_{j\neq i} h_{ij}X_j\right) - d\left(\sum\nolimits_{j\neq \pi(i)} H_{\pi(i)j}X_j\right) &(41)\\
&\leq K + \sum_{i=1}^{K} d\left(\sum\nolimits_{j} h_{ij}X_j\right) - d\left(\sum\nolimits_{j\neq i} h_{ij}X_j\right) - d(X_i) &(42)\\
&\leq K, &(43)
\end{aligned}$$

where (42) follows from (8), (11) and (12), and (43) is due to (13). ■

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

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

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 ([3])**.** $\mathsf{DoF}(\mathbf{H}) = \frac{K}{2}$ *for Lebesgue-a.e.* $\mathbf{H}$.

Proof sketch of Theorem 4 based on (27): construct input distributions depending only on the off-diagonal entries of $\mathbf{H}$ such that (39) are satisfied. Using the projection theorem [10, Theorem 1] with respect to the diagonal entries, (40) 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 $\mathsf{DoF}(\mathbf{H}) < \frac{3}{2}$.*

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

$$\mathsf{DoF}\left(\begin{bmatrix}1 & 2 & 0\\ 0 & 1 & 2\\ 2 & 0 & 1\end{bmatrix}\right) = \frac{3}{2}, \tag{44}$$

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. An example of lower-triangular channel matrix

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

$$\mathbf{H}_\lambda \triangleq \begin{bmatrix}1 & 0 & 0\\ 1 & \lambda & 0\\ 1 & 1 & 1\end{bmatrix}. \tag{45}$$

Theorem 6.

$$\mathsf{DoF}(\mathbf{H}_\lambda) = 1 + \sup_{X_1,X_2} d(X_1 + \lambda X_2) - d(X_1 + X_2). \tag{46}$$

Moreover,

1) $\mathsf{DoF}(\mathbf{H}_\lambda) \geq 1$ *with equality if and only if $\lambda = 0$ or 1;* $\mathsf{DoF}(\mathbf{H}_\lambda) \leq \frac{3}{2}$ *with equality if and only if λ is irrational.*
2) *For all $\lambda \neq 0$,*
$$\mathsf{DoF}(\mathbf{H}_\lambda) = \mathsf{DoF}(\mathbf{H}_{\lambda^{-1}}). \tag{47}$$
3) *For integer $\lambda = 2, 3, 4, \ldots$,*
$$\frac{3}{2} - \frac{1}{\lambda\log\lambda}\sum_{i=1}^{\lambda-1}\frac{i}{\lambda}\log\frac{\lambda}{i} \leq \mathsf{DoF}(\mathbf{H}_\lambda) \tag{48}$$
$$\leq \frac{3}{2} - \frac{1}{12\log\lambda + 50} \tag{49}$$
Therefore as $\lambda \to \infty$, $\mathsf{DoF}(\mathbf{H}_\lambda) = \frac{3}{2} - \Theta\left(\frac{1}{\log\lambda}\right)$. For $\lambda = 2$, (48) can be sharpened to
$$\mathsf{DoF}(\mathbf{H}_2) \geq 1 + \log_6\phi \approx 1.27. \tag{50}$$
where $\phi = \frac{1+\sqrt{5}}{2}$ denotes the golden ratio.

Since λ is the channel gain of the direct link for the second user, it seems that $\mathsf{DoF}(\mathbf{H}_\lambda)$ should be increasing in $|\lambda|$. However, (47) shows that this is not the case.

Proof of (46) and (50):

$$\begin{aligned}
&\mathsf{dof}(X^3,\mathbf{H}_\lambda)\\
&\geq d(X_1) + d(X_1+\lambda X_2) - d(X_1)\\
&\quad + d(X_1+X_2+X_3) - d(X_1+X_2) &(51)\\
&= d(X_1+X_2+X_3) + d(X_1+\lambda X_2) - d(X_1+X_2). &(52)
\end{aligned}$$

To maximize (52), 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 (46). To achieve (50) for $\lambda = 2$, consider the following singular input distributions:

$$X_1 = \sum_{i\geq 1} U_i 6^{-i}, \quad X_2 = \sum_{i\geq 1} V_i 6^{-i}, \tag{53}$$

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 $\{0, \ldots, 3\}$ and $\{0, \ldots, 5\}$ respectively. By the entropy-rate definition of information dimension in (6), we have

$$d(X_1 + X_2) = \frac{H(U + V)}{\log 6} \tag{54}$$

$$d(X_1 + 2X_2) = \frac{H(U + 2V)}{\log 6} = \frac{H(U) + H(V)}{\log 6}. \tag{55}$$

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 (54)–(55), gives (50). ■

V. Degrees of Freedom Region

In addition to the sum-rate degrees of freedom, the degrees of freedom region $\mathcal{DoF}(\mathbf{H})$ obtained by allowing different weights for the rates in (2) (see [4, Section II] for definition) is characterized by the following theorem:

Theorem 7. *$\mathcal{DoF}(\mathbf{H})$ is the collection of all $r^K \in [0, 1]^K$, such that for any probability vector w^K,*

$$\langle r^K, w^K\rangle \leq \sup_{X^K} \sum_{i=1}^{K} w_i d\left(\sum\nolimits_j h_{ij} X_j\right) - w_i d\left(\sum\nolimits_{j\neq i} h_{ij} X_j\right), \tag{56}$$

where the supremum is over independent $X_1, \ldots, X_K$ such that all information dimensions appearing in (56) exist.

The following results are analogous to Theorems 2 and 4:

Theorem 8. *For any fully connected $\mathbf{H}$,*

$$\mathcal{DoF}(\mathbf{H}) \subset \mathbf{co}\left\{\mathbf{e}_1, \ldots, \mathbf{e}_K, \frac{1}{2}\mathbf{1}\right\}, \tag{57}$$

where $\mathbf{e}_1, \ldots, \mathbf{e}_K$ are standard basis and $\mathbf{co}$ denotes the convex hull. Moreover, (57) holds with equality for Lebesgue-a.e. $\mathbf{H}$.

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 (49) are pure additive-combinatorial results, whereas the counterpart in [1] relies on additional techniques of deterministic channel approximation.
- 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 (50) 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 Trans. Inf. 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. S. Motahari, S. O. Gharan, M. A. Maddah-Ali, and A. K. Khandani, "Real interference alignment: Exploiting the potential of single antenna systems," submitted to *IEEE Trans. on Inf. Theory*, Nov. 2009. [Online]. Available: http://arxiv.org/abs/0908.2282

[4] V. R. Cadambe and S. A. Jafar, "Interference Alignment and the Degrees of Freedom for the K-User Interference Channel," *IEEE Trans. Inf. 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*, vol. 49, no. 4, pp. 1603–1621, 2000.

[9] 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.

[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*, Tsahkadsor, 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 sumsets," 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, 1998.

[17] K. Falconer, *Fractal Geometry: Mathematical Foundations and Applications*, 2nd ed. New York: Wiley, 2003.