The Degrees-of-Freedom of the K-User Gaussian Interference Channel Is Discontinuous at Rational Channel Coefficients

Raúl H. Etkin, *Member, IEEE*, and Erik Ordentlich, *Senior Member, IEEE*

***Abstract*—The degrees-of-freedom of a K-user Gaussian interference channel (GIC) has been defined to be the multiple of $(1/2)\log_2 P$ at which the maximum sum of achievable rates grows with increasing power P. In this paper, we establish that the degrees-of-freedom of three or more user, real, scalar GICs, viewed as a function of the channel coefficients, is discontinuous at points where all of the coefficients are nonzero rational numbers. More specifically, for all $K > 2$, we find a class of K-user GICs that is dense in the GIC parameter space for which $K/2$ degrees-of-freedom are exactly achievable, and we show that the degrees-of-freedom for any GIC with nonzero rational coefficients is strictly smaller than $K/2$. These results are proved using new connections with number theory and additive combinatorics.**

***Index Terms*—Additive combinatorics, interference alignment, lattices, sum sets.**

I. Introduction

THE time-invariant, real, scalar K-user Gaussian interference channel (GIC), as introduced in [2], involves K transmitter–receiver pairs in which each transmitter attempts to communicate a uniformly distributed, finite-valued message to its corresponding receiver by sending a signal comprised of n real numbers. Each receiver observes a component-wise linear combination of possibly *all* of the transmitted signals plus additive memoryless Gaussian noise, and seeks to decode, with probability close to one, the message of its corresponding transmitter, in spite of the interfering signals and noise. The time averages of the squares of the transmitted signal values are required to not exceed certain power constraints. A K-tuple of rates $(R_1, \ldots, R_K)$ is said to be achievable for a GIC if the transmitters can increase the sizes of their message sets as 2^{nR_i} with the signal length n, and signal in such a way that the power constraints are met and the receivers are able to correctly decode their corresponding messages with probability converging to 1, as n grows to infinity. The set of all achievable K-tuples of rates is known as the capacity region of the GIC. Determining it, as a function of the channel coefficients (specifying the linear combinations mentioned above), power constraints, and noise variances, has been an open problem in information theory for over 30 years.

Manuscript received June 18, 2008; revised July 11, 2009. Current version published October 21, 2009. The material in this paper was presented in part at the IEEE International Symposium on Information Theory (ISIT), Seoul, Korea, July 2009.

The authors are with Hewlett-Packard Laboratories, Palo Alto, CA 94304 USA (e-mail: raul.etkin@hp.com; erik.ordentlich@hp.com).

Communicated by G. Kramer, Associate Editor for Shannon Theory.

Digital Object Identifier 10.1109/TIT.2009.2030473

A complete solution for even the two-user case, which has received the most attention to date, is still out of reach. The best known coding scheme for two users is that presented in [3]. In some ranges of channel coefficients, such as for strong interference, the capacity region is completely known for two users [2]. Still for other ranges, the maximum achievable sum-of-rates is known [4]–[7]. For the general two-user case, the strongest known result is that of [8], which determines the capacity region to within a 1/2 bit margin (1 bit for the complex case) using a carefully chosen version of the scheme of [3], and a new genie-aided outer bound. Recently, the outer bound of [8] was improved in [5]–[7] and [9], but the worst case 1/2 bit gap between inner and outer bounds of [8] persists.

The case of $K > 2$ users has, until very recently, received less attention. Much of the recent effort on $K > 2$, beginning with [10] and continuing in, e.g., [12], [13] has focused on characterizing the growth of the capacity region in the limit of increasing signal-to-noise ratio (SNR) corresponding, for example, to fixing the noise variances and channel coefficients and letting the power constraints tend to infinity. Specific attention has been directed at the growth of the maximum sum of achievable rates. If there were no interference, the maximum achievable rate corresponding to each transmitter–receiver pair would grow like $(1/2)\log_2 P$ in the limit of increasing power, which follows from the well-known formula of $(1/2)\log_2(1 + P/N)$ for the capacity of a single-user additive Gaussian noise channel with power constraint P and noise variance N. Thus, the maximum sum of achievable rates would grow as $(K/2)\log_2 P$ if there were no interference. This motivates the expectation that, in the general case, the maximum sum of achievable rates would grow as $(d/2)\log_2 P$ for some constant $d \leq K$, depending on the channel coefficients, where d has been dubbed the degrees-of-freedom of the underlying GIC. Although determining d for a given GIC is, in principle, simpler than determining the capacity region, it has turned out to be a difficult problem in its own right, for $K > 2$.[1]

A positive development in the study of the degrees-of-freedom of GICs with more than two users has been the discovery of a new coding technique known as *interference alignment*, which involves carefully choosing the transmitted signals so that the interfering signals "align" benignly at each receiver [11], [12]. Interference alignment has been shown, under some conditions which we summarize below, to achieve nearly $d = K/2$ degrees-of-freedom, which is half of the

[1]The degrees-of-freedom is known to be 1 for all two-user GICs, unless there is no interference [10].

degrees-of-freedom in the case of no interference at all. Interference alignment is not possible to implement for two users[2] and became relevant only as the focus shifted to more users. Another new phenomenon in network information theory that has recently emerged as the number of users studied was increased, is the technique of *indirect decoding*, which is crucial for achieving the capacity region of certain three-user broadcast channels [15]. This technique is not relevant in the two-user case, and could not have been discovered in the study thereof.

In this paper, we find a new information-theoretic phenomenon concerning interference channels that is not manifest in the two-user case. In particular, we find that the degrees-of-freedom (and therefore the capacity region at high SNR) of real, scalar GICs with $K > 2$ users is very sensitive to whether the channel coefficients determining the linear combinations of signals at each receiver are rational or irrational numbers. Next, we formally explain our results and their significance in the context of the growing literature on $K > 2$ user GICs.

We shall use a matrix H to denote the direct and cross gains of a time-invariant, real, scalar K-user (GIC) [2] with the (i,j)th entry $h_{i,j}$ specifying the channel gain from transmitter i to receiver j. Thus, the signal observed by receiver $j \in \{1,\dots,K\}$ at time index $t = 1,2,\dots$ is given by $z_{j,t}+\sum_{i=1}^{K} x_{i,t}h_{i,j}$ where $x_{i,t}$ is the real-valued signal of transmitter $i \in \{1,\dots,K\}$ at time t and $z_{j,t}$ is additive Gaussian noise with variance σ_j^2, independent across time and users. Fixing a block length n, the transmitted signals $\{x_{i,t}\}$ are required to satisfy the average power constraints $\sum_{t=1}^{n} x_{i,t}^2 \leq nP_i$ for some collection of powers $P_1,\dots,P_K$. For $H \in \mathbb{R}^{K\times K}$, and $\boldsymbol{\sigma}, \boldsymbol{P} \in \mathbb{R}_+^K$, we let $\mathcal{C}(H,\boldsymbol{\sigma},\boldsymbol{P})$ denote the capacity region of a GIC with gain matrix H, receiver noise variances given by the corresponding components of $\boldsymbol{\sigma}$, and average (per codeword) power constraints given by the components of $\boldsymbol{P}$. Following [10], we define the degrees-of-freedom of H as

$$\mathrm{DoF}(H) = \limsup_{P\to\infty} \frac{\max_{\boldsymbol{R}\in\mathcal{C}(H,\mathbf{1},P\mathbf{1})} \mathbf{1}^t\boldsymbol{R}}{(1/2)\log_2 P} \tag{1}$$

where $\mathbf{1}$ denotes the vector of all ones. The degrees-of-freedom of a GIC characterizes the behavior of the maximum achievable sum–rate as the SNR tends to infinity, with the gain matrix fixed.

A fully connected GIC is one for which $h_{i,j} \neq 0$ for all i and j. It was shown in [10] that for fully connected H, $\mathrm{DoF}(H) \leq K/2$. If H is not fully connected, the degrees-of-freedom can be as high as K, such as when H is the identity matrix where all cross gains are zero. Little was known about tightness of the $K/2$ bound for $K > 2$ until it was shown in [12] that for *vector* GICs and an appropriate generalization of $\mathrm{DoF}(\cdot)$ to include a normalization by the input/output vector dimension, the degrees-of-freedom of "almost all" fully connected vector GICs approaches $K/2$ when the vector dimension tends to infinity.[3] In addition, an example of a fully connected two-dimensional vector GIC achieving exactly $K/2$ degrees-of-freedom was also given in [12]. The key tool used in [12] to establish these results is the technique of interference alignment, which involves the transmitters signaling over linear subspaces that, after component-wise scaling by the cross gains, *align* into interfering subspaces which are linearly independent with the directly received subspaces, allowing for many interference-free dimensions over which to communicate. For real, scalar GICs, it was shown in [13], using a different type of interference alignment, that the degrees-of-freedom of certain fully connected GICs also approaches $K/2$ when the cross gains tend to zero. Yet a different type of interference alignment is used in [14] to find new achievable rates for a non-fully connected GIC in which interference occurs only at one receiver. To our knowledge, the problem of determining or computing the degrees-of-freedom of general GICs is still open.

As in [13], in this paper we consider only fully connected scalar, real GICs and establish the following results on the degrees-of-freedom.

Theorem 1: If all diagonal components of a fully connected H are irrational algebraic numbers and all off-diagonal components are rational numbers then $\mathrm{DoF}(H) = K/2$.

Theorem 2: For $K > 2$, if all elements of a fully connected H are rational numbers then $\mathrm{DoF}(H) < K/2$.

The following corollary is then immediate from Theorems 1 and 2, and the well-known fact that irrational algebraic numbers are dense in the real numbers.

Corollary 1: For $K > 2$, the function $\mathrm{DoF}(H)$ is discontinuous at all fully connected H with rational components.

Theorem 1 demonstrates the existence of fully connected, real K-user GICs with exactly $K/2$ degrees-of-freedom. In contrast to the result of [13], Theorem 1 is nonasymptotic (in H). The underlying achievability scheme is based on an interference alignment phenomenon that differs from the ones used in [12] and [13], and relies on number-theoretic lower bounds on the approximability of irrational algebraic numbers by rationals.

Theorem 2 reveals a surprising limitation on $\mathrm{DoF}(H)$ when the components are nonzero rational numbers (up to arbitrary pre-post multiplication by diagonal matrices—see Lemma 1 in Section III). In this case, $\mathrm{DoF}(H)$ is strictly bounded away from $K/2$. Previously known techniques for finding outer bounds to the capacity regions of GICs, such as cooperative encoding and decoding [16], [17], genie-aided decoding [8], [18], [19], and multiple-access bounds [2], [20] are not sensitive to the rationality of the channel parameters and hence do not suffice to establish Theorem 2. Instead, our proof of this theorem is based on a new connection between GICs with rational H and results from additive combinatorics [23], a branch of combinatorics that is concerned with the cardinalities of sum sets, or sets obtained by adding (assuming an underlying group structure) any element of a set A to any element of a set B.

The remainder of the paper is organized as follows. The next section clarifies some notation and gives the formal definition of the capacity region of a GIC that will apply in this paper. In Section III, we present the proof of Theorem 1. This is followed by the proof of Theorem 2 in Section IV, which further consists of subsections collecting various intermediate results. Each of

[2]We refer specifically to two-user interference channels. As shown in [11], interference alignment is beneficial in the two-user multiple-antenna X-channel, which is an interference channel where each transmitter sends a message to each receiver (i.e., four messages instead of two).

[3]The $K/2$ bound of [10] extends to the fully connected vector case as well.

these sections is prefaced with a high level outline of the respective proofs. In Section V, we determine lower and upper bounds on $\mathrm{DoF}(H)$ for a simple three-user rational H by improving on the scheme of [13], and evaluating an upper bound implicit in the proof of Theorem 2. We conclude in Section VI with some final observations and directions for future work.

II. Notation and Definitions

We adopt the usual notation for the information-theoretic quantities of discrete and differential entropy (resp., $H(X)$ and $h(X)$), and mutual information ($I(X;Y)$), which shall all be measured in bits (i.e., involve logarithms to the base two) [21]. We shall use the standard notation $\lceil x \rceil$ and $\lfloor x \rfloor$ to, respectively, denote the smallest integer not smaller than x and the greatest integer not larger than x. The cardinality of a set $\mathcal{A}$ shall be denoted as $|\mathcal{A}|$.

Next, we review the definition of the capacity region of a K-user GIC with power constraints $\boldsymbol{P} = (P_1, \ldots, P_K)$ and noise variances $\boldsymbol{\sigma} = (\sigma_1^2, \ldots, \sigma_K^2)$ that will apply in this paper. Fixing a block length n and a rate-tuple $R_1, \ldots, R_K$, the random message W_i of the ith transmitter is assumed to be uniformly distributed in the set $\mathcal{W}_i \stackrel{\Delta}{=} \{1, \ldots, 2^{\lceil nR_i \rceil}\}$. The messages are further assumed to be independent from one user to the next. A coding scheme consists of K encoding functions $f_1, \ldots, f_K$ where f_i maps $\mathcal{W}_i$ into the n-dimensional ball of radius $\sqrt{nP_i}$ of real vectors, the components of which specify the signal value $x_{i,t}$ that the ith transmitter will send at each time index.[4] The set $\{f_i(1), f_i(2), \ldots, f_i(2^{\lceil nR_i \rceil})\}$ constitutes the codebook of transmitter i. There is also a corresponding set of K decoding functions $g_1, \ldots, g_K$ where g_i maps n-dimensional real vectors into the message set $\mathcal{W}_i$. The function g_i is applied by receiver i to the n received signal values $\boldsymbol{y}_i^n = y_{i,1}, \ldots, y_{i,n}$, which, as specified in the Introduction, are formed as a component-wise linear combination, according the gain matrix H, of the transmitted signals and Gaussian noise.

Definition 1: The capacity region $\mathcal{C}(H, \boldsymbol{\sigma}, \boldsymbol{P})$ of the GIC is defined as the set of rate-tuples $R_1, \ldots, R_K$ for which there exists a sequence of block length n message sets and power-constrained coding schemes satisfying $\lim_{n\to\infty} \max_{1 \le i \le K} \Pr(W_i \neq g_i(\boldsymbol{y}_i^n)) = 0$, where the probability of error is taken with respect to the distribution induced by the random messages, the coding scheme, and the channel, as specified above.

The degrees-of-freedom $\mathrm{DoF}(H)$ of a GIC, as defined in (1) above, will be assumed to be based on this formal definition of $\mathcal{C}(H, \boldsymbol{\sigma}, \boldsymbol{P})$.

[4]For simplicity, in this paper, we formally adopt a per codeword average power constraint, as opposed to the more conventional *expected* average power constraint. We note that the degrees-of-freedom of a GIC is the same under both types of power constraints. This can be informally proved as follows. For any block length n, starting from a set of codebooks satisfying an expected average power constraint P we form new codebooks by discarding the half of the codewords of each codebook that have the largest Euclidean norm, resulting in a set of codebooks satisfying a per codeword average power constraint $2P$, with a loss in rate of $1/n$ bits per user, or K/n bits in sum–rate. The resulting decoding error probability of the new codebooks can be bounded along the same lines of Lemma 5. The constant factor increase in the power constraint and asymptotically vanishing (in n) loss in rate do not modify the degrees-of-freedom.

III. Real, Scalar GICs With Exactly $K/2$ Degrees-of-Freedom

In this section, we prove Theorem 1 demonstrating the existence of fully connected, real, scalar K-user GICs with exactly $K/2$ degrees-of-freedom. An outline of the proof is as follows. First, we prove a simple lemma (which will also be useful in the next section) showing that $\mathrm{DoF}(H) = \mathrm{DoF}(D_t H D_r)$ for any diagonal matrices D_t and D_r with positive diagonal components. This, in turn, implies that we can transform any H satisfying the assumptions of Theorem 1 to one with irrational, algebraic numbers along the diagonal and integer values in off-diagonal components, while preserving the degrees-of-freedom. We then focus on coding schemes for the new H in which each transmitter is restricted to signaling over the scalar lattice $\{zP^{1/4+\epsilon} : z \in \mathbb{Z}\}$ intersected with the interval $[-P^{1/2}, P^{1/2}]$. The idea is that the integer-valued cross gains guarantee that the interfering signal values at each receiver will also be confined to this scalar lattice (though may fall outside of the $P^{1/2}$ interval), while the irrational direct gains place the directly transmitted signal values on a scaled lattice that "stands out" from the interfering lattice. Specifically, this scaled lattice has the property that offsetting the interfering lattice (equal to the original lattice) by each point in the scaled lattice results in disjoint sets. A nonempty intersection would imply that the direct gain could be written as the ratio of two integers, which would contradict its irrationality. An even stronger property holds for *algebraic* irrational direct gains: the distance between any pair of points obtained by adding a point from the scaled lattice to a point from the interfering lattice actually grows with P. This is shown to follow from a major result in number theory stating that for any irrational algebraic number α and any $\gamma > 0$, a rational p/q approximation will have an error of at least $\delta/q^{2+\gamma}$ for some δ depending only on α and γ.[5] The next step in the proof is to deal with the noise by coupling this inter-point distance growth with Fano's inequality to show that the mutual information induced between each transmitter–receiver pair by independent, uniform distributions on the original power-constrained lattices, taking interference into account, grows like $(1/4 - \epsilon)\log_2 P$, for arbitrarily small ϵ. This, in turn, implies the existence of a sequence of block codes (with symbols from the original lattice) with sum–rate approaching $(K/4)\log_2 P$, and which are correctly decodeable, with high probability, by treating interference as noise.

As mentioned, we begin with an invariance property of $\mathrm{DoF}(H)$.

Lemma 1 (Invariance Property): For any matrix H and diagonal matrices D_t and D_r with positive diagonal components $\mathrm{DoF}(D_t H D_r) = \mathrm{DoF}(H)$.[6]

[5]For irrational algebraic numbers of degree two (solutions to quadratic equations with integer coefficients), such as $\sqrt{2}$, the approximation bound holds with $\gamma = 0$ and is known as Liouville's theorem (established in 1844). The validity of the bound for general algebraic numbers was a long-standing open problem in number theory and was finally established in 1955 by K. F. Roth, for which he was awarded the Fields Medal.

[6]The matrices D_t and D_r need only have nonzero diagonal components for the result to hold. We assume positivity for simplicity, as this is all we shall require in this paper.

Proof: Let

$$D_t = \mathrm{diag}(d_{t1}, \ldots, d_{tK})$$
$$D_r = \mathrm{diag}(d_{r1}, \ldots, d_{rK}).$$

In the matrix multiplication $D_t H D_r$, d_{ti} scales the channel gains from transmitter i to the different receivers, while d_{rj} scales the channel gains from all transmitters to receiver j. By scaling the input signals and noise variances instead of the channel gains, we can write

$$\mathcal{C}(D_t H D_r, \mathbf{1}, P\mathbf{1}) = \mathcal{C}\left(H, \mathbf{1}^t D_r^{-2}, P\mathbf{1}^t D_t^2\right). \quad (2)$$

Let $\check{d}_t = \min_{1\le i\le K} d_{ti}$, $\hat{d}_t = \max_{1\le i\le K} d_{ti}$, $\check{d}_r = \min_{1\le i\le K} d_{ri}$, $\hat{d}_r = \max_{1\le i\le K} d_{ri}$. Since relaxing the power constraints and reducing the noise variances cannot reduce the capacity region of the GIC we have

$$\begin{aligned}\mathcal{C}\left(H, \left(1/\check{d}_r^2\right)\mathbf{1}, P\check{d}_t^2\mathbf{1}\right) &\subseteq \mathcal{C}\left(H, \mathbf{1}^t D_r^{-2}, P\mathbf{1}^t D_t^2\right)\\ &\subseteq \mathcal{C}\left(H, \left(1/\hat{d}_r^2\right)\mathbf{1}, P\hat{d}_t^2\mathbf{1}\right). \quad (3)\end{aligned}$$

Furthermore, once all the noise variances are equal, they can be normalized to 1 by scaling the power constraints, leading to

$$\begin{aligned}\mathcal{C}\left(H, \left(1/\check{d}_r^2\right)\mathbf{1}, P\check{d}_t^2\mathbf{1}\right) &= \mathcal{C}\left(H, \mathbf{1}, P\check{d}_r^2\check{d}_t^2\mathbf{1}\right)\\ \mathcal{C}\left(H, \left(1/\hat{d}_r^2\right)\mathbf{1}, P\hat{d}_t^2\mathbf{1}\right) &= \mathcal{C}\left(H, \mathbf{1}, P\hat{d}_r^2\hat{d}_t^2\mathbf{1}\right). \quad (4)\end{aligned}$$

Using (2)–(4), we can write

$$\begin{aligned}\frac{\max_{\boldsymbol{R}\in\mathcal{C}\left(H,\mathbf{1},P\check{d}_r^2\check{d}_t^2\mathbf{1}\right)}\mathbf{1}^t\boldsymbol{R}}{\frac{1}{2}\log_2 P} &\le \frac{\max_{\boldsymbol{R}\in\mathcal{C}(D_tHD_R,\mathbf{1},P\mathbf{1})}\mathbf{1}^t\boldsymbol{R}}{\frac{1}{2}\log_2 P}\\ &\le \frac{\max_{\boldsymbol{R}\in\mathcal{C}\left(H,\mathbf{1},P\hat{d}_r^2\hat{d}_t^2\mathbf{1}\right)}\mathbf{1}^t\boldsymbol{R}}{\frac{1}{2}\log_2 P}\end{aligned}$$

which can be rewritten as

$$\begin{aligned}&\frac{\frac{1}{2}\log_2\left(P\check{d}_r^2\check{d}_t^2\right)}{\frac{1}{2}\log_2 P}\frac{\max_{\boldsymbol{R}\in\mathcal{C}\left(H,\mathbf{1},P\check{d}_r^2\check{d}_t^2\mathbf{1}\right)}\mathbf{1}^t\boldsymbol{R}}{\frac{1}{2}\log_2\left(P\check{d}_r^2\check{d}_t^2\right)}\\ &\le \frac{\max_{\boldsymbol{R}\in\mathcal{C}(D_tHD_R,\mathbf{1},P)}\mathbf{1}^t\boldsymbol{R}}{\frac{1}{2}\log_2 P}\\ &\le \frac{\frac{1}{2}\log_2\left(P\hat{d}_r^2\hat{d}_t^2\right)}{\frac{1}{2}\log_2 P}\frac{\max_{\boldsymbol{R}\in\mathcal{C}\left(H,\mathbf{1},P\hat{d}_r^2\hat{d}_t^2\mathbf{1}\right)}\mathbf{1}^t\boldsymbol{R}}{\frac{1}{2}\log_2\left(P\hat{d}_r^2\hat{d}_t^2\right)}.\end{aligned}$$

Taking $\limsup$ of all three terms as $P\to\infty$ implies $\mathrm{DoF}(H) \le \mathrm{DoF}(D_tHD_r) \le \mathrm{DoF}(H)$. □

Proof: (of Theorem 1): By Lemma 1, we can scale H (by post-multiplying by an integer valued D_r) so that all off-diagonal elements are integers and all diagonal elements remain irrational algebraic. In addition, from (1), we only need to consider channels where all the inputs have the same power constraint P and all the noise processes have variance 1.

For any $\epsilon > 0$, we will present a communication scheme that achieves $\mathbf{1}^t\boldsymbol{R} = (K/4 - K\epsilon)\log_2 P - o(\log_2 P)$, implying that $\mathrm{DoF}(H) \ge K/2$. Consider the scalar lattice

$$\Lambda_{P,\epsilon} = \{x : x = P^{1/4+\epsilon}z, z \in \mathbb{Z}\}$$

and let $\Theta_{P,\epsilon} = \Lambda_{P,\epsilon} \cap [-\sqrt{P}, \sqrt{P}]$. Note that

$$|\Theta_{P,\epsilon}| = 2\left\lfloor\frac{\sqrt{P}}{P^{1/4+\epsilon}}\right\rfloor + 1 \le 2P^{1/4-\epsilon} + 1. \quad (5)$$

The users communicate using codebooks of block length n, obtained by uniform i.i.d. sampling from $\Theta_{P,\epsilon}$. Note that due to the truncation of the lattice to the interval $[-\sqrt{P}, \sqrt{P}]$, the symbol power $(x_{i,t}^2)$ never exceeds P at any time index, and hence the average codeword power does not exceed P. Each receiver decodes the signal of its transmitter, treating the interfering signals as i.i.d. noise. With this scheme, as $n\to\infty$ we can achieve

$$R_i = I(X_i; Y_i) = H(X_i) - H(X_i|Y_i), \quad i = 1, \ldots, K$$

where $X_i \sim \mathrm{Uniform}(\Theta_{P,\epsilon})$, $Y_i = \sum_{j=1}^K h_{ji}X_j + Z_i$, and $Z_i \sim \mathcal{N}(0,1)$, $i = 1, \ldots, K$.

First we note that

$$H(X_i) = \log_2|\Theta_{P,\epsilon}| \approx (\frac{1}{4} - \epsilon)\log_2 P + \log_2 2.$$

We will show that

$$\limsup_{P\to\infty} H(X_i|Y_i) \le 1, \quad i = 1, \ldots, K,$$

and as a result, $R_i = (\frac{1}{4} - \epsilon)\log_2 P - o(\log_2 P)$ can be achieved. It would then follow that $\mathrm{DoF}(H) \ge K/2$, and, from the upper bound of [10], that $\mathrm{DoF}(H) = K/2$.

We will use the following lemma to upper-bound $H(X_i|Y_i)$ for $i = 1, \ldots, K$.

Lemma 2: Let $\Sigma_{P,\epsilon} = \{\alpha x + s : x \in \Theta_{P,\epsilon}, s \in \Lambda_{P,\epsilon}\}$, with α being any real, irrational, and algebraic number. For any $y \in \Sigma_{P,\epsilon}$, there exists a unique pair $(x, s) \in \Theta_{P,\epsilon} \times \Lambda_{P,\epsilon}$ such that $y = \alpha x + s$. In addition, if $y_1, y_2 \in \Sigma_{P,\epsilon}$, $y_1 \ne y_2$, then $|y_1 - y_2| > P^{\epsilon}$ for any given $\epsilon > 0$ and large enough P.

Proof: Let $y = \alpha x + s$ with $x \in \Theta_{P,\epsilon}$, $s \in \Lambda_{P,\epsilon}$. To get a contradiction, assume that there exists $(\tilde{x}, \tilde{s}) \in \Theta_{P,\epsilon} \times \Lambda_{P,\epsilon}$ with $(\tilde{x}, \tilde{s}) \ne (x, s)$ such that $\alpha\tilde{x} + \tilde{s} = y$. Without loss of generality, we can assume $\tilde{x} \ge x$. Since $\alpha \ne 0$ we have $\tilde{s} \ne s$ and $\tilde{x} > x$. In addition, since by assumption $\alpha x + s = \alpha\tilde{x} + \tilde{s}$, we have

$$\alpha = \frac{s - \tilde{s}}{\tilde{x} - x} = \frac{(z_s - z_{\tilde{s}})P^{1/4+\epsilon}}{(z_{\tilde{x}} - z_x)P^{1/4+\epsilon}} = \frac{z_s - z_{\tilde{s}}}{z_{\tilde{x}} - z_x} \in \mathbb{Q}$$

where $z_x, z_{\tilde{x}}, z_s, z_{\tilde{s}} \in \mathbb{Z}$, which contradicts the assumption of irrational α.

To prove the second part of the lemma, let $\hat{y} = \alpha\hat{x} + \hat{s}$, where $\hat{x} \in \Theta_{P,\epsilon}$, $\hat{s} \in \Lambda_{P,\epsilon}$, and $\hat{y} \in \Sigma_{P,\epsilon}$, with $\hat{y} \ne y$. If $\hat{s} = s$ then

$$|\hat{y} - y| = \alpha|\hat{x} - x| = \alpha|z_{\hat{x}} - z_x|P^{1/4+\epsilon} > P^{\epsilon}$$

where $z_x, z_{\hat{x}} \in \mathbb{Z}$, as long as P is sufficiently large. Similarly, if $\hat{x} = x$ and P is large enough we have

$$|\hat{y} - y| = |\hat{s} - s| = |z_{\hat{s}} - z_s|P^{1/4+\epsilon} > P^{\epsilon}$$

where $z_s, z_{\hat{s}} \in \mathbb{Z}$. So it remains to consider the case $\hat{x} \ne x$ and $\hat{s} \ne s$. Without loss of generality we can assume $\hat{x} > x$. To get

a contradiction, we assume that $|\hat{y}-y| \leq P^{\epsilon}$, and write

$$\begin{aligned} |\hat{y}-y| &\leq P^{\epsilon} \\ |\alpha\hat{x}+\hat{s}-\alpha x-s| &\leq P^{\epsilon} \\ |\alpha z_{\hat{x}}+z_{\hat{s}}-\alpha z_x-z_s| &\leq \frac{P^{\epsilon}}{P^{1/4+\epsilon}} \\ \left|\alpha-\frac{z_s-z_{\hat{s}}}{z_{\hat{x}}-z_x}\right| &\leq \frac{P^{-1/4}}{z_{\hat{x}}-z_x} \end{aligned} \tag{6}$$

where $z_x, z_{\hat{x}}, z_s, z_{\hat{s}} \in \mathbb{Z}$.

On the other hand, there are bounds on how well an irrational algebraic number can be approximated with a rational number. The most refined of those bounds, due to Roth, 1955, states that for any irrational algebraic α, and any $\gamma > 0$, there exists $\delta > 0$ such that

$$\left|\alpha-\frac{p}{q}\right| > \frac{\delta}{q^{2+\gamma}} \tag{7}$$

for all $p, q \in \mathbb{Z}$, $q > 0$ [22].

Combining (6) and (7) we have

$$\frac{\delta}{(z_{\hat{x}}-z_x)^{2+\gamma}} < \left|\alpha-\frac{z_s-z_{\hat{s}}}{z_{\hat{x}}-z_x}\right| \leq \frac{P^{-1/4}}{z_{\hat{x}}-z_x}$$

so that

$$\begin{aligned} 0 < \delta < P^{-1/4}(z_{\hat{x}}-z_x)^{1+\gamma} &\overset{(a)}{\leq} P^{-1/4}\left(2P^{1/4-\epsilon}+1\right)^{1+\gamma} \\ &= 2^{1+\gamma}P^{\gamma/4-\epsilon(1+\gamma)}+o(1) \end{aligned} \tag{8}$$

where we used (5) in step (a). But the right-hand side of (8) goes to 0 as $P \to \infty$ whenever $\epsilon \geq 1/4$ or $\gamma < \epsilon/(1/4-\epsilon)$. Since we can choose any $\gamma > 0$, we can obtain a contradiction in (8) for any $\epsilon > 0$, for large enough P. □

We will use Lemma 2 to build an estimator that can identify X_i in Y_i with high probability.

Let $S_i \triangleq \sum_{j \neq i} h_{ji}X_j$, and note that since $h_{ji} \in \mathbb{Z}$ for $j \neq i$ we have that $S_i \in \Lambda_{P,\epsilon}$. In addition, let $\Sigma_{P,\epsilon,i} = \{h_{ii}x+y : x \in \Theta_{P,\epsilon}, y \in \Lambda_{P,\epsilon}\}$, and let $v_i : \Theta_{P,\epsilon} \times \Lambda_{P,\epsilon} \to \Sigma_{P,\epsilon,i}$ be defined as $v_i(x,s) = h_{ii}x+s$. Using Lemma 2 and the fact that h_{ii} is real, algebraic, and irrational we have that v_i is invertible, i.e., there exists $v_i^{-1} : \Sigma_{P,\epsilon,i} \to \Theta_{P,\epsilon} \times \Lambda_{P,\epsilon}$ such that $v_i^{-1}(v_i(x,s)) = (x,s)$ for any $(x,s) \in \Theta_{P,\epsilon} \times \Lambda_{P,\epsilon}$.

Let $u : \mathbb{R}^2 \to \mathbb{R}$ be defined as $u(x,s) = x$, and let $\hat{X}_i = u(v_i^{-1}(\arg\min_{x \in \Sigma_{P,\epsilon,i}} |x-Y_i|))$. We have $\hat{X}_i \neq X_i$ whenever Y_i is closer to some other point in $\Sigma_{P,\epsilon,i}$ than it is to $h_{ii}X_i+S_i$. From Lemma 2, this can only occur if $|Z_i| \geq P^{\epsilon}/2$ for large enough P. It follows that

$$\begin{aligned} \Pr(\hat{X}_i \neq X_i) &\leq Pr\left(|Z_i| \geq \frac{P^{\epsilon}}{2}\right) \\ &= 2Q_{\mathcal{N}(0,1)}\left(\frac{P^{\epsilon}}{2}\right) \leq 2\exp\left(-\frac{P^{2\epsilon}}{8}\right) \end{aligned}$$

where $Q_{\mathcal{N}(0,1)}(x)$ is the probability that a Gaussian random variable with zero mean and variance one exceeds x. Using the data processing and Fano's inequalities we obtain

$$\begin{aligned} H(X_i|Y_i) &\leq H(X_i|\hat{X}_i) \\ &\leq 1+\Pr(\hat{X}_i \neq X_i)\log(|\Theta_{P,\epsilon}|) \\ &\leq 1+2\exp\left(-\frac{P^{2\epsilon}}{8}\right) \\ &\quad \times \left[\left(\frac{1}{4}-\epsilon\right)\log_2 P+\log_2 2+o(1)\right] \end{aligned} \tag{9}$$

which goes to 1 as $P \to \infty$. □

IV. Degrees-of-Freedom for Rational H

In this section, we give the proof of Theorem 2, establishing that the degrees-of-freedom of any fully connected, real, scalar GIC is bounded strictly below $K/2$, for $K > 2$. As in the previous section, we begin with a sketch of the proof.

Most of the work in the proof is to establish the theorem for $K = 3$ users. The theorem for $K > 3$ will then follow from an extension of the averaging argument of [10], used therein to obtain the $K/2$ degrees-of-freedom upper bound from a bound of 1 on the degrees-of-freedom for $K = 2$. In this case, the $K = 3$ bound is averaged over all three-tuples of users (transmitters and corresponding receivers), as opposed to pairs of users in [10].

Given a $K = 3$ user GIC with fully connected, rational H, using Lemma 1 (invariance property), and eliminating some cross links, we can upper-bound $\mathrm{DoF}(H)$ by $\mathrm{DoF}(\tilde{H})$ where

$$\tilde{H} = [\tilde{h}_{ij}] = \begin{bmatrix} 1 & 0 & 0 \\ 1 & p & 0 \\ 1 & q & 1 \end{bmatrix}$$

where p and q are integers (see Fig. 2 in Subsection IV-C). The main step in the proof of the overall theorem is establishing that $\mathrm{DoF}(\tilde{H}) < 3/2$, which is formally carried out in Lemma 11 below, and proceeds as follows. First, it is shown (Lemma 4) that a deterministic channel obtained by eliminating all noise sources and restricting the power-constrained codewords to have integer valued symbols results in at most a power-constraint-independent loss in the achievable sum rate. Therefore, the degrees-of-freedom (according to the obvious generalization) of this deterministic interference channel (IC) is no smaller than $\mathrm{DoF}(\tilde{H})$. Next, it is shown using a Fano's inequality based argument that if the degrees-of-freedom of the deterministic IC is at least $3/2$ there would exist finite sets of n-dimensional integer-valued vectors $\mathcal{X}_2$ and $\mathcal{X}_3$ such that the corresponding independent random variables $\boldsymbol{x}_2^n$ and $\boldsymbol{x}_3^n$, uniformly distributed on these sets, induce discrete entropies satisfying $H(\boldsymbol{x}_2^n) \approx n(1/4)\log_2 P$, $H(\boldsymbol{x}_3^n) \approx n(1/4)\log_2 P$, $H(\boldsymbol{x}_2^n+\boldsymbol{x}_3^n) \approx n((1/4)+\epsilon)\log_2 P$, and $H(p \cdot \boldsymbol{x}_2^n + q \cdot \boldsymbol{x}_3^n) \approx n((1/2)-\epsilon)\log_2 P$, for the integers p and q defining the channel and ϵ arbitrarily small. These entropy relations suggest that the cardinality of the support of $p \cdot \boldsymbol{x}_2^n + q \cdot \boldsymbol{x}_3^n$ is much larger than that of $\boldsymbol{x}_2^n + \boldsymbol{x}_3^n$. However, tools from additive combinatorics (through Lemma 7) can be used to show that this is impossible for integer-valued p and q, leading to a contradiction, and thereby implying that the deterministic channel must have degrees-of-freedom strictly smaller than $3/2$. Unfortunately, the link between the entropy and the cardinality of the support of a sum of independent, uniformly distributed random variables is sufficiently weak that a somewhat more involved argument (incorporating Lemma 10 and Theorem 3) is ultimately required to reach the above

conclusions. The overall intuition behind the proof, however, is as outlined.

The rest of the section is organized as follows. Sections IV-A and IV-B, respectively, collect supporting results of an information-theoretic nature and results from additive combinatorics. The proof of the main lemma on the $K = 3$ user IC is presented in Section IV-C. Finally, the extension to $K > 3$ is presented in Section IV-D. Throughout, the capacity region of a K-user GIC will be taken as in Definition 1.

A. Supporting Information-Theoretic Results

Lemma 3: The capacity region of a K-user memoryless IC,[7] where the codebook of user i is subject to an average power constraint P_i, is given by the limiting expression given in (10) at the bottom of the page.

Proof: The lemma can be proved by extending the argument of [25] (see also [26] and [27]) to K-user memoryless ICs with possibly continuous alphabets and average power constraints on the inputs. The details are omitted. □

Lemma 4: Given a gain matrix H and power constraints $\boldsymbol{P} = (P_1, \ldots, P_K)$, let $\mathcal{C}_D(H, \boldsymbol{P})$ denote the capacity region of the deterministic IC defined by

$$\bar{y}_i(t) = \sum_{j=1}^{K} h_{ji} \bar{x}_j(t), \quad i = 1, \ldots, K$$

where the inputs are constrained to be integers (i.e., $\bar{x}_i(t) \in \mathbb{Z}$, $i = 1, \ldots, K$, $t = 1, 2 \ldots$) and satisfy an average power constraint $\frac{1}{n}\sum_{t=1}^{n} \bar{x}_i(t)^2 \leq P_i$ for all i.

Then, $\boldsymbol{R} \in \mathcal{C}(H, \mathbf{1}, \boldsymbol{P}) \Rightarrow (\boldsymbol{R} - \boldsymbol{\Delta}) \in \mathcal{C}_D(H, \boldsymbol{P})$ with $\boldsymbol{\Delta} = (\delta_1, \ldots, \delta_K)$, $\delta_i = \frac{1}{2}\log_2(1 + 2\sum_{j=1}^{K} h_{ji}^2)$, $i = 1, \ldots, K$, where $\mathcal{C}(H, \mathbf{1}, \boldsymbol{P})$ is the capacity region of the corresponding GIC (see Definition 1).

Proof: Let $f : \mathbb{R} \to \mathbb{R}$ be defined as

$$f(x) \triangleq \lfloor x \rfloor \cdot 1(x > 0) + \lceil x \rceil \cdot 1(x < 0)$$

and let $g : \mathbb{R} \to \mathbb{R}$ be defined as $g(x) = x - f(x)$. In addition, let $x_{i1} = f(x_i)$, $x_{i2} = g(x_i)$, $y_{i1} = \sum_{j=1}^{K} h_{ji} x_{j1} + z_{i1}$, and $y_{i2} = \sum_{j=1^K} h_{ji} x_{j2} + z_{i2}$, where $z_{i1}, z_{i2} \sim \mathcal{N}(0, 1/2)$ are independent. Then the outputs of the K-user Gaussian IC can be written as $y_i = y_{i1} + y_{i2}$, for $i = 1, \ldots, K$ (see Fig. 1).

If $\bar{\boldsymbol{R}} = (R_1, \ldots, R_K) \in \mathcal{C}(H, \mathbf{1}, \bar{\boldsymbol{P}})$, then for any $\eta > 0$ there exists a family of codebooks $\{C_{1,n}, \ldots, C_{K,n}\}_n$ satisfying the average power constraints, and decoding functions $\{g_{1,n}, \ldots, g_{K,n}\}_n$ with average decoding error probability

[7]Here we are considering more general ICs than the Gaussian case. Definition 1 still applies, but with the appropriate conditional probability distribution of channel outputs given channel inputs.

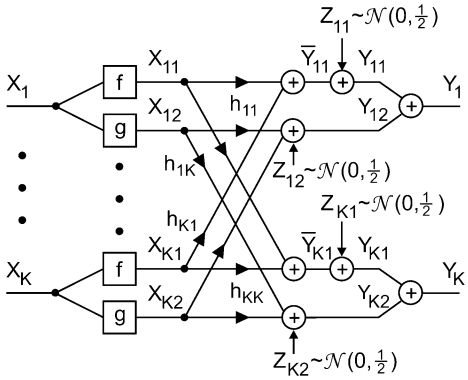

Fig. 1. A decomposition of a K-user Gaussian IC.

going to 0 as $n \to \infty$, such that $\lim_{n\to\infty} \frac{1}{n}\log_2 |C_{i,n}| \geq R_i - \eta$. For block-length n we have

$$\begin{aligned}
&n(R_i - \eta - \epsilon_n) \\
&\quad \leq I(\boldsymbol{x}_i^n; \boldsymbol{y}_i^n) && (11)\\
&\quad \leq I(\boldsymbol{x}_{i1}^n, \boldsymbol{x}_{i2}^n; \boldsymbol{y}_{i1}^n, \boldsymbol{y}_{i2}^n) && (12)\\
&\quad = h(\boldsymbol{y}_{i1}^n, \boldsymbol{y}_{i2}^n) - h(\boldsymbol{y}_{i1}^n, \boldsymbol{y}_{i2}^n | \boldsymbol{x}_{i1}^n, \boldsymbol{x}_{i2}^n) \\
&\quad = h(\boldsymbol{y}_{i1}^n) + h(\boldsymbol{y}_{i2}^n | \boldsymbol{y}_{i1}^n) \\
&\qquad - h\left(\sum_{\substack{j=1\\ j\neq i}}^{K} h_{ji}\boldsymbol{x}_{j1}^n + \boldsymbol{z}_{i1}^n, \sum_{\substack{j=1\\ j\neq i}}^{K} h_{ji}\boldsymbol{x}_{j2}^n + \boldsymbol{z}_{i2}^n\right) \\
&\quad \leq h(\boldsymbol{y}_{i1}^n) + h(\boldsymbol{y}_{i2}^n) - h\left(\sum_{\substack{j=1\\ j\neq i}}^{K} h_{ji}\boldsymbol{x}_{j1}^n + \boldsymbol{z}_{i1}^n\right) \\
&\qquad - h\left(\sum_{\substack{j=1\\ j\neq i}}^{K} h_{ji}\boldsymbol{x}_{j2}^n + \boldsymbol{z}_{i2}^n \,\middle|\, \sum_{\substack{j=1\\ j\neq i}}^{K} h_{ji}\boldsymbol{x}_{j1}^n + \boldsymbol{z}_{i1}^n\right) && (13)\\
&\quad \leq h(\boldsymbol{y}_{i1}^n) + h(\boldsymbol{y}_{i2}^n) - h\left(\sum_{\substack{j=1\\ j\neq i}}^{K} h_{ji}\boldsymbol{x}_{j1}^n + \boldsymbol{z}_{i1}^n\right) \\
&\qquad - h\left(\sum_{\substack{j=1\\ j\neq i}}^{K} h_{ji}\boldsymbol{x}_{j2}^n + \boldsymbol{z}_{i2}^n \,\middle|\, \sum_{\substack{j=1\\ j\neq i}}^{K} h_{ji}\boldsymbol{x}_{j1}^n + \boldsymbol{z}_{i1}^n,\ \boldsymbol{x}_{12}^n, \ldots, \boldsymbol{x}_{K2}^n\right) && (14)\\
&\quad = h(\boldsymbol{y}_{i1}^n) + h\left(\sum_{j=1}^{K} h_{ji}\boldsymbol{x}_{j2}^n + \boldsymbol{z}_{i2}^n\right) \\
&\qquad - h\left(\sum_{\substack{j=1\\ j\neq i}}^{K} h_{ji}\boldsymbol{x}_{j1}^n + \boldsymbol{z}_{i1}^n\right) - h(\boldsymbol{z}_{i2}^n)
\end{aligned}$$

$$\mathcal{C}_{\mathrm{IC}} = \bigcup_{n=1}^{\infty} \bigcup_{\substack{P_{\boldsymbol{x}_1^n \cdots \boldsymbol{x}_K^n} = P_{\boldsymbol{x}_1^n} \cdots P_{\boldsymbol{x}_K^n} \\ \Pr(\|\boldsymbol{x}_i^n\|_2^2 \leq nP_i) = 1, i = 1, \ldots, K}} \left\{ \boldsymbol{R} \in \mathbb{R}_+^K : R_i \leq \frac{1}{n} I(\boldsymbol{x}_i^n; \boldsymbol{y}_i^n), i = 1, \ldots, K \right\} \quad (10)$$

$$\leq I\left(\boldsymbol{x}_{i1}^n;\boldsymbol{y}_{i1}^n\right)+\frac{n}{2}\log_2\left[2\pi e\left(\sum_{j=1}^{K}h_{ji}^2+\frac{1}{2}\right)\right]-\frac{n}{2}\log_2\left(2\pi e\frac{1}{2}\right) \tag{15}$$

$$\leq I\left(\boldsymbol{x}_{i1}^n;\bar{\boldsymbol{y}}_{i1}^n\right)+\frac{n}{2}\log_2\left(1+2\sum_{j=1}^{K}h_{ji}^2\right) \tag{16}$$

where $\epsilon_n \to 0$ as $n \to \infty$. We used Fano's inequality in (11), the data processing inequality in (12), the fact that conditioning reduces entropy in (13) and (14), the Gaussian bound for differential entropies in (15), noting that $|x_{j2}(t)| \leq 1, t = 1, \ldots, n$, and used the data processing inequality in (16), where we defined $\bar{\boldsymbol{y}}_{i1}^n \triangleq \sum_{j=1}^{n} h_{ji}\boldsymbol{x}_{j1}^n$.

Since $\|\boldsymbol{x}_{i1}^n\|_2^2 \leq \|\boldsymbol{x}_i^n\|_2^2 \leq nP_i$ when the channel is used with codebooks satisfying the power constraints, it follows that the codebooks $\{C_{1,n},\ldots,C_{K,n}\}_n$ induce distributions on $\boldsymbol{x}_{i1}^n$, $i = 1,\ldots,K$, that satisfy the power constraints $\{P_i\}_{i=1}^K$. These distributions can be used in (10) to conclude that $(R_1-\eta-\delta_1,\ldots,R_K-\eta-\delta_K,\ldots)$ is an achievable rate vector in $\mathcal{C}_D(H,\boldsymbol{P})$. Since $\eta > 0$ is arbitrary, the result follows. □

Lemma 5: Let $\{(C_{i,n},g_{i,n},P_{e,i,n})\}_{i=1}^K$ denote a block length n coding scheme for a K-user interference channel satisfying average (per codeword) power constraints $\{P_i\}_{i=1}^K$ with rates $\{R_i\}_{i=1}^K$ and average error probabilities $\{P_{e,i,n}\}_{i=1}^K$. If $\{\tilde{C}_{1,n},\ldots,\tilde{C}_{K,n}\}$ is any set of codebooks with $\tilde{C}_{i,n}\subseteq C_{i,n}$, $|\tilde{C}_{i,n}| \geq |C_{i,n}|/\alpha_i$ and $\alpha_i \geq 1$, for all $i = 1,\ldots,K$, then $\{(\tilde{C}_{i,n},g_{i,n},\tilde{P}_{e,i,n})\}_{i=1}^K$ is a coding scheme with rates no smaller than $\{R_i-\frac{1}{n}\log_2\alpha_i\}_{i=1}^K$ and average error probabilities $\{\tilde{P}_{e,i,n}\}_{i=1}^K$ satisfying $\tilde{P}_{e,i,n} \leq (\prod_{j=1}^K\alpha_j)P_{e,i,n}$, and also satisfying the power constraints $\{P_i\}_{i=1}^K$.

Proof: Since $C_{i,n}$ has rate R_i, we have that for $i = 1,\ldots,K$ the rate $\tilde{R}_i$ of $\tilde{C}_{i,n}$ satisfies

$$\tilde{R}_i=\frac{1}{n}\log_2|\tilde{C}_{i,n}|\geq\frac{1}{n}\log_2\left(\frac{|C_{i,n}|}{\alpha_i}\right)=R_i-\frac{1}{n}\log_2\alpha_i. \tag{17}$$

During communication with the codebooks $\{C_{1,n},\ldots,C_{K,n}\}$, the transmitted messages (and hence the codewords) are chosen uniformly and independently, and as a result we have

$$\begin{aligned}P_{e,i,n}&=\frac{1}{\prod_{j=1}^K|C_{j,n}|}\sum_{\boldsymbol{c}_1\in C_{1,n}}\cdots\sum_{\boldsymbol{c}_K\in C_{K,n}}P_{e,i,n}(\boldsymbol{c}_1,\ldots,\boldsymbol{c}_K)\\&\geq\frac{1}{\prod_{j=1}^K|C_{j,n}|}\sum_{\boldsymbol{c}_1\in\tilde{C}_{1,n}}\cdots\sum_{\boldsymbol{c}_K\in\tilde{C}_{K,n}}P_{e,i,n}(\boldsymbol{c}_1,\ldots,\boldsymbol{c}_K)\\&=\frac{1}{\prod_{j=1}^K\alpha_j}\frac{\prod_{j=1}^K\alpha_j}{\prod_{j=1}^K|C_{j,n}|}\times\sum_{\boldsymbol{c}_1\in\tilde{C}_{1,n}}\cdots\sum_{\boldsymbol{c}_K\in\tilde{C}_{K,n}}P_{e,i,n}(\boldsymbol{c}_1,\ldots,\boldsymbol{c}_K)\\&\geq\frac{1}{\prod_{j=1}^K\alpha_j}\frac{1}{\prod_{j=1}^K|\tilde{C}_{j,n}|}\times\sum_{\boldsymbol{c}_1\in\tilde{C}_{1,n}}\cdots\sum_{\boldsymbol{c}_K\in\tilde{C}_{K,n}}P_{e,i,n}(\boldsymbol{c}_1,\ldots,\boldsymbol{c}_K)\\&=\frac{1}{\prod_{j=1}^K\alpha_j}\tilde{P}_{e,i,n}\end{aligned} \tag{18}$$

where we denoted by $P_{e,i,n}(\boldsymbol{c}_1,\ldots,\boldsymbol{c}_K)$ the probability of decoding error when the codewords $\boldsymbol{c}_1,\ldots,\boldsymbol{c}_K$ are transmitted.

Finally, since every codeword of $C_{i,n}$ satisfies the power constraint P_i, the codewords of $\tilde{C}_{i,n}$ satisfy the power constraint P_i. □

Lemma 6: Any achievable rate vector in a K-user IC can be achieved by codebooks with no repeated codewords, i.e., for every $n = 1,2,\ldots$ and $k = 1,\ldots,K$, the codebook $C_{k,n}$ is such that $\boldsymbol{c}_i,\boldsymbol{c}_j \in C_{k,n} \Rightarrow \boldsymbol{c}_i \neq \boldsymbol{c}_j$.

Proof: Since $(R_1,\ldots,R_K)$ is achievable, for any $\eta > 0$ there exists a family of codebooks $\{C_{1,n},\ldots,C_{K,n}\}_n$ with $\frac{1}{n}\log_2|C_{i,n}| \geq R_i-\eta$ satisfying the average power constraints, and a family of decoding functions achieving average error probabilities $P_{e,i,n}$ going to 0 as $n \to \infty$ for every $i = 1,\ldots,K$.

Consider the single-user channel between transmitter i and receiver i obtained from the interference channel by removing all the interfering signals at receiver i. Since the interference cannot help receiver i decode the message of its own transmitter, it follows that R_i can be achieved in the single-user channel with the family of codebooks $\{C_{i,n}\}_{i,n}$, for some decoding functions $\{g'_{i,n}\}_{i,n}$ with average error probabilities no larger than $P_{e,i,n}$, $i = 1,\ldots,K, n = 1,2,\ldots$. Let $\tilde{C}_{i,n} \subset C_{i,n}$ be obtained by removing the worst (i.e., leading to the largest error probability in the single-user channel) half of the codewords in $C_{i,n}$. It is easy to see that $\tilde{C}_{i,n}$ and $g'_{i,n}$ achieve a maximal error probability in the single-user channel no larger than $2P_{e,i,n}$, and in particular, $\tilde{C}_{i,n}$ has no repeated codewords for n large enough. The result follows by using Lemma 5 with $\{\tilde{C}_{i,n}\}_{i,n}$ as defined here, and $\alpha_i = 2, i = 1,\ldots,K$, noting that as $n \to \infty$, $\frac{1}{n}\log_2\alpha_i \to 0$ and $(\prod_{j=1}^K\alpha_j)P_{e,i,n} \to 0$ for $i = 1,\ldots,K$. □

B. Supporting Results From Additive Combinatorics

Given an Abelian group G and two sets $A, B \subseteq G$ let $A+B$ denote the set of sums obtainable by adding one element from A to one element from B.[8] Formally, $A+B=\{a+b : a\in A, b\in B\}$. The set of differences $A-B$ can be defined analogously. For any integer p and set $A \subseteq G$, we denote by $p\cdot A$ the set consisting of all p-multiples of elements of A or $p\cdot A=\{pa : a\in A\}$. For a nonnegative integer p and $A \subseteq G$, we denote by $p\star A$ the set of p-fold sums of A or $p\star A=\{a_1+a_2+\cdots+a_p : a_i\in A \text{ for } i=1,\ldots,p\}$. For negative integers p, $p\star A$ will denote $|p|\star(-A)$. We shall also require the concept of a partial sum set. Given $A, B \subseteq G$, let $F \subseteq A\times B$. The partial sum set of A and B with respect to F, denoted as $A\overset{F}{+}B$, is defined as $A\overset{F}{+}B=\{a+b : (a,b)\in F\}$. If $F = A\times B$, then $A\overset{F}{+}B = A+B$.

We shall later need the following result on sum sets.

Lemma 7: Let $p,q\in\mathbb{Z}$ and $K\in\mathbb{R}$ with $K\geq 1$. If $|A+B|\leq K|A|^{1/2}|B|^{1/2}$, then $|p\cdot A+q\cdot B|\leq K^{d(p,q)}|A|^{1/2}|B|^{1/2}$ for $d(p,q)=2\max\{|p|,|q|\}+5$.

[8]We shall apply these results to the group of vectors of integers with component-wise addition.

Our proof, stated below, follows fairly standard arguments from additive combinatorics (see [23, Chs. 2 and 6]) and is based on the following key results concerning nonempty subsets $A, B \subseteq G$. Proofs of these results can be found in [23].

Lemma 8 (Rusza's Covering Lemma): There exists a subset $X \subseteq B$ such that $|X| \leq |A+B|/|A|$ and $B \subseteq A - A + X$.

Lemma 9 (Plünnecke–Rusza Inequality): For positive integers p, q and any real-valued $\tilde{K} \geq 1$, if $|A+B| \leq \tilde{K}|A|$ then $|p \star B - q \star B| \leq \tilde{K}^{p+q}|A|$.

Proof (of Lemma 7.): By Lemma 8, there exists a set $X \subseteq q \cdot B$ with $|X| \leq |A+q \cdot B|/|A|$ satisfying $q \cdot B \subseteq A - A + X$. This, in turn, implies that $p \cdot A + q \cdot B \subseteq p \cdot A + A - A + X$. Therefore

$$\begin{aligned} |p \cdot A + q \cdot B| &\leq |p \cdot A + A - A||X| \\ &\leq |p \cdot A + A - A|\frac{|A + q \cdot B|}{|A|} \end{aligned} \tag{19}$$

where we have used the trivial inequality $|S+T| \leq |S||T|$ in the first step. Again, by Lemma 8, there exists a set $Y \subseteq A$ with $|Y| \leq |A+B|/|B|$ satisfying $A \subseteq B - B + Y$, which implies that $A + q \cdot B \subseteq q \cdot B + B - B + Y$. Proceeding as above, we then have

$$\begin{aligned} |A + q \cdot B| &\leq |q \cdot B + B - B||Y| \\ &\leq |q \cdot B + B - B|\frac{|A+B|}{|B|}. \end{aligned} \tag{20}$$

Combining (19) and (20) gives

$$\begin{aligned} &|p \cdot A + q \cdot B| \\ &\quad \leq |p \cdot A + A - A||q \cdot B + B - B|\frac{|A+B|}{|A||B|} \\ &\quad \leq |p \cdot A + A - A||q \cdot B + B - B|K|A|^{-1/2}|B|^{-1/2} \quad (21) \\ &\quad \leq |p \star A + A - A||q \star B + B - B|K|A|^{-1/2}|B|^{-1/2} \quad (22) \end{aligned}$$

where (21) follows from the assumption of the lemma and (22) follows from the trivial inclusions $p \cdot A + A - A \subseteq p \star A + A - A$ and $q \cdot B + B - B \subseteq q \star B + B - B$. Rewriting the assumption of the lemma as $|A+B| \leq K|A|^{-1/2}|B|^{1/2}|A|$, we can apply Lemma 9 with $\tilde{K} = K|A|^{-1/2}|B|^{1/2}$ to conclude that

$$|q \star B + B - B| \leq K^{|q|+2}|A|^{-(|q|+2)/2}|B|^{(|q|+2)/2}|A|. \tag{23}$$

Similarly, rewriting the assumption of the lemma as $|A+B| \leq K|A|^{1/2}|B|^{-1/2}|B|$, we can apply Lemma 9, with the roles of A and B switched, to obtain

$$|p \star A + A - A| \leq K^{|p|+2}|A|^{(|p|+2)/2}|B|^{-(|p|+2)/2}|B|. \tag{24}$$

Combining (22) with (23) and (24) gives

$$\begin{aligned} &|p \cdot A + q \cdot B| \\ &\quad \leq K^{|p|+2}|A|^{(|p|+2)/2}|B|^{-(|p|+2)/2}K^{|q|+2} \\ &\qquad \times |A|^{-(|q|+2)/2}|B|^{(|q|+2)/2}K|A|^{1/2}|B|^{1/2} \\ &\quad \leq K^{2\max\{|p|,|q|\}+5}|A|^{1/2}|B|^{1/2} \end{aligned} \tag{25}$$

where (25) follows from $K|A|^{-1/2}|B|^{1/2} \geq 1$ and $K|A|^{1/2}|B|^{-1/2} \geq 1$, which in turn follow from the lemma's assumption

$$|A+B| \leq K|A|^{-1/2}|B|^{1/2}|A| = K|A|^{1/2}|B|^{-1/2}|B|$$

together with the obvious relations $|A+B| \geq |A|$ and $|A+B| \geq |B|$. The lemma is thus established with $d(p,q) = 2\max\{|p|,|q|\} + 5$. □

Remark 1: A considerably smaller $d(p,q)$ for larger p and q can be obtained by applying the bounds of [24] to the factors $|p \cdot A + A - A|$ and $|q \cdot B + B - B|$ appearing in (21). These bounds are obtained through a more sophisticated application of Lemma 9, that takes greater advantage of the structure of sets like $p_1 \cdot A + \cdots + p_m \cdot A$. The resulting $d(p,q)$ grows logarithmically in $|p|$ and $|q|$. It is likely that a direct application of the technique of [24] to $p \cdot A + q \cdot B$ would even further improve $d(p,q)$ for large p and q.

We shall also make use of the following lemma relating the entropy of a sum of uniformly distributed independent random variables to a partial sum set involving their supports.

Lemma 10: Let X and Y be independent uniform random variables with support sets $A \subseteq G$ and $B \subseteq G$ for some Abelian group G, with $|A| \geq |B|$, such that

$$H(X+Y) \leq (1+\epsilon)\log_2|A|$$

for some $\epsilon > 0$. Then, for any given $c > 1$ there exists a set $F \subseteq A \times B$ such that

$$\begin{aligned} |F| &\geq |A||B|\frac{c-1}{c} \text{ and} \\ |A \overset{F}{+} B| &\leq \left(|A|^{1/2+c\epsilon}|B|^{-1/2}\right)|A|^{1/2}|B|^{1/2}. \end{aligned}$$

Proof: Define $T(s) = \{(a,b) \in A \times B : a+b = s\}$. Define $S \triangleq \{s : |T(s)| \geq |B||A|^{-c\epsilon}\}$ and $F \triangleq \{(a,b) \in A \times B : a+b \in S\}$. From these definitions we have $|A \overset{F}{+} B| = |S|$. In addition

$$\begin{aligned} |A||B| &\geq \sum_{s \in S}|T(s)| \\ &\geq |S||B||A|^{-c\epsilon} \end{aligned}$$

where the last step follows from the definition of S. As a result, $|S| \leq |A|^{1+c\epsilon}$, giving the required upper bound for $|A \overset{F}{+} B| = |S|$.

To get a lower bound on $|F|$ we start by rewriting it as follows:

$$\begin{aligned} |F| &= \sum_{(x,y) \in A \times B} 1((x,y) \in F) \\ &= |A||B|\Pr((X,Y) \in F) \\ &= |A||B|\Pr(X+Y \in S) \\ &= |A||B|[1 - \Pr(X+Y \in S^c)] \\ &= |A||B|[1 - \Pr(|T(X+Y)| < |B||A|^{-c\epsilon})]. \end{aligned} \tag{26}$$

The probability term can be upper-bounded using Markov's inequality, noting that $|B| \geq |T(s)|$ for all s

$$\begin{aligned}&\Pr\left(|T(X+Y)| < |B||A|^{-c\epsilon}\right)\\&\quad= \Pr\left(\log_2 \frac{|B|}{|T(X+Y)|} > c\epsilon \log_2 |A|\right)\\&\quad\leq \frac{E\left[\log_2 \frac{|B|}{|T(X+Y)|}\right]}{c\epsilon \log_2 |A|}.\end{aligned} \tag{27}$$

To bound the expectation in (27) we note that for any $(x,y) \in A \times B$, $\Pr(X=x, Y=y) = \frac{1}{|A||B|}$, and expand $I(X+Y;Y)$ in two different ways

$$\begin{aligned}I(X+Y;Y) &= H(X+Y) - H(X+Y|Y)\\&= H(X+Y) - H(X)\\&= H(X+Y) - \log_2 |A|\\&\leq \epsilon \log_2 |A|\\I(X+Y;Y)&\\&= H(Y) - H(Y|X+Y)\\&= \log_2 |B| - E[-\log_2 p_{Y|X+Y}(Y|X+Y)]\\&= \log_2 |B| - E\left[-\log_2 \frac{p_{Y,X+Y}(Y,X+Y)}{p_{X+Y}(X+Y)}\right]\\&= \log_2 |B| - E\left[-\log_2 \frac{1}{|A||B|}\frac{|A||B|}{|T(X+Y)|}\right]\\&= E\left[\log_2 \frac{|B|}{|T(X+Y)|}\right]\end{aligned}$$

to obtain

$$E\left[\log_2 \frac{|B|}{|T(X+Y)|}\right] \leq \epsilon \log_2 |A|. \tag{28}$$

From (26), (27) and (28) we obtain:

$$|F| \geq |A||B|\left(1 - \frac{1}{c}\right). \qquad \square$$

Finally, we shall also rely on the following important theorem from additive combinatorics, as stated in [23], relating partial sum sets to full sum sets.

Theorem 3 (Balog–Szemerédi–Gowers Theorem): Let $A \subseteq G$ and $B \subseteq G$ for some Abelian group G and let $F \subseteq A \times B$ be such that

$$|F| \geq |A||B|/K \quad \text{and} \quad |A \overset{F}{+} B| \leq K'|A|^{1/2}|B|^{1/2}$$

for some $K \geq 1$ and $K' > 0$. Then there exists $A' \subseteq A, B' \subseteq B$ such that

$$\begin{aligned}|A'| &\geq \frac{|A|}{4\sqrt{2}K}\\|B'| &\geq \frac{|B|}{4K}\\|A' + B'| &\leq 2^{12}K^5(K')^3|A|^{1/2}|B|^{1/2}.\end{aligned}$$

Theorem 3 is proved in [23, Ch. 6].

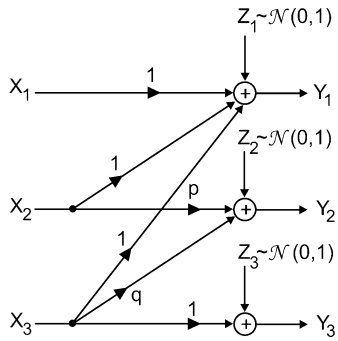

Fig. 2. A three-user Gaussian IC with channel matrix $\bar{H}$.

C. Main Lemma

In this subsection, we state and prove the main lemma at the core of our proof of Theorem 2.

Lemma 11: Let $p, q \in \mathbb{Z}$, $p, q \neq 0$, and

$$\tilde{H} = [\tilde{h}_{ij}] = \begin{bmatrix}1 & 0 & 0\\1 & p & 0\\1 & q & 1\end{bmatrix}$$

with the corresponding GIC depicted in Fig. 2. Then $\mathrm{DoF}(\tilde{H}) \leq \frac{3}{2} - \epsilon(p,q)$, with $\epsilon(p,q) > 0$. In particular, this holds for

$$\epsilon(p,q) = \frac{1}{12d(p,q)+2},$$

where $d(p,q)$ is as in Lemma 7.

Proof: We start by extending the definition of degrees-of-freedom to deterministic interference channels. Consider a K-user deterministic interference channel with input and output alphabets $\{\mathcal{X}_i\}_{i=1}^K$,$\{\mathcal{Y}_i\}_{i=1}^K$, defined by

$$y_i(t) = v_i(x_1(t), \ldots, x_K(t)), \quad i = 1, \ldots, K; t = 1, 2, \ldots$$

where for each i, $x_i(t) \in \mathcal{X}_i$ must satisfy an average power constraint $\sum_{t=1}^n x_i^2(t) \leq nP$, and $v_i : \mathcal{X}_1 \times \cdots \times \mathcal{X}_K \to \mathcal{Y}_i$ is a deterministic function. Let $\mathcal{C}(P)$ be the capacity region of the channel with power constraint P. We define the degrees-of-freedom of the deterministic channel by

$$\mathrm{DoF} \triangleq \limsup_{P\to\infty} \frac{\max_{\boldsymbol{R}\in\mathcal{C}(P)} \mathbf{1}^t\boldsymbol{R}}{(1/2)\log_2 P}.$$

Due to Lemma 4, the degrees-of-freedom of the GIC with channel matrix $\tilde{H}$ is upper-bounded by the degrees-of-freedom of the deterministic channel

$$y_i(t) = \sum_{j=1}^3 \tilde{h}_{ji} x_j(t), \quad i = 1, 2, 3, \tag{29}$$

where $x_i(t), y_i(t) \in \mathbb{Z}$, and where the channel inputs are subject to the average power constraint $\frac{1}{n}\sum_{t=1}^n x_i^2(t) \leq P$, for $i = 1, 2, 3$. We will prove by contradiction that the degrees-of-freedom of this deterministic channel is strictly smaller than $3/2$. Therefore, to get a contradiction, we assume that as P goes to infinity, there are achievable rates $(R_1(P), R_2(P), R_3(P))$ satisfying

$$\limsup_{P\to\infty} \frac{R_1(P) + R_2(P) + R_3(P)}{(1/2)\log_2 P} = \frac{3}{2}.$$

This implies that there exists an increasing sequence of power constraints $\{P_m\}_{m=1}^{\infty}$ with $\lim_{m\to\infty} P_m = \infty$, such that

$$\lim_{m\to\infty} \frac{R_1(P_m)+R_2(P_m)+R_3(P_m)}{(1/2)\log_2 P_m} = \frac{3}{2}. \tag{30}$$

For a power constraint P_m, consider a set of codebooks $\{\mathcal{X}_{1,n,m}, \mathcal{X}_{2,n,m}, \mathcal{X}_{3,n,m}\}$ of block-length n with rates $\{R_{1,m}, R_{2,m}, R_{3,m}\}$ in the deterministic channel (29), and with decoding functions $\{g_{1,n,m}, g_{2,n,m}, g_{3,n,m}\}$ that achieve average error probabilities $\{P_{e,1,n,m}, P_{e,2,n,m}, P_{e,3,n,m}\}$. We assume that $\{R_{1,m}, R_{2,m}, R_{3,m}\}$ satisfy (30) as $m \to \infty$ and that $\{P_{e,1,n,m}, P_{e,2,n,m}, P_{e,3,n,m}\}$ go to 0 as $n \to \infty$ for fixed m. In addition, let $\boldsymbol{x}_i^n$, $i = 1,2,3$, be the independent random vectors induced by the codebooks resulting from the uniform distribution of the messages. It follows that $\mathcal{X}_{i,n,m} \in \mathbb{Z}^n$ is the support set of $\boldsymbol{x}_i^n$, $i = 1,2,3$. Due to Lemma 6 we can assume without loss of generality that the codebooks do not have repeated codewords. This implies that each $\boldsymbol{x}_i^n$ is chosen uniformly in $\mathcal{X}_{i,n,m}$, or equivalently, that $\Pr(\boldsymbol{x}_i^n) = \frac{1}{|\mathcal{X}_{i,n,m}|}$, for $\boldsymbol{x}_i^n \in \mathcal{X}_{i,n,m}$, $i = 1,2,3$.

Using Fano's inequality, we write (where, for simplicity, we suppress the dependence on m of the variables $\boldsymbol{x}_1, \boldsymbol{x}_2$, etc.)

$$\begin{aligned} n(R_{1,m}+R_{2,m}+R_{3,m}-\delta_n) &\le \sum_{i=1}^{3} I(\boldsymbol{x}_i^n;\boldsymbol{y}_i^n) \\ &= \sum_{i=1}^{3}[H(\boldsymbol{y}_i^n) - H(\boldsymbol{y}_i^n|\boldsymbol{x}_i^n)] \\ &= H(\boldsymbol{x}_1^n+\boldsymbol{x}_2^n+\boldsymbol{x}_3^n) - H(\boldsymbol{x}_2^n+\boldsymbol{x}_3^n) \\ &\quad + H(p\boldsymbol{x}_2^n+q\boldsymbol{x}_3^n) - H(q\boldsymbol{x}_3^n) + H(\boldsymbol{x}_3^n) \\ &= H(\boldsymbol{x}_1^n+\boldsymbol{x}_2^n+\boldsymbol{x}_3^n) - H(\boldsymbol{x}_2^n+\boldsymbol{x}_3^n) + H(p\boldsymbol{x}_2^n+q\boldsymbol{x}_3^n). \end{aligned} \tag{31}$$

To handle the first term of (31), we use the following lemma, which follows from [21, Exercise 8.7] and Jensen's inequality.

Lemma 12: Let $\boldsymbol{X}^n = (X_1, \ldots, X_n)$ be a discrete random vector on $\mathbb{Z}^n$. Then

$$H(\boldsymbol{X}^n) \le \frac{n}{2}\log_2\left[2\pi e\left(\frac{1}{n}\sum_{i=1}^{n} Var(X_i) + \frac{1}{12}\right)\right]$$

where $\mathrm{Var}(X_i) = E(X_i^2) - E^2(X_i)$ is the variance of X_i.

Therefore, using the independence among the input signals and the fact that $\sum_{t=1}^{n} \mathrm{Var}(x_{i,t}) \le E\|\boldsymbol{x}_i^n\|_2^2 \le nP_m$ we obtain

$$H(\boldsymbol{x}_1^n+\boldsymbol{x}_2^n+\boldsymbol{x}_3^n) \le \frac{n}{2}\log_2\left[2\pi e\left(3P_m+\frac{1}{12}\right)\right]. \tag{32}$$

The remaining two terms in (31) will be bounded in two different ways. First, we use the following simple bounds on $H(aX+bY)$, valid for any $a, b \neq 0$ and independent random variables X and Y:

$$\begin{aligned} H(aX+bY) &\le H(aX+bY, aX) \\ &= H(aX) + H(aX+bY|aX) \\ &= H(X) + H(Y) \\ H(aX+bY) &\ge H(aX+bY|bY) \\ &= H(aX) = H(X). \end{aligned} \tag{33}$$

Using these bounds we get

$$\begin{aligned} &H(p\boldsymbol{x}_2^n+q\boldsymbol{x}_3^n) - H(\boldsymbol{x}_2^n+\boldsymbol{x}_3^n) \\ &\quad \le H(\boldsymbol{x}_2^n) + H(\boldsymbol{x}_3^n) - \max\{H(\boldsymbol{x}_2^n); H(\boldsymbol{x}_3^n)\} \\ &\quad = \min\{H(\boldsymbol{x}_2^n); H(\boldsymbol{x}_3^n)\} \\ &\quad = \log_2 \min\{|\mathcal{X}_{2,n,m}|; |\mathcal{X}_{3,n,m}|\}. \end{aligned} \tag{34}$$

We define $f_{\min}(n,m)$ such that

$$\log_2 \min\{|\mathcal{X}_{2,n,m}|; |\mathcal{X}_{3,n,m}|\} = \left(\frac{1}{4} + f_{\min}(n,m)\right) n \log_2 P_m.$$

Note that by the assumed dependence of $\mathcal{X}_{i,n,m}$ on R_i, $\lim_{n\to\infty} f_{\min}(n,m)$ exists for each m. Then, using (30), (31), (32), and (34), we have

$$\begin{aligned} \frac{3}{2} &= \lim_{m\to\infty} \frac{\lim_{n\to\infty}(R_{1,m}+R_{2,m}+R_{3,m}-\delta_n)}{(1/2)\log_2 P_m} \\ &\le 1 + \frac{1}{2} + \liminf_{m\to\infty} \frac{\lim_{n\to\infty} f_{\min}(n,m)\log_2 P_m}{(1/2)\log_2 P_m} \end{aligned}$$

which implies

$$\liminf_{m\to\infty} \lim_{n\to\infty} f_{\min}(n,m) \ge 0. \tag{35}$$

As a second option, we use Lemma 12 to get the upper bound

$$H(p\boldsymbol{x}_2^n+q\boldsymbol{x}_3^n) \le \frac{n}{2}\log_2\left[2\pi e\left((p^2+q^2)P_m+\frac{1}{12}\right)\right] \tag{36}$$

which we use to get

$$\begin{aligned} &H(p\boldsymbol{x}_2^n+q\boldsymbol{x}_3^n) - H(\boldsymbol{x}_2^n+\boldsymbol{x}_3^n) \\ &\quad \le \frac{n}{2}\log_2\left[2\pi e\left((p^2+q^2)P_m+\frac{1}{12}\right)\right] \\ &\qquad - \log_2 \max\{|\mathcal{X}_{2,n,m}|; |\mathcal{X}_{3,n,m}|\}. \end{aligned} \tag{37}$$

We define $f_{\max}(n,m)$ such that

$$\log_2 \max\{|\mathcal{X}_{2,n,m}|; |\mathcal{X}_{3,n,m}|\} = \left(\frac{1}{4} + f_{\max}(n,m)\right) n \log_2 P_m.$$

Note, as above, that $\lim_{n\to\infty} f_{\max}(n,m)$ exists for each m. Then, using (30), (31), (32), and (37), we have

$$\begin{aligned} \frac{3}{2} &= \lim_{m\to\infty} \frac{\lim_{n\to\infty}(R_{1,m}+R_{2,m}+R_{3,m}-\delta_n)}{(1/2)\log_2 P_m} \\ &\le 1 + 1 - \frac{1}{2} - \limsup_{m\to\infty} \frac{\lim_{n\to\infty} f_{\max}(n,m)\log_2 P_m}{(1/2)\log_2 P_m} \end{aligned}$$

which implies

$$\limsup_{m\to\infty} \lim_{n\to\infty} f_{\max}(n,m) \le 0. \tag{38}$$

Since $f_{\max}(n,m) \ge f_{\min}(n,m)$, (35) and (38) imply

$$\lim_{m\to\infty}\lim_{n\to\infty} f_{\max}(n,m) = \lim_{m\to\infty}\lim_{n\to\infty} f_{\min}(n,m) = 0 \tag{39}$$

and, as a result, for any $\xi > 0$, there exist $m_0(\xi)$ and $n_0(\xi,m)$ such that $|f_{\min}(n,m)| \le \xi$ and $|f_{\max}(n,m)| \le \xi$ for all $m \ge m_0(\xi)$ and $n \ge n_0(\xi,m)$. Therefore, for any $\xi > 0$, $m \ge m_0(\xi)$, and $n \ge n_0(\xi,m)$ we have

$$\left(\frac{1}{4} - \xi\right) n \log_2 P_m \le \log_2(\min\{|\mathcal{X}_{2,n,m}|; |\mathcal{X}_{3,n,m}|\})$$

$$\le \log_2\left(\max\{|\mathcal{X}_{2,n,m}|;|\mathcal{X}_{3,n,m}|\}\right)$$
$$\le \left(\frac{1}{4}+\xi\right) n \log_2 P_m. \tag{40}$$

We define $g(n,m) \ge 0$ such that $H(\boldsymbol{x}_2^n + \boldsymbol{x}_3^n) = (1+g(n,m))\log_2(\max\{|\mathcal{X}_{2,n,m}|;|\mathcal{X}_{3,n,m}|\})$. Then, using (30), (31), (36), and (39) we have (41) given at the bottom of the page, which, together with (39) and the condition $g(n,m)\ge 0$, imply

$$\lim_{m\to\infty} \limsup_{n\to\infty} g(n,m) = 0 \tag{42}$$

and, as a result, for any $\epsilon > 0$, there exist $m_0(\epsilon)$ and $n_0(\epsilon,m)$ such that, $0 \le g(n,m) \le \epsilon$ for all $m \ge m_0(\epsilon)$ and $n \ge n_0(\epsilon,m)$.

Therefore, for any $\epsilon > 0$, $m \ge m_0(\epsilon)$, and $n \ge n_0(\epsilon,m)$ we have

$$H(\boldsymbol{x}_2^n + \boldsymbol{x}_3^n) \le (1+\epsilon)\log_2 \max\{|\mathcal{X}_{2,n,m}|,|\mathcal{X}_{3,n,m}|\}.$$

For any given $c > 1$, Lemma 10 guarantees the existence of $F_{c,n,m} \subseteq \mathcal{X}_{2,n,m} \times \mathcal{X}_{3,n,m}$ such that

$$|F_{c,n,m}| \ge \frac{|\mathcal{X}_{2,n,m}||\mathcal{X}_{3,n,m}|}{K} \text{ and}$$
$$\left|\mathcal{X}_{2,n,m} \overset{F_{c,n,m}}{+} \mathcal{X}_{3,n,m}\right| \le K'_{n,m}|\mathcal{X}_{2,n,m}|^{1/2}|\mathcal{X}_{3,n,m}|^{1/2}$$

with $K = c/(c-1)$ and

$$K'_{n,m} = (\max\{|\mathcal{X}_{2,n,m}|;|\mathcal{X}_{3,n,m}|\})^{1/2+c\epsilon} \cdot (\min\{|\mathcal{X}_{2,n,m}|;|\mathcal{X}_{3,n,m}|\})^{-1/2}.$$

Using Theorem 3 with $F_{c,n,m}$, it follows that there exist $\mathcal{X}'_{2,n,m} \subseteq \mathcal{X}_{2,n,m}$ and $\mathcal{X}'_{3,n,m} \subseteq \mathcal{X}_{3,n,m}$ such that

$$|\mathcal{X}'_{2,n,m}| \ge \frac{|\mathcal{X}_{2,n,m}|}{4\sqrt{2}K} \tag{43}$$
$$|\mathcal{X}'_{3,n,m}| \ge \frac{|\mathcal{X}_{3,n,m}|}{4\sqrt{2}K} \tag{44}$$
$$|\mathcal{X}'_{2,n,m} + \mathcal{X}'_{3,n,m}| \le 2^{12}K^5(K'_{n,m})^3|\mathcal{X}_{2,n,m}|^{1/2}|\mathcal{X}_{3,n,m}|^{1/2}. \tag{45}$$

From (43) and (44) and Lemma 5 it follows that the set of codebooks $\{\mathcal{X}_{1,n,m}, \mathcal{X}'_{2,n,m}, \mathcal{X}'_{3,n,m}\}$ has rates $(R_{1,m}, R_{2,m} - \frac{1}{n}\log_2(4\sqrt{2}K), R_{3,m} - \frac{1}{n}\log_2(4\sqrt{2}K))$, and using the decoding functions $\{g_{1,n,m}, g_{2,n,m}, g_{3,n,m}\}$, the average error probabilities are no larger than $\{P_{e,1,n,m} * 32K^2, P_{e,2,n,m} * 32K^2, P_{e,3,n,m} * 32K^2\}$. Let $\tilde{\boldsymbol{x}}_1^n, \tilde{\boldsymbol{x}}_2^n, \tilde{\boldsymbol{x}}_3^n$ be the random vectors induced by the codebooks $\{\mathcal{X}_{1,n,m}, \mathcal{X}'_{2,n,m}, \mathcal{X}'_{3,n,m}\}$. Using Fano's inequality and absorbing in δ'_n all the constants that vanish with $n \to \infty$ we write

$$\begin{aligned} n(R_{1,m} + R_{2,m} + R_{3,m} - \delta'_n) &\le \sum_{i=1}^{3} I(\tilde{\boldsymbol{x}}_i^n; \boldsymbol{y}_i^n) \\ &= \sum_{i=1}^{3} [H(\boldsymbol{y}_i^n) - H(\boldsymbol{y}_i^n|\tilde{\boldsymbol{x}}_i^n)] \\ &= H(\tilde{\boldsymbol{x}}_1^n + \tilde{\boldsymbol{x}}_2^n + \tilde{\boldsymbol{x}}_3^n) - H(\tilde{\boldsymbol{x}}_2^n + \tilde{\boldsymbol{x}}_3^n) \\ &\quad + H(p\tilde{\boldsymbol{x}}_2^n + q\tilde{\boldsymbol{x}}_3^n) - H(q\tilde{\boldsymbol{x}}_3^n) + H(\tilde{\boldsymbol{x}}_3^n) \\ &= H(\tilde{\boldsymbol{x}}_1^n + \tilde{\boldsymbol{x}}_2^n + \tilde{\boldsymbol{x}}_3^n) - H(\tilde{\boldsymbol{x}}_2^n + \tilde{\boldsymbol{x}}_3^n) \\ &\quad + H(p\tilde{\boldsymbol{x}}_2^n + q\tilde{\boldsymbol{x}}_3^n). \end{aligned} \tag{46}$$

To bound the first term of (46), as before, we use Lemma 12, obtaining

$$H(\tilde{\boldsymbol{x}}_1^n + \tilde{\boldsymbol{x}}_2^n + \tilde{\boldsymbol{x}}_3^n) \le \frac{n}{2}\log_2\left[2\pi e\left(3P_m + \frac{1}{12}\right)\right]. \tag{47}$$

To bound the remaining two terms of (46) we use techniques from additive combinatorics to upper-bound the cardinality of the support set of $(p\tilde{\boldsymbol{x}}_2^n + q\tilde{\boldsymbol{x}}_3^n)$ in terms of the cardinality of the support set of $(\tilde{\boldsymbol{x}}_2^n + \tilde{\boldsymbol{x}}_3^n)$. From Lemma 7, (43), (44), and (45) we have

$$\begin{aligned} &|p\cdot\mathcal{X}'_{2,n,m} + q\cdot\mathcal{X}'_{3,n,m}| \\ &\quad \le |\mathcal{X}'_{2,n,m} + \mathcal{X}'_{3,n,m}|^{d(p,q)}|\mathcal{X}'_{2,n,m}|^{[1-d(p,q)]/2} \\ &\qquad \times |\mathcal{X}'_{3,n,m}|^{[1-d(p,q)]/2} \\ &\quad \le \{2^{12}K^5(K'_{n,m})^3|\mathcal{X}_{2,n,m}|^{1/2}|\mathcal{X}_{3,n,m}|^{1/2}\}^{d(p,q)} \\ &\qquad \cdot \left(\frac{|\mathcal{X}_{2,n,m}|}{4\sqrt{2}K}\right)^{[1-d(p,q)]/2} \left(\frac{|\mathcal{X}_{3,n,m}|}{4\sqrt{2}K}\right)^{[1-d(p,q)]/2} \end{aligned} \tag{48}$$

which together with the bound $H(X) \le \log_2|\mathcal{X}|$ results in

$$\begin{aligned} H(p\tilde{\boldsymbol{x}}_2^n + q\tilde{\boldsymbol{x}}_3^n) &\le \frac{1}{2}\log_2|\mathcal{X}_{2,n,m}| + \frac{1}{2}\log_2|\mathcal{X}_{3,n,m}| \\ &\quad + 3d(p,q)\log_2 K'_{n,m} + \tilde{K}_{c,p,q} \\ &= \frac{1}{2}\log_2|\mathcal{X}_{2,n,m}| + \frac{1}{2}\log_2|\mathcal{X}_{3,n,m}| \\ &\quad + \frac{3}{2}d(p,q)(1+2c\epsilon)\log_2(\max\{|\mathcal{X}_{2,n,m}|;|\mathcal{X}_{3,n,m}|\}) \\ &\quad - \frac{3}{2}d(p,q)\log_2(\min\{|\mathcal{X}_{2,n,m}|;|\mathcal{X}_{3,n,m}|\}) + \tilde{K}_{c,p,q} \end{aligned} \tag{49}$$

where $\tilde{K}_{c,p,q}$ is some constant independent of n and m.

$$\begin{aligned} \frac{3}{2} &= \lim_{m\to\infty} \frac{\lim_{n\to\infty}(R_{1,m} + R_{2,m} + R_{3,m} - \delta_n)}{(1/2)\log_2 P_m} \\ &\le 1 + 1 - \frac{1}{2} - 0 - \limsup_{m\to\infty} \frac{\limsup_{n\to\infty} g(n,m)((1/4) + f_{\max}(n,m))\log_2 P_m}{(1/2)\log_2 P_m} \end{aligned} \tag{41}$$

On the other hand, using (43) and (44) we have

$$\begin{aligned} &H\left(\tilde{\boldsymbol{x}}_2^n+\tilde{\boldsymbol{x}}_3^n\right)\\ &\quad\geq \max\left\{H\left(\tilde{\boldsymbol{x}}_2^n\right);H\left(\tilde{\boldsymbol{x}}_3^n\right)\right\}\\ &\quad= \log_2(\max\{|\mathcal{X}'_{2,n,m}|;|\mathcal{X}'_{3,n,m}|\})\\ &\quad\geq \log_2(\max\{|\mathcal{X}_{2,n,m}|;|\mathcal{X}_{3,n,m}|\})-\log_2(4\sqrt{2}K). \end{aligned} \quad (50)$$

Therefore

$$\begin{aligned} &H\left(p\tilde{\boldsymbol{x}}_2^n+q\tilde{\boldsymbol{x}}_3^n\right)-H\left(\tilde{\boldsymbol{x}}_2^n+\tilde{\boldsymbol{x}}_3^n\right)\\ &\quad\leq \frac{1}{2}\log_2(\min\{|\mathcal{X}_{2,n,m}|;|\mathcal{X}_{3,n,m}|\})\\ &\qquad-\frac{1}{2}\log_2(\max\{|\mathcal{X}_{2,n,m}|;|\mathcal{X}_{3,n,m}|\})\\ &\qquad+\frac{3}{2}d(p,q)(1+2c\epsilon)\log_2(\max\{|\mathcal{X}_{2,n,m}|;|\mathcal{X}_{3,n,m}|\})\\ &\qquad-\frac{3}{2}d(p,q)\log_2(\min\{|\mathcal{X}_{2,n,m}|;|\mathcal{X}_{3,n,m}|\})+\tilde{K}'_{c,p,q}\\ &\quad\leq 0+\frac{3}{2}d(p,q)(1+2c\epsilon)\log_2(\max\{|\mathcal{X}_{2,n,m}|;|\mathcal{X}_{3,n,m}|\})\\ &\qquad-\frac{3}{2}d(p,q)\log_2(\min\{|\mathcal{X}_{2,n,m}|;|\mathcal{X}_{3,n,m}|\})+\tilde{K}'_{c,p,q} \end{aligned} \quad (51)$$

where $\tilde{K}'_{c,p,q}$ is some other constant independent of n and m. Using the bounds in (40) we obtain

$$\begin{aligned} &H\left(p\tilde{\boldsymbol{x}}_2^n+q\tilde{\boldsymbol{x}}_3^n\right)-H\left(\tilde{\boldsymbol{x}}_2^n+\tilde{\boldsymbol{x}}_3^n\right)\\ &\quad\leq\left[3\xi d(p,q)+3c\epsilon d(p,q)\left(\frac{1}{4}+\xi\right)\right]n\log_2 P_m+\tilde{K}'_{c,p,q} \end{aligned} \quad (52)$$

which, together with (46) and (47), imply (53) at the bottom of the page, which is strictly smaller than $(3/2)$ for small enough ξ and ϵ, and large enough m and n. This contradicts (30).

We can refine the above analysis to find the smallest $\epsilon(p,q)$ such that, assuming that $\mathrm{DoF}(\tilde{H})=3/2-\epsilon(p,q)$, does not lead to a contradiction. The expression for $\epsilon(p,q)$ in the statement of the lemma is obtained by avoiding bounds on $H(\boldsymbol{x}_2^n+\boldsymbol{x}_3^n)$ until the last step in the analysis. For example, Lemma 10 is applied with

$$\begin{aligned} H\left(\boldsymbol{x}_2^n+\boldsymbol{x}_3^n\right)=&\left(\frac{H\left(\boldsymbol{x}_2^n+\boldsymbol{x}_3^n\right)}{\log_2(\max\{|\mathcal{X}_{2,n,m}|;|\mathcal{X}_{3,n,m}|\})}\right)\\ &\times\log_2(\max\{|\mathcal{X}_{2,n,m}|;|\mathcal{X}_{3,n,m}|\}). \end{aligned}$$

Note that $H(\boldsymbol{x}_2^n+\boldsymbol{x}_3^n)\neq H(\tilde{\boldsymbol{x}}_2^n+\tilde{\boldsymbol{x}}_3^n)$ and we still rely on the lower bound (50) for the latter entropy. The entropy $H(\boldsymbol{x}_2^n+\boldsymbol{x}_3^n)$ and cardinalities in the final expression are bounded by assuming that $\mathrm{DoF}(\tilde{H})=3/2-\epsilon(p,q)$, which can easily be shown, following steps similar to those leading to (40), to imply

$$\frac{H(\boldsymbol{x}_2^n+\boldsymbol{x}_3^n)}{n\log_2 P}\leq\left(\frac{1}{4}+\frac{\epsilon(p,q)}{2}\right)+o(1)$$

and

$$\begin{aligned} \left(\frac{1}{4}-\frac{\epsilon(p,q)}{2}\right)+o(1)&\leq\frac{\log_2(\min\{|\mathcal{X}_{2,n,m}|;|\mathcal{X}_{3,n,m}|\})}{n\log_2 P}\\ &\leq\frac{\log_2(\max\{|\mathcal{X}_{2,n,m}|;|\mathcal{X}_{3,n,m}|\})}{n\log_2 P}\\ &\leq\left(\frac{1}{4}+\frac{\epsilon(p,q)}{2}\right)+o(1). \end{aligned}$$

The details are omitted. □

D. Proof of Theorem 2

By assumption, all the entries of H are nonzero and rational, and therefore, there exists a diagonal matrix D_r with positive diagonal entries such that $\bar{H}\triangleq HD_r$ has nonzero integer entries. From Lemma 1 it follows that

$$\mathrm{DoF}(H)=\mathrm{DoF}(\bar{H}). \quad (54)$$

Consider the channel formed by the transmitters and receivers of users i,j, and k. Let $\bar{H}_{i,j,k}\in\mathbb{Z}^{3\times 3}$ be the principal minor (matrix) of $\bar{H}$ corresponding to the (i,j,k)th rows and columns. Due to the independence of the signals of the different users it follows that

$$\begin{aligned} &\max_{(R_1,\ldots,R_K)\in\mathcal{C}(\bar{H},\mathbf{1},P\mathbf{1})}R_i+R_j+R_k\\ &\qquad=\max_{(R_i,R_j,R_k)\in\mathcal{C}(\bar{H}_{i,j,k},\mathbf{1},P\mathbf{1})}R_i+R_j+R_k \end{aligned} \quad (55)$$

i.e., the other users cannot help users i, j, and k to improve their rates. In addition, since interference cannot help a given receiver in decoding the signal of interest it follows that we can set some of the cross-gains in $\bar{H}_{i,j,k}$ to zero without reducing the maximum achievable sum–rate. More specifically, defining $\hat{H}_{i,j,k}=[\hat{h}_{i,j,k}(m,n)]$ by $\hat{h}_{i,j,k}(m,n)\triangleq\bar{h}_{i,j,k}(m,n)\cdot 1(m\geq n)$ we have

$$\begin{aligned} &\max_{(R_i,R_j,R_k)\in\mathcal{C}(\bar{H}_{i,j,k},\mathbf{1},P\mathbf{1})}R_i+R_j+R_k\\ &\qquad\leq\max_{(R_i,R_j,R_k)\in\mathcal{C}(\hat{H}_{i,j,k},\mathbf{1},P\mathbf{1})}R_i+R_j+R_k. \end{aligned} \quad (56)$$

Furthermore, it is easy to see [9] that there exist diagonal matrices $\hat{D}_t$, $\hat{D}_r$ with positive diagonal entries such that $\hat{D}_t\hat{H}_{i,j,k}\hat{D}_r=\tilde{H}_{i,j,k}$ where

$$\tilde{H}_{i,j,k}=\begin{bmatrix}1&0&0\\1&p_{i,j,k}&0\\1&q_{i,j,k}&1\end{bmatrix}$$

[9]Letting $\hat{H}_{i,j,k}=[a,0,0;b,c,0;d,e,f]$ in Matlab matrix notation, we can choose $\hat{D}_t=[bd,0,0;0,ad,0;0,0,ab]$ and $\hat{D}_r=[1/(abd),0,0;0,1/a,0;0,0,1/(abf)]$.

$$\frac{R_{1,m}+R_{2,m}+R_{3,m}}{(1/2)\log_2(P_m)}\leq\frac{\frac{1}{2}\log_2\left[2\pi e\left(3P_m+\frac{1}{12}\right)\right]+\left[3\xi d(p,q)+3c\epsilon d(p,q)\left(\frac{1}{4}+\xi\right)\right]\log_2 P_m+\frac{\tilde{K}'_{c,p,q}}{n}+\delta'_n}{(1/2)\log_2 P_m} \quad (53)$$

for some $p_{i,j,k}, q_{i,j,k} \in \mathbb{Z}$, $p_{i,j,k}, q_{i,j,k} \neq 0$. Using (56), Lemma 1, and Lemma 11 we have

$$\begin{aligned}\mathrm{DoF}(\bar{H}_{i,j,k}) &\leq \mathrm{DoF}(\hat{H}_{i,j,k}) \\ &= \mathrm{DoF}(\tilde{H}_{i,j,k}) \leq \frac{3}{2} - \epsilon(p_{i,j,k}, q_{i,j,k}) \end{aligned} \quad (57)$$

where $\epsilon(p_{i,j,k}, q_{i,j,k}) > 0$.

Considering every possible subset of users $\{i,j,k\} \subseteq \{1,\ldots,K\}$ and adding the corresponding sum–rates, the rate of each user appears $\binom{K-1}{2}$ times in the sum. Therefore, we have (58) at the bottom of the following page, where (a) is due to (54), (b) follows from (55), (c) is obtained from (57), and where we defined $\delta \triangleq \min_{\{i,j,k\}\subseteq\{1,\ldots,K\}} \epsilon(p_{i,j,k}, q_{i,j,k}) > 0$. We finally obtain

$$\mathrm{DoF}(H) \leq \frac{K}{2} - \frac{K}{3}\delta < \frac{K}{2} \quad (59)$$

establishing the theorem.

V. A Three-User Rational GIC Example

In this section, we will derive lower and upper bounds on the degrees-of-freedom of the three-user GIC with channel matrix

$$H = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

represented in Fig. 3. This channel is a special case of the one considered in Lemma 11, with $p = 2$ and $q = 1$. From this lemma, we obtain

$$\mathrm{DoF}(H) \leq \frac{3}{2} - \frac{1}{12d(2,1)+2} \quad (60)$$

where $d(p,q) = 2\max\{|p|,|q|\} + 5$ was obtained in Lemma 7. For the special case of $q = 1$, the result of Lemma 7 can be easily strengthened to get $d(p,1) = 2|p| + 3$. Evaluating (60) with $d(2,1) = 7$ we obtain $\mathrm{DoF}(H) \leq 1.4884$. It should be possible to improve this bound in a number of ways, such as by improving on $d(2,1) = 7$, and possibly by improving on the power 3 in the term $(K')^3$ appearing in Theorem 3. Another possibility might be to forgo this theorem for a different approach, such as one based on [23, Exercise 2.5.4].

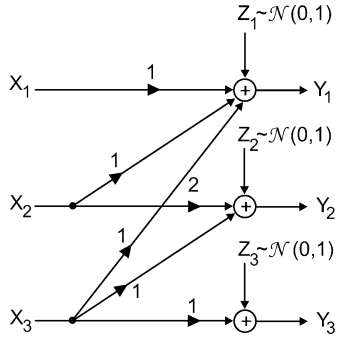

Fig. 3. Three-user GIC of the example in Section V.

To get a lower bound on $\mathrm{DoF}(H)$, we will describe a communication scheme that aims to achieve good interference alignment at receiver 1 by aligning the interfering signals of transmitters 2 and 3, while achieving good separation at receiver 2 between the signal of transmitter 2 and the interference of transmitter 3.

As was done in [13], we design the communication scheme for a deterministic interference channel and later show how to extend the scheme to the Gaussian channel. We derive the deterministic channel from the Gaussian IC by removing the Gaussian noise and constraining the inputs to be integers. Let $A_1 = \{0,1\}$, $A_2 = \{0,2,4\}$, $A_3 = \{0,2\}$, and $Q = 8$. We communicate information independently over L levels, without coding over time. The signal of user i at time t is given by

$$x_i(t) = \sum_{\ell=1}^{L} m_{i,\ell}(t) Q^{\ell-1} \quad (61)$$

where $m_{i,\ell}(t) \in A_i$ is the message of user i in level ℓ at time t.

Since $A_1 + A_2 + A_3 = \{0,1,2,3,4,5,6,7\}$, the signal at receiver 1 can be written as

$$y_1(t) = \sum_{i=1}^{3} x_i(t) = \sum_{\ell=1}^{L} w_{1,\ell}(t) Q^{\ell-1}$$

with $w_{1,\ell}(t) = m_{1,\ell}(t) + m_{2,\ell}(t) + m_{3,\ell}(t)$. Therefore, by computing the Q-ary decomposition of $y_1(t)$ we can recover the sums $m_{1,\ell}(t) + m_{2,\ell}(t) + m_{3,\ell}(t)$ at each of the L levels.

$$\begin{aligned}\binom{K-1}{2}\mathrm{DoF}(H) &\overset{(a)}{=} \binom{K-1}{2}\mathrm{DoF}(\bar{H}) \\ &= \limsup_{P\to\infty} \frac{\max_{(R_1,\ldots,R_K)\in\mathcal{C}(\bar{H},\mathbf{1},P\mathbf{1})} \binom{K-1}{2} \sum_{i=1}^{K} R_i}{\frac{1}{2}\log_2 P} \\ &\leq \limsup_{P\to\infty} \frac{\sum_{\{i,j,k\}\subseteq\{1,\ldots,K\}} \max_{(R_1,\ldots,R_K)\in\mathcal{C}(\bar{H},\mathbf{1},P\mathbf{1})} (R_i+R_j+R_k)}{\frac{1}{2}\log_2 P} \\ &\overset{(b)}{=} \limsup_{P\to\infty} \frac{\sum_{\{i,j,k\}\subseteq\{1,\ldots,K\}} \max_{(R_i,R_j,R_k)\in\mathcal{C}(\bar{H}_{i,j,k},\mathbf{1},P\mathbf{1})} (R_i+R_j+R_k)}{\frac{1}{2}\log_2 P} \\ &\leq \sum_{\{i,j,k\}\subseteq\{1,\ldots,K\}} \mathrm{DoF}(\bar{H}_{i,j,k}) \\ &\overset{(c)}{\leq} \binom{K}{3}\left(\frac{3}{2} - \delta\right) \end{aligned} \quad (58)$$

In addition, since $A_2 + A_3 = \{0,2,4,6\}$, we have $m_{1,\ell}(t) = 1(w_{1,\ell}(t) \in \{1,3,5,7\})$, so we can directly determine $m_{1,\ell}(t)$ from $w_{1,\ell}(t)$.

Similarly, at receiver 2 we compute

$$\frac{y_2(t)}{2} = x_2(t) + \frac{1}{2}x_3(t) = \sum_{\ell=1}^{L} w_{2,\ell}(t)Q^{\ell-1}$$

with $w_{2,\ell}(t) = m_{2,\ell}(t)+(1/2)m_{3,\ell}(t) \in \{0,1,2,3,4,5\}$, from which we can compute $m_{2,\ell}(t) = w_{2,\ell}(t) - [w_{2,\ell}(t) \bmod 2]$.

Finally, receiver 3 can directly recover $m_{3,\ell}(t)$ at all levels from the received signal $y_3(t) = x_3(t)$.

To compute the achievable degrees-of-freedom of this scheme we note that since $|x_i(t)| < Q^L$, the transmission power at each transmitter is smaller than Q^{2L}. On the other hand, the rate of users 1 and 3 is $L\log_2 2$ while the rate of user 2 is $L\log_2 3$. Therefore, we obtain for the deterministic channel

$$\begin{aligned}\mathrm{DoF} &\geq \frac{2L\log_2 2 + L\log_2 3}{\frac{2L}{2}\log_2 8} \\ &= \frac{2+\log_2 3}{3} \approx 1.19499.\end{aligned}$$

We now informally argue that the same degrees-of-freedom can be achieved in the Gaussian channel. We essentially use the same multilevel coding scheme, but we now encode the signals of each level over long blocks of time. The lower levels may be severely affected by noise, but as the level ℓ increases, the influence of the noise becomes smaller, ultimately being insignificant. As a result, the amount of redundancy that needs to be added to the signal of level ℓ to ensure low probability of decoding error goes to 0 as ℓ grows to infinity. It follows that for ℓ large enough, the achievable rates in the Gaussian channel at level ℓ approach the achievable rates in the deterministic channel, and since the rates of the lower levels do not affect the degrees-of-freedom, we conclude that $\mathrm{DoF}(H) \geq \frac{2+\log_2 3}{3}$ (see [13] for a similar argument).

In summary, using the lower and upper bounds that we derived we have

$$1.19499 \leq \mathrm{DoF}(H) \leq 1.4884.$$

Remark 2: The achievable scheme that we described is simple to analyze because there are no "carryovers" across the different levels, and the signals and interference are "orthogonal" in the sense that there is no need to code over time to ensure reliable decoding in the deterministic channel. In choosing the sets A_1, A_2, and A_3 we tried to obtain small $|A_2 + A_3|$ to align the interference at receiver 1 and simultaneously obtain large $|2A_2 + A_3|$ to achieve good signal–interference separation at receiver 2. With these design guidelines, one could optimize the sets A_1, A_2, and A_3 (with possibly larger Q) in order to improve the achievable degrees-of-freedom. In addition, Han–Kobayashi-type schemes [3] at each level where part of the interference is decoded and subtracted may result in better performance than purely orthogonal schemes.

VI. Conclusion

We have shown that the degrees-of-freedom of $K > 2$ user, real, scalar GICs is sensitive to whether the channel gains have rational or irrational values, and it is, in fact, discontinuous at all fully connected, rational gain matrices (up to the invariance property of Lemma 1). Specifically, Theorem 1 shows that certain fully connected real, scalar GICs with irrational, algebraic coefficients have degrees-of-freedom exactly equal to the known upper bound of $K/2$, the first such examples for real, scalar GICs. Theorem 2, on the other hand, shows that if all coefficients are nonzero rationals, the degrees-of-freedom is strictly bounded away from $K/2$, for $K > 2$. These theorems are established by appealing to major results in mathematics on the inapproximability of irrational, algebraic numbers by rational numbers, in the case of Theorem 1, and on the combinatorics of additive sets, in the case of Theorem 2. In the latter case, previously used information-theoretic converse techniques, which are not sensitive to the rationality of channel coefficients, do not suffice. We believe these results may have some implications for real GICs under channel parameter uncertainty, since in this case, channel coefficients with irrational and rational coefficients would have to be dealt with simultaneously. Additionally, in practical systems, computations for encoding and decoding are ultimately restricted to finite precision, and hence rational numbers, suggesting that additive combinatorics based bounds on achievable rates may have practical relevance.

Throughout this paper, we have been concerned with real, scalar GICs. Theorem 1 can be readily extended to the complex and vector cases, revealing an additional class of $K/2$ degrees-of-freedom vector GICs, complementing those already known [12]. The extension of Theorem 2 to complex and vector GICs seems less trivial. For instance, the example of a $K/2$ degrees-of-freedom achieving K-user two-dimensional vector GIC in [12] actually has integer coefficients, though they are a very special choice. Any extension of Theorem 2 would have to avoid such special cases. More significantly, our crucial Lemma 11 can be shown, using the interference alignment technique of [12], not to hold in the complex or vector cases. Nevertheless, we conjecture that for any fixed (complex) vector dimension, limitations on the degrees-of-freedom similar to Theorem 2 do exist for a sufficiently large number of users K. As noted, establishing such a result would require analyzing few-user GICs that are more complicated than the three-user channel of Lemma 11. It is likely that tools from additive combinatorics will still prove useful, though they would need to be applied differently from the proof of Lemma 11. The topic of characterizing the degrees-of-freedom of rational vector GICs and scalar, complex GICs having channel gains with rational real and imaginary parts is left for future work.

References

[1] R. Etkin and E. Ordentlich, "On the degrees-of-freedom of the K-user Gaussian interference channel," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Seoul, Korea, Jun./Jul. 2009, pp. 1919–1923.

[2] A. Carleial, "Interference channels," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 1, pp. 60–70, Jan. 1978.

[3] T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inf. Theory*, vol. IT-27, no. 1, pp. 49–60, Jan. 1981.

[4] I. Sason, "On achievable rate regions for the Gaussian interference channel," *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 1345–1356, Jun. 2004.

[5] X. Shang, G. Kramer, and B. Chen, "A new outer bound and noisy-interference sum-rate capacity for Gaussian interference channels," *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 689–699, Feb. 2009.
[6] A. S. Motahari and A. K. Khandani, "Capacity bounds for the Gaussian interference channel," *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 620–643, Feb. 2009.
[7] V. S. Annapureddy and V. V. Veeravalli, "Gaussian interference networks: Sum capacity in the low-interference regime and new outer bounds on the capacity region," *IEEE Trans. Inf. Theory*, vol. 55, no. 7, pp. 3032–3050, Jul. 2009.
[8] R. Etkin, D. M. C. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *IEEE Trans. Inf. Theory*, vol. 54, no. 12, pp. 5534–5562, Dec. 2008.
[9] R. Etkin, "New sum-rate upper bound for the two-user Gaussian interference channel," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Seoul, Korea, Jun./Jul. 2009, pp. 2582–2586.
[10] A. Høst-Madsen and A. Nosratinia, "The multiplexing gain of wireless networks," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Adelaide, Australia, Sep. 4–9, 2005, pp. 2065–2069.
[11] M. A. Maddah-Ali, A. S. Motahari, and A. K. Khandani, "Communication over MIMO X channels: Interference alignment, decomposition, and performance analysis," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3457–3470, Aug. 2008.
[12] 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.
[13] V. Cadambe, S. A. Jafar, and S. Shamai (Shitz), "Interference alignment on the deterministic channel and application to fully connected Gaussian interference networks," *IEEE Trans. Inf. Theory*, vol. 55, no. 1, pp. 269–274, Jan. 2009.
[14] G. Bresler, A. Parekh, and D. N. C. Tse, "The approximate capacity of the many-to-one and one-to-many Gaussian interference channels," in *Proc. Allerton Conf. Communications, Control and Computing*, Monticello, IL, Sep. 2007, pp. 791–801.
[15] C. Nair and A. El Gamal, "The capacity region of a class of 3-receiver broadcast channels with degraded message sets," *IEEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 4479–4493, Oct. 2009.
[16] H. Sato, "Two-user communication channels," *IEEE Trans. Inf. Theory*, vol. IT-23, no. 3, pp. 295–304, May 1977.
[17] H. Sato, "On degraded Gaussian two-user channels," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 5, pp. 637–640, Sep. 1978.
[18] A. B. Carleial, "Outer bounds on the capacity of interference channels," *IEEE Trans. Inf. Theory*, vol. IT-29, no. 4, pp. 602–606, Jul. 1983.
[19] G. Kramer, "Outer bounds on the capacity of Gaussian interference channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 3, pp. 581–586, Mar. 2004.
[20] 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.
[21] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New York: Wiley-Interscience, 2006.
[22] H. Stark, *An Introduction to Number Theory*, 4th ed. Cambridge, MA: MIT Press, 1984.
[23] T. Tao and V. H. Vu, *Additive Combinatorics*. Cambridge, U.K.: Cambridge Univ. Press, 2006.
[24] B. Bukh, "Sums of dilates," *Comb., Probab. Comput.*, vol. 17, no. 5, pp. 627–639, Sep. 2008.
[25] R. Ahlswede, "Multi-way communication channels," in *Proc. 2nd Int. Symp. Information Theory*, Tsahkadsor, Armenia, U.S.S.R., Sep.. 1971, pp. 23–52.
[26] R. Cheng and S. Verdú, "On limiting characterizations of memoryless multiuser capacity regions," *IEEE Trans. Inf. Theory*, vol. IT-32, no. 2, pp. 609–612, Mar. 1993.
[27] G. Kramer, "Capacity results for the discrete memoryless network," *IEEE Trans. Inf. Theory*, vol. 49, no. 1, pp. 4–21, Jan. 2003.

Raúl H. Etkin (S'05–M'07) received the B.S. degree in electrical engineering (with honors) from the University of Buenos Aires, Buenos Aires, Argentina, in 1998 and the M.S. and Ph.D. degrees in electrical engineering from the University of California, Berkeley, in 2003 and 2006, respectively.

During the summers of 2002 and 2003 he worked in the Corporate R&D Department of Qualcomm Inc., where he filed four patent applications. Since 2006, he has been a Researcher at Hewlett-Packard Laboratories, Palo Alto, CA, working with the information theory group, currently part of the Information and Quantum Systems Lab. His current research interests include multiuser information theory, wireless communications, and spectrum sharing.

Dr. Etkin received the Gold Medal award from the University of Buenos Aires, the Best Engineers award from the Argentine National Engineering Academy in 1999, and a Fulbright fellowship in 2000. From 2000 to 2002, he was supported by the University of California's Regents Fellowship.

Erik Ordentlich (S'92–M'96–SM'06) received the S.B. and S.M. degrees in electrical engineering from the Massachusetts Institute of Technology, Cambridge, MA, in 1990 and the Ph.D. degree, also in electrical engineering, from Stanford University, Stanford, CA, in 1996.

He is a Senior Research Scientist in the Information Theory Research Group at Hewlett-Packard Laboratories, Palo Alto, CA. He has been with Hewlett-Packard Laboratories since 1996, with the exception of a period in 1999–2002, when he was with iCompression, Inc., Santa Clara, CA. His work has addressed multiple topics in signal processing and information theory. He is a coinventor on 22 U.S. Patents and contributed technology to the ISO JPEG 2000 image compression standard.

Dr. Ordentlich was a corecipient of the 2006 IEEE Joint Communications/ Information Theory Paper Award and is a member of Phi Beta Kappa and Tau Beta Pi. He currently serves as Associate Editor for Source Coding for the IEEE TRANSACTIONS ON INFORMATION THEORY.