Degrees of Freedom in Vector Interference Channels

David Stotz and Helmut Bölcskei
Dept. of IT & EE, ETH Zurich, Switzerland
Email: {dstotz, boelcskei}@nari.ee.ethz.ch

***Abstract*—This paper continues the Wu-Shamai-Verdú program [1] on characterizing the degrees of freedom (DoF) of interference channels (ICs) through Rényi information dimension. Concretely, we find a general formula for the DoF of vector ICs, encompassing multiple-input multiple-output (MIMO) ICs, time- and/or frequency-selective ICs, and combinations thereof, as well as constant single-antenna ICs considered in [1]. As in the case of constant single-antenna ICs, achieving full DoF requires the use of singular input distributions. Strikingly, in the vector case it suffices to enforce singularity on the joint distribution of individual transmit vectors. This can be realized through signaling in subspaces of the ambient signal space, which is in accordance with the idea of interference alignment, and, most importantly, allows the scalar components of the transmit vectors to have non-singular distributions. We recover the result by Cadambe and Jafar on the non-separability of parallel ICs [2] and we show that almost all parallel ICs are separable. Finally, our results extend the main finding in [1] to the complex case.**

I. Introduction

Sparked by the surprising finding of Cadambe and Jafar [3] stating that $K/2$ degrees of freedom (DoF) can be realized in K-user interference channels (ICs) through interference alignment, the study of DoF in wireless networks has seen significant activity in recent years. The essence of interference alignment is to exploit channel variations in time/frequency/space to align interference in low-dimensional subspaces. This is accomplished through symbol extension and a clever choice of the resulting vector-valued transmit signals.

Following the discovery in [3] it turned out that the basic idea of aligning interference to realize DoF can be applied to numerous further settings [4]. Most surprisingly, it was shown in [5], [6] that $K/2$ DoF can be realized in K-user constant single-antenna ICs, i.e., in complete absence of channel variations. The schemes introduced in [5], [6], referred to in [5] as "real interference alignment", use Diophantine approximation or lattice structures to design the transmit signals and rely on number-theoretic properties of the channel coefficients to achieve $K/2$ DoF for almost all channels [5].

Wu et al. [1] showed recently that (Rényi) information dimension [7] is a suitable tool for characterizing the DoF achievable in constant single-antenna ICs in a unified manner. Concretely, based on a single-letter characterization of the DoF of the constant single-antenna IC, it is shown in [1] that the real interference alignment schemes proposed in [5], [6] correspond to the use of singular input distributions. Strikingly, Wu et al. [1] found that full, i.e., $K/2$ in a K-user IC, DoF can be achieved *only* by input distributions that have a singular component. This is a strong, negative, result as input distributions with a singular component are not realizable in practice. On the other hand, it is well known that full DoF can be realized in ICs with channel variations in time/frequency or in constant multiple-antenna ICs, even if the marginals of the transmit vectors do not have singular components. It is therefore of interest to reconcile these two lines of results. We will see that singularity of input distributions plays a fundamental role in achieving full DoF in ICs with channel variations and/or multiple-antenna users as well, albeit in a much more benign form.

Contributions: We continue the Wu-Shamai-Verdú program [1] by showing how information dimension can be used to systematically characterize the DoF of vector ICs. This extension, mentioned in [1] as an avenue of interesting further research, is relevant as "classical" interference alignment relies on vector-valued signaling to realize full DoF. The vector IC considered in this paper contains, as special cases, the MIMO IC, time- and/or frequency-selective ICs, and combinations thereof, as well as the constant single-antenna IC studied in [1], [5], [6]. The Wu-Shamai-Verdú theory builds on a little known, but highly useful result by Guionnet and Shlyakthenko [8, Thm. 2.7], stating that the DoF achieved by a particular input distribution in a scalar additive noise channel are given by the input distribution's information dimension. We present an extension of this result to the vector case, formalized in Theorem 2 in Section IV, with the corresponding proof provided in [9]. This extension is the key ingredient of our main result, a general formula for the DoF in vector ICs.

The formula is in the spirit of [1] and allows us to conclude that, while input distributions with a singular component are still needed to achieve full DoF, it suffices to enforce singularity on the joint distribution of individual transmit vectors. This form of singularity is easy to realize in practice by taking, e.g., the transmit vectors to live in lower-dimensional subspaces of the ambient signal space, as is, in fact, done in interference alignment. We illustrate, by way of selected examples, how the main result in this paper allows to treat different interference alignment concepts in a unified manner. Furthermore, we recover the result by Cadambe and Jafar on the non-separability of parallel ICs [2] and we show that almost all parallel ICs are separable in terms of DoF. Finally, our results apply to complex channel matrices and signal vectors, thereby also extending the main result in [1] to the complex case. Most results in this paper are stated without proof. Omitted proofs are reported in [9].

Notation: Random vectors are represented by uppercase letters, deterministic vectors by lowercase letters, in both cases using letters from the end of the alphabet. Boldface uppercase letters are used to indicate matrices. The diagonal $n \times n$ matrix with entries $a_1, ..., a_n$ on its main diagonal is denoted by $\mathrm{diag}(a_1, ..., a_n)$. For $x \in \mathbb{R}$, we write $\lfloor x \rfloor$ for the largest integer not greater than x. For $k \in \mathbb{N} \setminus \{0\}$, we set $\langle x \rangle_k := \lfloor kx \rfloor / k$. The last two notation conventions are extended to real vectors by applying $\lfloor \cdot \rfloor$ and $\langle \cdot \rangle_k$ componentwise. For a discrete random vector X, we write $H(X)$ for its entropy. All logarithms are to the base e. $\mathbb{E}[\cdot]$ denotes the expectation operator.

II. Setup and Definitions

We consider a memoryless K-user Gaussian vector IC with input-output relation

$$Y_i = \sqrt{\mathsf{snr}} \sum_{j=1}^{K} \mathbf{H}_{i,j} X_j + W_i, \quad i = 1, ..., K, \tag{1}$$

where $X_i = (X_i[1] \dots X_i[M])^T \in \mathbb{R}^M$ and $Y_i = (Y_i[1] \dots Y_i[M])^T \in \mathbb{R}^M$ is the transmit and receive vector, respectively, corresponding to user i, M is the dimension of the input/output signal space, $\mathbf{H}_{i,j} \in \mathbb{R}^{M \times M}$ denotes the channel matrix between transmitter j and receiver i, and $(W_i)_{i=1,...,K}$ are i.i.d. zero mean Gaussian random vectors with identity covariance matrix.[1] Defining the $KM \times KM$ matrix

$$\mathbf{H} := \begin{pmatrix} \mathbf{H}_{1,1} & \cdots & \mathbf{H}_{1,K} \\ \vdots & \ddots & \vdots \\ \mathbf{H}_{K,1} & \cdots & \mathbf{H}_{K,K} \end{pmatrix},$$

we can rewrite (1) as

$$\begin{pmatrix} Y_1 \\ \vdots \\ Y_K \end{pmatrix} = \sqrt{\mathsf{snr}}\, \mathbf{H} \begin{pmatrix} X_1 \\ \vdots \\ X_K \end{pmatrix} + \begin{pmatrix} W_1 \\ \vdots \\ W_K \end{pmatrix}. \tag{2}$$

The channel matrix $\mathbf{H}$ is assumed to be known perfectly at all transmitters and receivers. We impose the average power constraint[2]

$$\frac{1}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} \left(x_i^{(n)}[m] \right)^2 \leqslant 1$$

on codeword matrices $\left(x_i^{(1)} \dots x_i^{(N)} \right)$ of block length N transmitted by user $i = 1, ..., K$, where $x_i^{(n)} = \left(x_i^{(n)}[1] \dots x_i^{(n)}[M] \right)^T \in \mathbb{R}^M$.

This setting encompasses the MIMO IC with M antennas at each user and $\mathbf{H}_{i,j}$ denoting the MIMO channel matrix between transmitter j and receiver i. It furthermore contains time- and/or frequency-selective ICs, where intersymbol/intercarrier interference is transformed into a memoryless vector channel through the use of guard periods/bands. When all the $\mathbf{H}_{i,j}$ are diagonal, obtained, e.g., when OFDM is used in frequency-selective channels [10], the model reduces to that of M parallel ICs. The constant single-antenna setting considered in [1] is covered by the special case $M = 1$. In contrast to the channel model in [1], where the action of each subchannel is represented through scaling by a single coefficient, in our setting each subchannel acts as a linear operator (represented by a matrix) on the corresponding transmit vector.

Let $\overline{C}(\mathbf{H}; \mathsf{snr})$ be the sum-capacity[3] for a given snr. The degrees of freedom[4] of the IC (1) are then defined as

$$\mathsf{DoF}(\mathbf{H}) := \limsup_{\mathsf{snr} \to \infty} \frac{\overline{C}(\mathbf{H}; \mathsf{snr})}{\frac{1}{2} \log \mathsf{snr}}. \tag{3}$$

We call $\mathsf{DoF}(\mathbf{H})/M$ the normalized DoF.

III. Rényi Information Dimension

It was recognized in [1] that information dimension is a suitable tool for characterizing the DoF in constant single-antenna ICs. The main conceptual contribution of the present paper is to show that information dimension is a natural tool for analyzing the DoF in vector ICs as well. This eventually leads to a unified framework for real interference alignment results as in [5], [6], that hinge on intricate properties of input distributions (singularity) and channel coefficients (number-theoretic), and "classical" interference alignment results, that rely on channel variations and vector signaling [3].

We start by collecting basic properties of information dimension used in the remainder of the paper.

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

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

respectively. 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.

Lemma 1 ([7],[1]):

1) Let X be a random vector in $\mathbb{R}^n$. Then $\underline{d}(X), \overline{d}(X) < \infty$ if and only if $H(\lfloor X \rfloor) < \infty$. Moreover, in this case, we have

$$0 \leqslant \underline{d}(X) \leqslant \overline{d}(X) \leqslant n.$$

2) Let $X_1, ..., X_K$ be independent random vectors, such that $d(X_i)$ exists for $i = 1, ..., K$. Then

$$d \begin{pmatrix} X_1 \\ \vdots \\ X_K \end{pmatrix} = \sum_{i=1}^{K} d(X_i). \tag{4}$$

3) Let X and Y be independent random vectors in $\mathbb{R}^n$. Then

$$\max\{\overline{d}(X), \overline{d}(Y)\} \leqslant \overline{d}(X + Y). \tag{5}$$

4) Let X be a random vector in $\mathbb{R}^n$ and $F\colon \mathbb{R}^n \to \mathbb{R}^n$ a linear isomorphism. Then

$$\overline{d}(FX) = \overline{d}(X). \tag{6}$$

[1] We show in Section V-C how the results in this paper can be extended to the complex case.

[2] Note that the average is taken across vector symbols.

[3] For a precise definition of $\overline{C}(\mathbf{H}; \mathsf{snr})$ see [11, Sect. 6.1].

[4] For motivation on why to study this quantity see [4, App. A].

The statements (5) and (6) also hold for lower information dimension.

Remark 1: An immediate consequence of (4) and (6) is the following. Suppose $(v_1, ..., v_n)$ are linearly independent vectors in $\mathbb{R}^n$ and $\widetilde{X}_1, ..., \widetilde{X}_n$ are independent random variables with $d(\widetilde{X}_i) = 1$, for all i. Then

$$d\Big(\widetilde{X}_1 v_1 + \ldots + \widetilde{X}_n v_n\Big) \overset{(6)}{=} d\begin{pmatrix}\widetilde{X}_1\\ \vdots \\ \widetilde{X}_n\end{pmatrix} \overset{(4)}{=} \sum_{i=1}^{n} d(\widetilde{X}_i) = n. \quad (7)$$

The nature of general probability distributions on $\mathbb{R}^n$ can often be understood better by invoking the following decomposition result.

Proposition 1 (e.g., [12, Thm. 2.7.19, Ex. 2.9.14]): Every probability distribution μ on $\mathbb{R}^n$ can be decomposed uniquely as

$$\mu = \alpha\mu_{\text{ac}} + \beta\mu_{\text{d}} + \gamma\mu_s,$$

where μ_{ac} is an absolutely continuous, μ_{d} a discrete, and μ_{s} a singular probability distribution,[5] and $\alpha, \beta, \gamma \geqslant 0$ satisfy $\alpha + \beta + \gamma = 1$.

Information dimension is typically difficult to compute analytically. However, for probability distributions that are discrete-continuous mixtures, there is a simple formula, which is an extension of [7, Thm. 3] to the vector case.

Proposition 2: Let X be a random vector in $\mathbb{R}^n$ with $H(\lfloor X \rfloor) < \infty$ and probability distribution μ that admits a decomposition $\mu = \alpha\mu_{\text{ac}} + (1-\alpha)\mu_{\text{d}}$, where μ_{ac} is an absolutely continuous and μ_{d} a discrete probability distribution, and $\alpha \in [0,1]$. Then $d(X)$ exists and is given by

$$d(X) = n\alpha. \quad (8)$$

Another class of distributions admitting an explicit formula for information dimension is given by so called self-similar homogeneous distributions [14]. As in [1], these special singular distributions, suitably extended to the vector case, play a central role in the proof of our general DoF-formula for vector ICs. For a more detailed discussion of this matter we refer to [9].

IV. Main Result

Theorem 1 (DoF-formula): For the channel (1) we have

$$\mathsf{DoF}(\mathbf{H}) = \sup_{X_1,\ldots,X_K} \mathsf{dof}(X_1, ..., X_K; \mathbf{H}), \quad (9)$$

where

$$\mathsf{dof}(X_1, ..., X_K; \mathbf{H}) := \sum_{i=1}^{K} d\left(\sum_{j=1}^{K} \mathbf{H}_{i,j} X_j\right) - d\left(\sum_{j\neq i}^{K} \mathbf{H}_{i,j} X_j\right),$$

and where the supremum is taken over all independent $X_1, ..., X_K$ such that $H(\lfloor X_i \rfloor) < \infty$, for all i, and such that all appearing information dimensions exist.[6]

The proof of Theorem 1, as the proof of the corresponding result for the constant single-antenna IC in [1], builds on two results, whose extensions to the vector case are provided next. The first one relates information dimension to the high-snr asymptotics of mutual information in additive noise vector channels.

Theorem 2: Let X and W be independent random vectors such that W has an absolutely continuous distribution with $h(W) > -\infty$ and $H(\lfloor W \rfloor) < \infty$. Then

$$\limsup_{\mathsf{snr}\to\infty} \frac{I(X; \sqrt{\mathsf{snr}}X + W)}{\frac{1}{2}\log \mathsf{snr}} = \overline{d}(X).$$

Remark 2: It is not hard to show that $\mathbb{E}[X^T X] < \infty$ implies $H(\lfloor X \rfloor) < \infty$ [9]. Consequently, finite second moment of X and W and $h(W) > -\infty$ is sufficient for the conditions in Theorem 2 to be satisfied.

For the scalar case, Theorem 2 was proved in [8, Thm. 2.7], under the slightly stronger assumptions $h(W) > -\infty$, $\mathbb{E}[1 + |W|] < \infty, \mathbb{E}[1 + |X|] < \infty$. Wu [15, Sect. 2.7] extended the result in [8, Thm. 2.7] to scalar W with finite non-Gaussianness, also allowing $H(\lfloor X \rfloor) = \infty$. Particularizing Theorem 2 above to the scalar case also covers certain noise distributions with infinite non-Gaussianness, e.g., W can have a Cauchy distribution.

The second result we build on is the following multi-letter characterization of the vector IC sum-capacity, which is obtained by extending [16, Sect. 2: Lem. 1] to continuous vector alphabets with an average power constraint.

Proposition 3: The sum-capacity of the channel (1) is

$$\overline{C}(\mathbf{H}; \mathsf{snr}) = \lim_{N\to\infty} \frac{1}{N} \sup_{\mathbf{X}_1^N,\ldots,\mathbf{X}_K^N} \sum_{i=1}^{K} I\big(\mathbf{X}_i^N; \mathbf{Y}_i^N\big),$$

where $\mathbf{X}_i^N = \Big(X_i^{(1)} \ldots X_i^{(N)}\Big)$ for random vectors $X_i^{(n)}$ in $\mathbb{R}^M$ and the supremum is taken over all independent $\mathbf{X}_1^N, ..., \mathbf{X}_K^N$ satisfying

$$\frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N} \mathbb{E}\left[\Big(X_i^{(n)}[m]\Big)^2\right] \leqslant 1, \quad \text{for all } i.$$

V. Implications of the DoF-Formula

We now discuss implications of Theorem 1.

Proposition 4: Suppose there is a permutation σ of $\{1, ..., K\}$ such that $\sigma(i) \neq i$ and $\det \mathbf{H}_{\sigma(i),i} \neq 0$, for all i. Then we have

$$\frac{\mathsf{DoF}(\mathbf{H})}{M} \leqslant \frac{K}{2}. \quad (10)$$

Proposition 4, which is the extension of [1, Thm. 5] to vector ICs, implies that for almost all channel matrices no more than $K/2$ normalized DoF can be achieved.

[5] "Absolutely continuous" is to be understood with respect to Lebesgue measure. A discrete distribution is one that is supported on a countable set of points, whereas a singular distribution is concentrated on a set of Lebesgue measure zero and, in addition, does not have any point masses. An example of a singular distribution on $\mathbb{R}$ is the Cantor distribution [13, Ex. 1.2.4].

[6] It turns out that it is not necessary to restrict the supremum to inputs satisfying the power constraint [9].

A. Singular distributions

It was observed in [1] that input distributions with a singular component play a central role in achieving full DoF, i.e., the supremum in (9), for the constant single-antenna IC. The natural extension of this result to the vector case holds, albeit with vastly different consequences as we shall see next.

Proposition 5: Suppose $X_1, ..., X_K$ are independent random vectors in $\mathbb{R}^M$, whose probability distributions are discrete-continuous mixtures. If $\det \mathbf{H}_{i,j} \neq 0$, for all i, j, then

$$\frac{\mathsf{dof}(X_1, ..., X_K; \mathbf{H})}{M} \leqslant 1. \tag{11}$$

The condition $\det \mathbf{H}_{i,j} \neq 0$, for all i, j, which is stronger than the condition of Proposition 4, is satisfied for almost all $\mathbf{H}$. Proposition 5 has far-reaching consequences as it says that restricting the transmit vectors to have distributions that do not exhibit a singular component, we can achieve no more than one normalized DoF. Singularity for scalar transmit symbols, as mandated in the constant single-antenna case, is impossible to realize in practice. In the case of vector ICs, however, a striking new feature appears. Singularity of the transmit vectors can be realized by simply choosing input distributions that are (continuously) supported on lower-dimensional subspaces of the ambient signal space. The following example illustrates this point.

Example 1: Let $K = 3$, $M = 2$, and take

$$\mathbf{H} = \left(\begin{array}{cc|cc|cc} 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 1 \\ \hline 1 & 0 & 1 & 0 & 1 & 0 \\ 2 & 2 & 0 & 1 & 1 & 1 \\ \hline 1 & 0 & 2 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{array}\right).$$

Choose the transmit vectors as

$$X_1 := \widetilde{X}_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad X_2 := \widetilde{X}_2 \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad X_3 := \widetilde{X}_3 \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \tag{12}$$

where $\widetilde{X}_1, \widetilde{X}_2, \widetilde{X}_3$ are independent random variables with absolutely continuous distributions and $H(\lfloor \widetilde{X}_i \rfloor) < \infty$, for $i = 1, 2, 3$. Note that the vectors X_1, X_2, X_3 are continuously distributed on lines in 2-dimensional space and hence have probability distributions supported on sets of Lebesgue measure zero; this renders the distributions singular. Therefore $\mathsf{dof}(X_1, X_2, X_3; \mathbf{H})/2$ is not bound by (11). Indeed, a simple calculation reveals that

$$\begin{aligned} &\mathsf{dof}(X_1, X_2, X_3; \mathbf{H}) = \\ &\quad d\left(\widetilde{X}_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + (\widetilde{X}_2 + \widetilde{X}_3) \begin{pmatrix} 1 \\ 3 \end{pmatrix}\right) - d\left((\widetilde{X}_2 + \widetilde{X}_3) \begin{pmatrix} 1 \\ 3 \end{pmatrix}\right) \\ &+ d\left(\widetilde{X}_2 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + (\widetilde{X}_1 + \widetilde{X}_3) \begin{pmatrix} 1 \\ 4 \end{pmatrix}\right) - d\left((\widetilde{X}_1 + \widetilde{X}_3) \begin{pmatrix} 1 \\ 4 \end{pmatrix}\right) \\ &+ d\left(\widetilde{X}_3 \begin{pmatrix} 4 \\ 3 \end{pmatrix} + (\widetilde{X}_1 + 2\widetilde{X}_2) \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) - d\left((\widetilde{X}_1 + 2\widetilde{X}_2) \begin{pmatrix} 1 \\ 1 \end{pmatrix}\right) \\ &= 2 - 1 + 2 - 1 + 2 - 1 = 3, \end{aligned}$$

and hence $\mathsf{DoF}(\mathbf{H})/2 \geqslant 3/2$. Here, we used (7), (5), and (8). Since, by Proposition 4, $\mathsf{DoF}(\mathbf{H})/2 \leqslant 3/2$, the transmit vectors in (12) achieve the supremum in (9), and hence full DoF. We emphasize that while the distributions of the transmit vectors in this example are singular, their marginals are non-singular; what is more, the marginals can be taken to be absolutely continuous, e.g., to have a Gaussian distribution. The reader can readily verify that this example follows the idea of "classical" interference alignment, as put forward in [3].

B. Parallel interference channels

The system model (2) contains parallel ICs as a special case, obtained when $\mathbf{H}_{i,j} = \operatorname{diag}(h_{i,j}[1], ..., h_{i,j}[M])$, for all i, j. For simplicity of exposition, we introduce the notation

$$\mathbf{H}[m] := \begin{pmatrix} h_{1,1}[m] & \cdots & h_{1,K}[m] \\ \vdots & \ddots & \vdots \\ h_{K,1}[m] & \cdots & h_{K,K}[m] \end{pmatrix}, \quad m = 1, ..., M.$$

Here $\mathbf{H}[m]$ is the interference matrix of the m-th subchannel, which is a constant single-antenna IC.

It was shown in [2] that there exist non-separable parallel ICs, i.e., parallel ICs whose capacity region is strictly larger than the (Minkowski) sum of the capacity regions of the individual subchannels. This phenomenon stands in contrast to the Gaussian multiple access and the Gaussian broadcast channel, both of which are separable [2]. In this paper, we deal with the (weaker) notion of separability in the sense of DoF, i.e., the question of whether full DoF can be achieved by transmit vectors that have independent components across subchannels.

Proposition 6: For parallel ICs, we have

$$\mathsf{DoF}(\mathbf{H}) \geqslant \sum_{m=1}^{M} \mathsf{DoF}(\mathbf{H}[m]). \tag{13}$$

There exist examples where the inequality in (13) is strict.

Proof: The inequality in (13) follows by restricting the supremization (9) to inputs X_i with independent components $X_i[1], ..., X_i[M]$, such that all information dimensions in $\mathsf{dof}(X_1[m], ..., X_K[m]; \mathbf{H}[m])$ exist, for $m = 1, ..., M$. Evaluating $\mathsf{dof}(X_1, ..., X_K; \mathbf{H})$ for this class of inputs yields

$$\begin{aligned} &\mathsf{dof}(X_1, ..., X_K; \mathbf{H}) \\ &= \sum_{i=1}^{K} d\begin{pmatrix} \sum_{j=1}^{K} h_{i,j}[1] X_j[1] \\ \vdots \\ \sum_{j=1}^{K} h_{i,j}[M] X_j[M] \end{pmatrix} - d\begin{pmatrix} \sum_{j\neq i}^{K} h_{i,j}[1] X_j[1] \\ \vdots \\ \sum_{j\neq i}^{K} h_{i,j}[M] X_j[M] \end{pmatrix} \\ &\overset{(4)}{=} \sum_{m=1}^{M} \sum_{i=1}^{K} d\left(\sum_{j=1}^{K} h_{i,j}[m] X_j[m]\right) - d\left(\sum_{j\neq i}^{K} h_{i,j}[m] X_j[m]\right) \\ &= \sum_{m=1}^{M} \mathsf{dof}(X_1[m], ..., X_K[m]; \mathbf{H}[m]), \end{aligned}$$

which proves (13). A class of examples that renders the inequality in (13) strict, and is different from the example reported in [2], is obtained as follows. Take $K = 3$ and $M = 2$ with

$$\mathbf{H}[m] = \begin{pmatrix} 1 & 0 & 0 \\ 1 & \lambda[m] & 0 \\ 1 & 1 & 1 \end{pmatrix}, \quad m = 1, 2,$$

where $\lambda[1]$, $\lambda[2]$ are nonzero, rational, and satisfy $\lambda[1] \neq \lambda[2]$. Using [1, Cor. 1], we get $\mathsf{DoF}(\mathbf{H}[m]) < 3/2$, for $m = 1, 2$. Now, consider the transmit vectors

$$X_1 := \widetilde{X}_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad X_2 := \widetilde{X}_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad X_3 := \widetilde{X}_3 \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

where $\widetilde{X}_1, \widetilde{X}_2, \widetilde{X}_3$ are independent random variables with $d(\widetilde{X}_i) = 1$, for $i = 1, 2, 3$. A calculation similar to the one in Example 1 reveals that $\mathsf{dof}(X_1, X_2, X_3; \mathbf{H}) = 3$, and therefore[7]

$$\mathsf{DoF}(\mathbf{H}) \geqslant 3, \tag{14}$$

which proves that $\mathsf{DoF}(\mathbf{H}) > \mathsf{DoF}(\mathbf{H}[1]) + \mathsf{DoF}(\mathbf{H}[2])$. ∎

For parallel ICs that satisfy the condition of Proposition 4, e.g., when all subchannel matrices are fully connected,[8] we find through application of Proposition 6 that

$$\frac{\mathsf{DoF}(\mathbf{H}[1]) + \ldots + \mathsf{DoF}(\mathbf{H}[M])}{M} \overset{(13)}{\leqslant} \frac{\mathsf{DoF}(\mathbf{H})}{M} \overset{(10)}{\leqslant} \frac{K}{2}. \tag{15}$$

Since we know that $\mathsf{DoF}(\mathbf{H}[m]) = K/2$ for almost all $\mathbf{H}[m]$ [5, Thm. 1], it follows from (15) that $\mathsf{DoF}(\mathbf{H})/M = K/2$ for almost all parallel ICs. Moreover, the left-hand side of (15) is achieved by transmit vectors with independent components across subchannels, i.e., joint coding across subchannels is not needed and therefore almost all parallel ICs are separable. The downside is, however, that by Proposition 5, the distributions of the scalar components of these transmit vectors are required to have a singular component. On the other hand, if we allow joint coding across subchannels, i.e., the use of dependent symbols across subchannels for the individual transmit vectors, full DoF can be achieved without having these symbols contain singular components, as the example in the proof of Proposition 6 demonstrates. We illustrate our findings through an example.

Example 2: We consider a parallel IC with $K = 3$, $M = 2$, and fully connected $\mathbf{H}[1]$ and $\mathbf{H}[2]$. The fully connected constant single-antenna IC (i.e., $M = 1$) was studied in [5] where it was shown that $K/2 = 3/2$ DoF are achievable for almost all channel matrices. However, the result in [5] is not constructive in the sense of providing explicit conditions on the channel matrix for the underlying IC to have $K/2 = 3/2$ DoF. In the following, we provide such explicit conditions on $\mathbf{H}[1]$ and $\mathbf{H}[2]$. We also identify conditions on $\mathbf{H}[1]$ and $\mathbf{H}[2]$ for the underlying IC to have strictly less than $3/2$ normalized DoF. We start by noting that $\mathsf{DoF}(\mathbf{H})$ is invariant to scaling of rows and columns of $\mathbf{H}[m]$, for any m [9]. It therefore suffices to consider the channel matrices

$$\mathbf{H}[m] = \begin{pmatrix} a[m] & 1 & 1 \\ 1 & b[m] & 1 \\ 1 & d[m] & c[m] \end{pmatrix}, \quad m = 1, 2, \tag{16}$$

for $m = 1, 2$ (called the "standard 3-user IC matrix" in [5, Sec. VI-B]). Since $\mathbf{H}[1]$ and $\mathbf{H}[2]$ are fully connected, $\det \mathbf{H}_{i,j} \neq 0$, for all i, j, and thus, by Proposition 4, we have $\mathsf{DoF}(\mathbf{H})/2 \leqslant 3/2$.

1) If $d[m]$ is rational and $a[m], b[m], c[m]$ are irrational, for $m = 1, 2$, then $\mathsf{DoF}(\mathbf{H}[1]) = \mathsf{DoF}(\mathbf{H}[2]) = 3/2$ by [1, Thm. 6]. Using Proposition 6, it then follows that

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

Here full DoF are achieved by transmit vectors that have independent (across subchannels) scalar components.

2) If $d[1] = d[2]$ and $a[1] \neq a[2]$, $b[1] \neq b[2]$, $c[1] \neq c[2]$, choosing each transmit vector to be continuously distributed along $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, a calculation similar to the one in Example 1 reveals that $\mathsf{dof}(X_1, X_2, X_3; \mathbf{H}) = 3$, and hence

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

This is achieved by joint coding, i.e., by using transmit vectors with scalar components that are dependent across subchannels.

3) If one of the entries on the main diagonal in (16) is rational and identical across the two subchannels, then

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

This follows from an extension of [1, Thm. 8], see [9].

C. The complex case

The results stated so far apply to real signals and channel matrices. We next outline the extension of our main result, Theorem 1, to the complex case, thereby also providing an extension of the main result in [1] to complex alphabets.

For a vector IC with transmit and receive signals in $\mathbb{C}^M$ and $\mathbf{H}_{i,j} \in \mathbb{C}^{M \times M}$, simply stack the real and imaginary parts of the transmit, receive, and noise vectors in (1), and stack the real and imaginary parts of the matrices $\mathbf{H}_{i,j}$ correspondingly. Then, apply the DoF-formula (9) to the resulting vector IC in $\mathbb{R}^{2M}$. Defining information dimension of a complex random vector accordingly as the information dimension of the vector obtained by stacking real and imaginary parts, we find that Theorem 1 applies directly to the complex case.

VI. Acknowledgments

We would like to thank G. Pope, R. Heckel, C. Bunte, Y. Wu, S. Verdú, and S. Shamai (Shitz) for inspiring discussions. Advice from A. Guionnet on the proof of her result [8, Thm. 2.7] is gratefully acknowledged.

[7] A more careful analysis shows that, in fact, we have equality in (14).
[8] We call a matrix fully connected when all its entries are nonzero.

REFERENCES

[1] Y. Wu, S. Shamai (Shitz), and S. Verdú, "A general formula for the degrees of freedom of the interference channel," *Submitted to IEEE Trans. Inf. Theory*, Aug. 2011.

[2] V. R. Cadambe and S. A. Jafar, "Parallel Gaussian interference channels are not always separable," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 3983–3990, Sep. 2009.

[3] ——, "Interference alignment and degrees of freedom of the K-user interference channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.

[4] 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.

[5] 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. Inf. Theory, arXiv:0908.2282v2*, vol. cs.IT, Nov. 2009.

[6] 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.

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

[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] D. Stotz and H. Bölcskei, "Sensitivity of degrees of freedom in interference channels," *in preparation*.

[10] A. Peled and A. Ruiz, "Frequency domain data transmission using reduced computational complexity algorithms," *Proc. IEEE Int. Conf. on Acoustics, Speech, and Sig. Proc.*, vol. 5, pp. 964–967, Apr. 1980.

[11] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge, UK: Cambridge University Press, 2012.

[12] J. Taylor, *An Introduction to Measure and Probability*. New York, NY: Springer, 1997.

[13] R. Durrett, *Probability: Theory and Examples*. New York, NY: Cambridge University Press, 2010.

[14] J. E. Hutchinson, "Fractals and self similarity," *Indiana University Mathematics Journal*, vol. 30, pp. 713–747, 1981.

[15] Y. Wu, "Shannon theory for compressed sensing," *Ph.D. dissertation, Princeton University, NJ*, Sep. 2011.

[16] R. Ahlswede, "Multi-way communication channels," *Proc. of the 2nd Int. Symp. on Inf. Theory*, pp. 23–52, Sep. 1971.