Aligned Image Sets Under Channel Uncertainty: Settling Conjectures on the Collapse of Degrees of Freedom Under Finite Precision CSIT

Arash Gholami Davoodi, *Student Member, IEEE*, and Syed Ali Jafar, *Fellow, IEEE*

Abstract—A conjecture made by Lapidoth *et al.* at Allerton 2005 (also an open problem presented at ITA 2006) states that the degrees of freedom (DoF) of a two user broadcast channel, where the transmitter is equipped with two antennas and each user is equipped with one antenna, must collapse under finite precision channel state information at the transmitter (CSIT). That this conjecture, which predates interference alignment, has remained unresolved, is emblematic of a pervasive lack of understanding of the DoF of wireless networks—including interference and X networks—under channel uncertainty at the transmitter(s). In this paper, we prove that the conjecture is true in all non-degenerate settings (e.g., where the probability density function of unknown channel coefficients exists and is bounded). The DoF collapse even when perfect channel knowledge for one user is available to the transmitter. This also settles a related recent conjecture by Tandon *et al.* The key to our proof is a bound on the number of codewords that can cast the same image (within noise distortion) at the undesired receiver whose channel is subject to finite precision CSIT, while remaining resolvable at the desired receiver whose channel is precisely known by the transmitter. We are also able to generalize the result along two directions. First, if the peak of the probability density function is allowed to scale as $O((\sqrt{P})^{\alpha})$, representing the concentration of probability density (improving CSIT) due to, e.g., quantized feedback at rate $(\alpha/2)\log(P)$, then the DoF is bounded above by $1+\alpha$, which is also achievable under quantized feedback. Second, we generalize the result to arbitrary number of antennas at the transmitter, arbitrary number of single-antenna users, and complex channels. The generalization directly implies a collapse of DoF to unity under non-degenerate channel uncertainty for the general K-user interference and $M \times N$ user X networks as well.

Index Terms—Degrees of freedom, broadcast channel, interference alignment, channel state information at the transmitter.

I. INTRODUCTION

INTERFERENCE alignment studies [1] have spurred much interest in the degrees of freedom (DoF) of wireless communication networks. While much progress has been

made under the assumption of perfect channel knowledge, the degrees of freedom under channel uncertainty at the transmitters have remained mostly a mystery. A prime example is the, heretofore unresolved, conjecture by Lapidoth, Shamai and Wigger from the Allerton conference in 2005 [2], also featured at the "*Open Problems Session*" at the Inaugural Information Theory and its Applications (ITA) workshop in 2006 [3], which claims that the DoF collapse under finite precision channel state information at the transmitter (CSIT). Specifically, Lapidoth et al. conjecture that the DoF of a 2 user multiple input single output (MISO) broadcast channel (BC) with 2 antennas at the transmitter and 1 antenna at each of the receivers, must collapse to unity (same as single user) if the probability distribution of the channel realizations, from the transmitter's perspective, is sufficiently well behaved that the differential entropy rate is bounded away from $-\infty$. The condition excludes not only settings where some or all channel coefficients are perfectly known, but also scenarios where some channel coefficients are functions of others, even if their values remain unknown. The best DoF outer bound under such channel uncertainty, also obtained by Lapidoth et al., is $\frac{4}{3}$. Deepening the mystery is the body of evidence on both sides of the conjecture. On the one hand, supporting evidence in favor of the collapse of DoF is available if the channel is essentially degraded, i.e., the users' channel vector directions are statistically indistinguishable from the transmitters' perspective [4], [5]. On the other hand, the idea of blind interference alignment introduced by Jafar in [6] shows that the 2 user MISO BC achieves $\frac{4}{3}$ DoF (which is also an outer bound, thus optimal), even without knowledge of channel realizations at the transmitter, provided that one user experiences time-selective fading and the other user experiences frequency-selective fading. Since the time-selective channel is assumed constant across frequency and the frequency-selective channel is assumed constant across time, it makes some channel coefficients functions of others (they are equal if they belong to the same coherence time/bandwidth interval), so that the model does not contradict the conjecture of Lapidoth et al. Thus, quite remarkably, this conjecture of Lapidoth, Shamai and Wigger, which predates interference alignment in wireless networks, has remained unresolved for nearly a decade.

Following in the footsteps of Lapidoth et al., subsequent works have made similar, sometimes even stronger conjectures, as well as partial attempts at proofs. For instance, the collapse of DoF of the MISO BC was also conjectured by

Weingarten, Shamai and Kramer in [7] under the finite state compound setting. However, this conjecture turned out to be too strong and was shown to be false by Gou, Jafar and Wang in [8], and by Maddah-Ali in [9], who showed that, once again, $\frac{4}{3}$ DoF are achievable (and optimal) for almost all realizations of the finite state compound MISO BC, regardless of how large (but finite) the number of states might be. Since the differential entropy of the channel process is not defined (approaches $-\infty$) for the finite state compound setting, this result also does not contradict the conjecture of Lapidoth et al. A related refinement of the conjecture, informally noted on several occasions (including by Shlomo Shamai at the ITA 2006 presentation) and mentioned most recently (although in the context of i.i.d. fading channels) by Tandon, Jafar, Shamai and Poor in [10] — is that the DoF should collapse even in the "PN" setting, where perfect (P) CSIT is available for one of the two users, while no (N) CSIT is available for the other user. A valiant attempt at proving this conjecture is made in [11], but it turns out to be unsuccessful because it relies critically on an incorrect use of the extremal inequality of [12] under channel uncertainty.[1] Thus the "PN" conjecture has also thus far remained unresolved.

That these conjectures remain unresolved, is emblematic of a broader lack of understanding of the DoF of wireless networks under non-degenerate forms of channel uncertainty. For instance, by extension, under non-degenerate channel uncertainty we also do not know the DoF of the vector broadcast channel with more than 2 users, or the DoF of interference networks, X networks, cellular, multi hop, or two-way relay networks, with or without multiple antennas, or any of a variety of settings with partial uncertainty, such as mixed [15], [16] or alternating [10] channel uncertainty. Thus, the resolution of these conjectures is likely to have a broad impact on our understanding of the "robustness" of the DoF of wireless networks. This is the motivation for our work in this paper.

A. Overview of Contribution

The main contribution of this work is to prove the conjecture of Lapidoth, Shamai and Wigger, thereby closing the ITA 2006 open problem, as well as the "PN" conjecture of Tandon et al., for all non-degenerate forms of finite precision CSIT, which includes all settings where density functions of the unknown channel realizations exist and are bounded. For all such settings, we show that the DoF collapse to unity as conjectured. Remarkably, this is the first result to show the total collapse of DoF under channel uncertainty without making assumptions of degradedness, or the (essentially) statistical equivalence of users.

Our approach, which is reminiscent of Korner and Marton's work on the images of a set in [17], is based on estimating the size of the images of the set of codewords as seen by the two users. Specifically, we bound the expected number of

codewords that are resolvable at their desired receiver whose images align (within bounded noise distortion) at the undesired receiver under finite precision CSIT. We show that this quantity is $\approx O((\log(P))^n)$ where n is the length of codewords, and P is the power constraint which defines the DoF limit as $P \to \infty$. This is negligible relative to the total number of resolvable codewords, which is $\approx O(P^{nd/2})$ when the desired information is sent at rate $\frac{d}{2}\log(P)$, i.e., with DoF $d > 0$ (normalization by $\frac{1}{2}\log(P)$ is because we initially deal with real channels). The difference between the entropy contributed by any set of codewords at their desired receiver (desired DoF) and the entropy contributed by the same set of codewords at the undesired receiver (DoF consumed by interference) tends to zero in the DoF sense. Under non-degenerate channel uncertainty, it is not possible to utilize the DoF at the desired receiver without sacrificing the same number of DoF at the undesired receiver due to interference. Therefore, the DoF are bounded above by unity, the same as with a single user.

We also generalize this result in several directions. First, we extend it to include CSIT that improves as $P \to \infty$, e.g., through quantized feedback at rate $\frac{\alpha}{2}\log(P)$, so that the probability density function of unknown channel coefficients concentrates around the correct realizations. This refinement of CSIT is captured by the growth in the peak value of the probability density function. We show that if the peak of the probability density function of unknown channel coefficients grows no faster than $O(P^{\frac{\alpha}{2}})$, representing e.g., improving channel quantization from feedback at rate $\frac{\alpha}{2}\log(P)$, then the total DoF are bounded above by $1 + \alpha$. Furthermore, with quantized feedback of rate $\frac{\alpha}{2}\log(P)$ this DoF bound is achievable.

Next, we go beyond 2 users and generalize the result to the K user MISO broadcast channel where the transmitter has K antennas and there are K users with a single antenna each. We also go beyond the restriction to real channels and generalize the results to complex channels. In all cases we prove that the DoF collapse to unity under non-degenerate channel uncertainty. Since the outer bound for this MISO BC is also an outer bound for the MISO BC with fewer than K antennas at the transmitter or fewer than K users, our result also establishes the collapse of DoF to unity for K user interference networks, and for $M \times N$ X channels, under non-degenerate channel uncertainty. Remarkably, the best known outer bounds for K user interference and $M \times N$ user X networks under non-degenerate channel uncertainty (except for essentially degraded settings) prior to this work were $\frac{K}{2}$ and $\frac{MN}{M+N-1}$ (same as with perfect CSIT). Thus, this work on finite precision CSIT and the work of Cadambe and Jafar in [18] where perfect CSIT was assumed, reveal a surprising contrast between the two sides of the same coin. In both cases the best previously known DoF outer bound was $K/2$ and the best previously known DoF inner bound was 1. Both works close this large gap. However, whereas under perfect CSIT, Cadambe and Jafar close the gap in the optimistic direction, showing that $K/2$ is optimal, in this work under finite-precision CSIT, we close the gap in the pessimistic direction, showing that 1 DoF is optimal.

---

[1]Similar problems arise in [13] and [14]. A simple counter example is the MISO BC with finitely many channel states, where the same arguments as used in these works would imply a collapse of DoF to 1, whereas this setting is known to have $\frac{4}{3}$ DoF as shown in [8] and [9].

B. Notation

We use the Landau $O(\cdot)$, $o(\cdot)$ and $\Theta(\cdot)$ notations as follows. For functions $f(x), g(x)$ from $\mathbb{R}$ to $\mathbb{R}$, $f(x) = O(g(x))$ denotes that $\limsup_{x \to \infty} \frac{|f(x)|}{|g(x)|} < \infty$. $f(x) = o(g(x))$ denotes that $\limsup_{x \to \infty} \frac{|f(x)|}{|g(x)|} = 0$. $f(x) = \Theta(g(x))$ denotes that there exists a positive finite constant, M, such that $\frac{1}{M} g(x) \leq f(x) \leq M g(x)$, $\forall x$. We use $\mathbb{P}(\cdot)$ to denote the probability function Prob($\cdot$). We define $\lfloor x \rfloor$ as the largest integer that is smaller than or equal to x when $x > 0$, the smallest integer that is larger than or equal to x when $x < 0$, and x itself when x is an integer. The index set $\{1, 2, \cdots, n\}$ is represented compactly as $[1 : n]$ or simply $[n]$ when it would cause no confusion. $X^{[s]}$ represents the random vector $(X(1), X(2), \cdots, X(s))$ and $\{X^{[s]}\}$ represents the set $\{X(t) : t \in [s]\}$. The cardinality of a set A is denoted as $|A|$. The support of a random variable is denoted as supp(X).

II. THE 2 USER MISO BC WITH PERFECT CSIT FOR ONE USER

To prove the collapse of DoF in the strongest sense possible, let us first enhance the 2 user MISO BC by allowing perfect CSIT for user 1. Consider the vector broadcast channel with 2 users where the transmitter is equipped with 2 antennas, each user is equipped with 1 receive antenna, and there are 2 independent messages W_1, W_2 that originate at the transmitter and are desired by users 1 and 2, respectively. The transmission takes place over n channel uses. The channel state information at the transmitter (CSIT) is denoted as $\mathcal{T}$, and includes perfect channel state information for the channel vector of user 1 but not for the channel vector of user 2. In the terminology of Tandon et al. [10], this is the PN setting, although not restricted to any statistical equivalence assumptions.

The best outer bound for the DoF of the PN setting based on known results so far is $\frac{3}{2}$, which is obtained from the finite state compound model by Weingarten, Shamai and Kramer in [7] and is applicable to finite precision CSIT as well. While Weingarten et al. conjectured that their outer bound was loose even in the finite state compound setting, predicting a collapse of DoF, this conjecture was shown to be false by Gou, Jafar and Wang in [8], who showed that $\frac{3}{2}$ DoF are achievable under the finite state compound model, through the DoF tuple $(d_1, d_2) = (1, 0.5)$. The key to achievability is to split user 1's 1 DoF into two parts that carry 0.5 DoF each. These parts align at user 2, consuming half the available signal space of user 2, while remaining resolvable at user 1. User 2's signal, carrying 0.5 DoF, is then sent in the null space of user 1's channel, and is resolvable from the 0.5 dimensional interference-free space at user 2. Note that zero forcing at user 1 is possible because perfect CSIT for user 1 is assumed to be available.

The $\frac{3}{2}$ DoF outer bound is also applicable in the blind interference alignment setting (BIA) introduced by Jafar in [6], where user 1 experiences time or frequency selective fading but user 2 experiences a relatively flat fading channel. Here also the outer bound is shown to be achievable through the pair $(d_1, d_2) = (1, 0.5)$. The key is to send two symbols for user 1, one from each antenna, repeated over two channel

realizations where the channel of user 1 changes but the channel of user 2 remains the same. Thus, user 1 sees two linear combinations of the two symbols from which both symbols can be resolved, whereas user 2 only sees the same linear combination over both channel uses. Thus the interference occupies only 0.5 DoF at user 2. The remaining 0.5 DoF at user 2 is utilized by sending his desired signal, carrying 0.5 DoF, into the null space of user 1.

The finite state compound setting and the blind interference alignment setting reveal some of the challenges of proving the collapse of DoF for the PN setting. Any attempt at proving a collapse of DoF must carefully exclude such scenarios from the channel model. With this cautionary note, we are now ready to introduce the channel model for our problem.

A. General Channel Model

Following [2], we will start with the real channel model, where all symbols take only real values. The extension to complex channels is cumbersome but conceptually straightforward, and will be presented in Section VII for the sake of completeness.

The channel is described as follows:

$$\begin{bmatrix} \tilde{Y}_1(t) \\ \tilde{Y}_2(t) \end{bmatrix} = \underbrace{\begin{bmatrix} \tilde{G}_{11}(t) & \tilde{G}_{12}(t) \\ \tilde{G}_{21}(t) & \tilde{G}_{22}(t) \end{bmatrix}}_{\tilde{\mathbf{G}}(t)} \begin{bmatrix} \tilde{X}_1(t) \\ \tilde{X}_2(t) \end{bmatrix} + \begin{bmatrix} \tilde{Z}_1(t) \\ \tilde{Z}_2(t) \end{bmatrix} \quad (1)$$

where all symbols are real. At time $t \in \mathbb{N}$, $\tilde{Y}_k(t)$ is the symbol received by user k, $\tilde{Z}_k(t) \sim \mathcal{N}(0, 1)$ is the real additive white Gaussian noise (AWGN), $\tilde{X}_k(t)$ is the symbol sent from transmit antenna k, and $\tilde{G}_{kj}(t)$ is the channel fading coefficient between the jth transmit antenna and user k. The channel coefficients are not restricted to i.i.d. realizations, but are assumed to be drawn from a continuous distribution such that the joint density of $\tilde{\mathbf{G}}(t)$ exists. The transmitter is subject to the power constraint:

$$\frac{1}{n} \sum_{t=1}^{n} [(\tilde{X}_1(t))^2 + (\tilde{X}_2(t))^2] \leq \tilde{P} = O(P), \quad (2)$$

To avoid degenerate situations we will assume that the range of values of each of the elements $\tilde{G}_{ij}$ is bounded away from zero and infinity, as is the determinant of the overall channel matrix — i.e., $|\tilde{G}_{ij}(t)|$, $\det(\tilde{\mathbf{G}}(t))$ are all $\Theta(1)$. Stated explicitly, there exists positive finite constant M, such that

$$\frac{1}{M} \leq |\tilde{G}_{ij}(t)|, \det(\tilde{\mathbf{G}}(t)) \leq M \quad (3)$$

Note that this is not a major restriction because by choosing the bounding constants large enough, the omitted neighborhoods can be reduced to a probability measure less than ϵ for arbitrarily small ϵ, and thus has only a vanishing impact on the DoF.

B. Canonical Form

For the purpose of deriving a DoF outer bound it suffices to work with a simplified channel with fewer parameters. (See Appendix B for justification). The simplified form of the
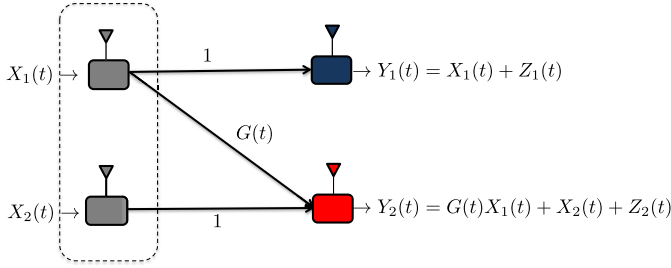
Fig. 1. Canonical form of the 2 user MISO BC with perfect CSIT for user 1.

channel model, shown in Fig. 1 has outputs $Y_1(t), Y_2(t) \in \mathbb{R}$, and inputs $X_1(t), X_2(t) \in \mathbb{R}$, so that:

$$\begin{bmatrix} Y_1(t) \\ Y_2(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ G(t) & 1 \end{bmatrix} \begin{bmatrix} X_1(t) \\ X_2(t) \end{bmatrix} + \begin{bmatrix} Z_1(t) \\ Z_2(t) \end{bmatrix} \qquad (4)$$

The channel coefficient $G(t)$ is also bounded away from zero and infinity, i.e., there exists finite positive M, such that $|G(t)| \in \left(\frac{1}{M}, M\right)$. The new power constraint is expressed as

$$\frac{1}{n} \sum_{t=1}^{n} [(X_1(t))^2 + (X_2(t))^2] \le P, \qquad (5)$$

where $P = \Theta(\tilde{P})$. Further, for notational convenience let us define the set of admissible inputs.

$$\mathcal{X}^{[n]} \triangleq \{(X_1^{[n]}, X_2^{[n]}) : \frac{1}{n} \sum_{t=1}^{n} [(X_1(t))^2 + (X_2(t))^2] \le P\}$$

C. Messages, Rates, Capacity, DoF

The messages W_1, W_2 are jointly encoded at the transmitter for transmission over n channel uses at rates R_1, R_2, respectively, into a $2^{nR_1+nR_2} \times n$ codebook matrix over the input alphabet. The codebook is denoted by $C(n, (R_1, R_2), P)$. For given power constraint parameter P, the rate vector (R_1, R_2) is said to be achievable if there exists a sequence of codebooks $\mathcal{C}(n, (R_1, R_2), P)$, indexed by n, such that the probability that all messages are correctly decoded by their desired receivers approaches 1 as n approaches infinity. The closure of achievable rate vectors is the capacity region $\mathcal{C}(P)$. The DoF pair (d_1, d_2) is said to be achievable if there exist $(R_1(P), R_2(P)) \in \mathcal{C}(P)$ such that

$$d_1 = \lim_{P \to \infty} \frac{R_1(P)}{\frac{1}{2} \log(P)}, \quad d_2 = \lim_{P \to \infty} \frac{R_2(P)}{\frac{1}{2} \log(P)}$$

The closure of all achievable DoF tuples (d_1, d_2) is called the DoF region, $\mathcal{D}$. The sum-DoF value is defined as

$$\mathcal{D}_{\Sigma} = \max_{(d_1,d_2) \in \mathcal{D}} (d_1 + d_2)$$

D. Non-Degenerate Channel Uncertainty

Let us denote by $\mathcal{T}$, all available channel state information at the transmitter. Then, non-degenerate channel uncertainty corresponds to the assumption that the conditional probability density functions of channel coefficients exist and are bounded, as explained next.

1) Peak of Density Function Is Bounded for Fixed P: For the G_{ij} the transmiters are only aware of the joint probability density function (PDF). $\forall k \in [2 : K]$, define the set of channel coefficient variables

$$\{G_k^{[n]}\} = \{G_{ki}(t) : i \in [1 : k-1], t \in [n]\}$$

Define $\{G^{[n]}\} = \bigcup_{k \in [1:K]} G_k^{[n]}$. Finite precision CSIT corresponds to the existence of bounded density functions. Precisely, the finite precision CSIT model assumes that there exists a finite positive constant f_{max},

$$0 < f_{max} < \infty$$

such that $\forall n \in \mathbb{N}$, and for all finite cardinality disjoint subsets $\{G_1\}, \{G_2\}$ of $\{G^{[n]}\}$,

$$\forall \{G_1\} \subset \{G^{[n]}\}, \ \{G_2\} \subset \{G^{[n]}\}, \{G_1\} \cap \{G_2\} = \varnothing,$$
$$|\{G_1\}| < \infty, |\{G_2\}| < \infty$$

the conditional PDF

$$f_{G_1|\mathcal{T},G_2}(g_1|\tau, g_2) \le f_{max}^{|\{G_1\}|}.$$

The condition implies that a zero measure set cannot carry a non-zero probability. So it precludes scenarios where, e.g., the channel is perfectly known or when one channel coefficient is a function of the rest. In all such cases, a zero measure set carries a non-zero probability, thus precluding the existence of a bounded constant f_{max} as defined above. This restriction essentially accomplishes the same goal as the restriction by Lapidoth et al. [2] that the differential entropy should be greater than $-\infty$.

For example, if conditioned on the available CSIT $\mathcal{T}$, the channel realizations are independent, then we can simply choose f_{max} as the peak value of the marginal density functions.

2) Peak of Density Function Is Allowed to Scale With P: To model CSIT that improves as a function of P, we allow $f_{max}(P)$ to scale as $O((\sqrt{P})^\alpha)$ for some $\alpha \in [0, 1]$.

$$f_{max}(P) = O(P^{\frac{\alpha}{2}}) \qquad (6)$$

The case studied by Lapidoth et al. in [2], where the density does not depend on P, is represented here by setting $\alpha = 0$. The positive values of α allow us to address settings where the CSIT improves with P, e.g., due to quantized channel feedback of rate $\frac{\alpha}{2} \log(P)$, so that the weight of the distribution is increasingly concentrated around the true channel realizations. Note that the maximum value of α is unity, because a feedback rate of $\frac{1}{2} \log(P)$, implying 1 real DoF worth of feedback, is sufficient to approach perfect CSIT performance over channels that take only real values [19], [20].

Since the receivers have full channel state information, $\mathcal{T}$ is globally known. For compact notation, we will suppress the conditioning, writing $f_{G^{[n]}}(g^{[n]})$ directly instead.

E. K User Extension

Extending beyond the 2 user case, the simplified channel model in the K user setting is described as follows.

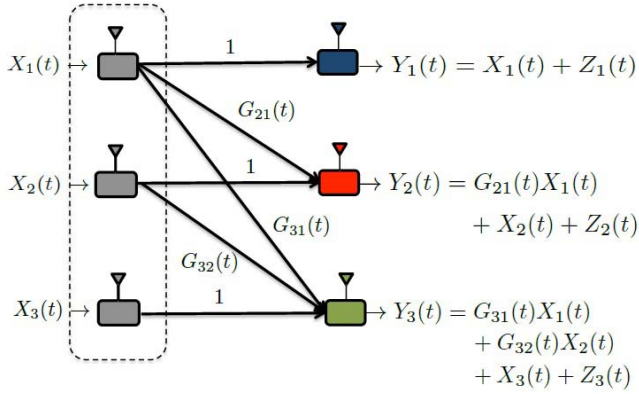$$Y_1(t) = X_1(t) + Z_1(t) \qquad (7)$$

Fig. 2. Canonical form of the 3 user MISO BC with Graded Channel Uncertainty.

$$Y_2(t) = G_{21}(t)X_1(t) + X_2(t) + Z_2(t) \qquad (8)$$
$$\vdots$$
$$Y_K(t) = G_{K1}(t)X_1(t) + G_{K2}(t)X_2(t) + \cdots$$
$$+ G_{K(K-1)}(t)X_{K-1}(t) + X_K(t) + Z_K(t) \qquad (9)$$

where the inputs, $X_k(t) \in \mathbb{R}$, are subject to the power constraint

$$\frac{1}{n}\sum_{t=1}^{n}\left((X_1(t))^2 + (X_2(t))^2 + \cdots + (X_K(t))^2\right) \le P \qquad (10)$$

The $G_{ij}(t)$ terms are known to the transmitter only up to finite precision and are assumed to be bounded away from 0 and infinity. Further, the density of the kth users' unknown channel coefficients, $k > 1$, is bounded by $f_{max,k}(P) = O\left(P^{\frac{a_k}{2}}\right)$.

III. RESULTS

We state the main result in its most general form, for K users. The 2 user case, corresponds to setting $\alpha_2 \triangleq \alpha$.

Theorem 1: For the K user MISO BC with non-degenerate channel uncertainty, the sum-GDoF are bounded above as

$$\mathcal{D}_\Sigma \le 1 + \alpha_2 + \alpha_3 + \cdots + \alpha_K \qquad (11)$$

A. Settling the Conjecture by Lapidoth et al. in [2]

The 2 user setting studied by Lapidoth et al., where the joint pdf is fixed, i.e., it does not depend on P, is captured here when $\alpha = 0$ (equivalently, $\alpha_2 = 0$). When $\alpha = 0$, the sum-GDoF are bounded above by unity, thus settling the conjecture of Lapidoth et al. for non-degenerate channel uncertainty models.

B. Settling the "PN" Conjecture

Since we allow perfect CSIT for one user, and one may assume (as a special case of our result) that the channels are i.i.d., the collapse of DoF for $\alpha = 0$, also proves the conjecture of Tandon et al. for the 2 user setting.

C. Interference and X Networks

Consider any one-hop wireless network where all receivers are equipped with a single antenna each. This includes all interference and X networks. Allowing the transmitters to cooperate produces a MISO BC setting. Since cooperation cannot hurt, the outer bound for the MISO BC under non-degenerate channel uncertainty applies to interference and X networks as well. In all cases, the DoF collapse to unity.

D. Limited Rate Feedback ($\alpha > 0$)

Consider the 2 user setting, with $\alpha = \alpha_2 > 0$. This case is interesting because it has direct implications to the achievable DoF under limited rate quantized channel state feedback for the channel vector of user 2. If the feedback link has α DoF, i.e., the feedback rate scales as $\frac{\alpha}{2}\log(P)$ bits per channel use, then this corresponds to $\sim P^{\frac{\alpha}{2}}$ channel quantization levels, so that the size of a quantization interval scales as $\frac{1}{(\sqrt{P})^\alpha}$ and the channel density restricted to a quantization interval, i.e., $f_{G^{[n]}}(g^{[n]}|\mathcal{T})$ scales as $P^{\frac{\alpha}{2}}$. Theorem 1 tells us that in this case the GDoF are bounded above as $\mathcal{D}_\Sigma \le 1 + \alpha$. It is also easy to see that under such quantized feedback, the DoF tuple $(d_1, d_2) = (1, \alpha)$ is achievable, simply by best-effort zero-forcing at the transmitter and treating residual interference as noise at the receiver 2. Thus, $\mathcal{D}_\Sigma = 1 + \alpha$ is the optimal sum-DoF value if the quantized channel state feedback is limited to rate $\frac{\alpha}{2}\log(P)$. This generalizes the results from [19] and [20] where it was shown that in order to achieve the same DoF as with perfect CSIT, i.e., $\mathcal{D}_\Sigma = 2$, the quantized feedback rate should scale as $\frac{1}{2}\log(P)$, i.e., carry one full degree of freedom ($\alpha = 1$). The bound for the K user extension is similarly tight as well.

IV. ALIGNED IMAGE SETS UNDER CHANNEL UNCERTAINTY

The main idea we want to illustrate intuitively in this section is a geometric notion of aligned images of codewords—loosely related to Korner and Marton's work on the images of a set in [17] but under a much more specialized setting—which is the key to our proof. As the proof in Section V will show, the problem boils down to the difference of two terms when only information to user 1 is being transmitted,

$$\mathcal{D}_\Sigma \le 1 + \limsup_{P\to\infty}\limsup_{n\to\infty}\frac{h(Y_1^{[n]}|G^{[n]}) - h(Y_2^{[n]}|G^{[n]})}{\frac{n}{2}\log(P)} \qquad (12)$$

The first term, $h(Y_1^{[n]}|G^{[n]})$, we wish to maximize because it represents the rate of desired information being sent to user 1. The second, $h(Y_2^{[n]}|G^{[n]}) = h(G^{[n]}X_1^{[n]} + X_2^{[n]} + Z_2^{[n]}|G^{[n]})$ we wish to minimize, because it represents the interference seen by user 2, due to the information being sent to user 1. If $G^{[n]}$ was perfectly available to the transmitter, then $X_2^{[n]}$ could be chosen to cancel $G^{[n]}X_1^{[n]}$ thus eliminating interference entirely at user 2. With only statistical knowledge of $G^{[n]}$, zero forcing is not possible. Indeed, the purpose of $X_2^{[n]}$ is mainly to align interference into as small a space as possible. However, instead of consolidating interference in the sense

of vector space dimensions, as is typically the case in DoF studies involving interference alignment, here the goal is for $X_2^{[n]}$ to minimize the size of the image, as seen by user 2, of the codewords that carry information for user 1. This is the new perspective that is the key to the proof.

A. Toy Setting to Introduce Aligned Image Sets

For illustrative purposes, let us start with a rather extreme over-simplification, by considering the case with $n = 1$ channel use, ignoring noise, and using the log of the cardinality of the codewords as a surrogate for the entropy. With this simplification, the quantity that we are interested in is the difference:

$$\log |\text{supp}(X_1)| - \log |\text{supp}(GX_1 + X_2)| \qquad (13)$$

averaged over G. Recall that by $|\text{supp}(A)|$ is meant the cardinality of the set of values taken by the random variable A.

The codebook is the set of (X_1, X_2) values. Note that $|\text{supp}(X_1)|$, the number of distinct values of X_1, is the number of distinct "codewords" that can be seen by user 1, who (once noise is ignored) only sees $Y_1 = X_1$, so that his "rate" is $\log |\text{supp}(X_1)|$. Given the set of X_1 values, we would like to associate each X_1 value with a corresponding X_2 value, such that the number of distinct values of $Y_2 = GX_1 + X_2$ is minimized. In other words, we wish to minimize the image of the set of codewords as seen by user 2, by choosing X_2 to be a suitable function of X_1.

Consider two codewords $(X_1, X_2) = (x_1, x_2)$ and $(X_1, X_2) = (x_1', x_2')$. If $x_1 \neq x_1'$ then these codewords are distinct from user 1's perspective, and thus capable of carrying information to user 1 via the transmitter's choice to transmit one or the other. Suppose the channel is G. Then for these two codewords to "*align*" where they cause interference, they must have the same image as seen by user 2. This gives us the condition for aligned images that is central to this work.

$$Gx_1 + x_2 = Gx_1' + x_2' \qquad (14)$$

$$\Rightarrow G = -\left(\frac{x_2' - x_2}{x_1' - x_1}\right) \qquad (15)$$

In other words, G must be the negative of the slope of the line connecting the codeword (x_1, x_2) to the codeword (x_1', x_2') in the X_1, X_2 plane. For a given channel realization G, all codewords that align with (x_1, x_2) (i.e., whose images align with the image of (x_1, x_2)) as seen by user 2, must lie on the same line that passes through (x_1, x_2) and has slope $-G$. Conversely, all codewords that lie on this line have images that align with the image of (x_1, x_2) at user 2. For any codeword that does not lie on this line, there is a parallel line with the same slope, $-G$, that represents the set of codewords whose images align with the image of that codeword. Thus, these lines of the same slope, $-G$, partition the set of codewords into equivalence classes, such that all codewords that lie on the same line have the same image at user 2. Also note that a different channel realization, G', gives rise to a different equivalent class partition, corresponding to lines with slope $-G'$. This is illustrated in Fig. 3. Since the X_2 values are functions of X_1 values, in the figure we label the codewords only on the
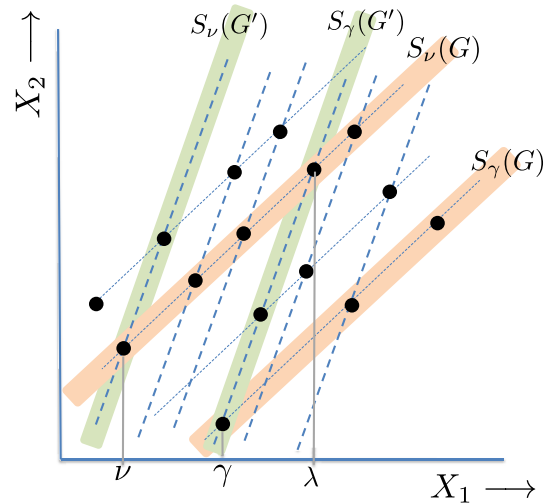

Fig. 3. Two codewords, corresponding to $X_1 = \nu$ and $X_1 = \gamma$, and their equivalence classes, S_ν and S_γ, containing all codewords (X_1, X_2) that have the same image at user 2 as the codewords corresponding to $X_1 = \nu$ and $X_1 = \gamma$, respectively. The partitioning into equivalence classes depends on the channel realization. The figure shows the distinct equivalence classes for two channel realizations, G and G'.

X_1 axis. So, for example, the codeword $(X_1, X_2) = (\nu, X_2(\nu))$ is simply referred to as the codeword ν. This codeword belongs to the equivalence class $S_\nu(G)$ under the channel realization G and to the equivalence class $S_\nu(G')$ under the channel realization G'. Also, note that two codewords that belong to the same equivalence class under one channel realization, cannot belong to the same equivalence class under any other channel realization. For instance, codewords λ and ν belong to the same equivalence class $S_\nu(G)$ under channel realization G, but they belong to different equivalence classes, $S_\nu(G')$ and $S_\gamma(G')$, under a different channel realization G'.

B. Sketch of Proof

Staying with the intuitive character of this section, let us conclude with an outline which will be useful to navigate the structure of the proof that appears in the subsequent section.

From the perspective of DoF studies, the presence of noise essentially imposes a resolution threshold, e.g., δ, such that the codewords with images that differ by less than δ, are unresolvable. As the first step of the proof, this effect is captured by discretizing the input and output alphabet and eliminating noise, as is done in a variety of deterministic channel models that have been used for DoF studies [21]–[23], so that instead of differential entropies we now need to deal only with entropies $H(\bar{Y}_1^{[n]}|G^{[n]})$ and $H(\bar{Y}_2^{[n]}|G^{[n]})$. Here $\bar{X}_1, \bar{X}_2$ represent the discretized inputs, $\bar{Y}_1, \bar{Y}_2$ the discretized outputs, and $\bar{Y}_1 = \bar{X}_1$.

The next step is to note that we are only interested in the maximum value of the difference $H(\bar{Y}_1^{[n]}|G^{[n]}) - H(\bar{Y}_2^{[n]}|G^{[n]})$. It then follows that without loss of generality, $\bar{X}_2^{[n]}$ can be made a function of $\bar{X}_1^{[n]}$, and therefore $\bar{Y}_2^{[n]}$ becomes a function of $\bar{Y}_1^{[n]}, G^{[n]}$. This implies that $H(\bar{Y}_1^{[n]}|G^{[n]}) = H(\bar{Y}_1^{[n]}, \bar{Y}_2^{[n]}|G^n) = H(\bar{Y}_2^{[n]}|G^{[n]}) + H(\bar{Y}_1^{[n]}|\bar{Y}_2^{[n]}, G^{[n]})$. Thus, the difference of entropies is equal to $H(\bar{Y}_1^{[n]}|\bar{Y}_2^{[n]}, G^{[n]}) = H(\bar{X}_1^{[n]}|\bar{Y}_2^{[n]}, G^{[n]})$.

Now, conditioned on $\bar{Y}_2^{[n]}, G^{[n]}$, the set of feasible values of $\bar{X}_1^{[n]}$ is an aligned image set $S(G^{[n]})$, i.e., all these $\bar{X}_1^{[n]}$ produce the same value of $\bar{Y}_2^{[n]}$ for the given channel realization $G^{[n]}$.

Since entropy is maximized by a uniform distribution,

$$H(\bar{X}_1^{[n]}|\bar{Y}_2^{[n]}, G^{[n]}) \leq \mathrm{E}_{G^n}\left[\log\left[|S_{\bar{X}_1^{[n]}}(G^{[n]})|\right]\right]$$
$$\leq \log\left(\mathrm{E}_{G^n}\left[|S_{\bar{X}_1^{[n]}}(G^{[n]})|\right]\right)$$

where the last step uses Jensen's inequality. Thus, the difference of entropies is bounded by the log of the expected cardinality of the aligned image sets.

The most critical step of the proof then is to bound the expected cardinality of aligned image sets. This is done by bounding the probability that two given $\bar{X}_1^{[n]}$ are in the same aligned image set, i.e., the probability of the set of channels for which the two produce the same image $\bar{Y}_2^{[n]}$. Recall that for two codewords to belong to the same aligned set in the absence of noise, the channel realization over each channel use must be the slope of the vector connecting the corresponding codeword vectors. The blurring of δ around the two codewords also blurs the slope of the line connecting them, but by no more than $\pm\delta/\Delta$, where Δ is the distance (difference in magnitudes) between the two codeword symbols over that channel use. Thus, the probability that the given two codewords that are resolvable at user 1 cast the same image at user 2 is bounded above by $\approx f_{\max}\frac{2\delta}{\Delta}$. The power constraint of P implies that there are at most $\approx \sqrt{P}/\delta$ resolvable codeword symbols per channel use. Summing over all possible resolvable codeword symbols, gives us $\approx f_{\max}\sum_{\Delta\in[0:\sqrt{P}/\delta]}\frac{2\delta}{\Delta} = f_{\max}\delta\log(P) + o(\log(P))$, per channel use, so that the average cardinality of an aligned image set, $E|S(G^{[n]})|$, turns out to be bounded above by $\approx (f_{\max}\delta\log(P))^n$, and $\log\left(\mathrm{E}\left[|S_{\bar{X}_1^{[n]}}(G^{[n]})|\right]\right)$ is bounded above by $\approx n\log(f_{\max}\delta) + n\log(\log(P))$.

Since $f_{\max} = O(P^{\frac{\alpha}{2}})$, normalizing by $\frac{n}{2}\log(P)$ and sending first n and then P to infinity sends this term to α. Thus, combining with (12) produces the sum-DoF outer bound value $1 + \alpha$, giving us the result of Theorem 1. Note that in the DoF limit, $\delta = \Theta(1)$, and it will be useful to think of it as 1 for simplicity, so that the inputs and outputs are restricted to integer values. With this sketch as the preamble, we now proceed to the actual proof.

V. Proof of Theorem 1 for $K = 2$ Users

For ease of exposition, the proof is divided into several key steps. The first step is the discretization of the channel to capture the effect of noise. This leads to a deterministic channel model. The DoF of the deterministic channel model will be shown to be an outer bound to the DoF of the canonical channel model, which in turn is an outer bound on the DoF of the general channel model.

1) Deterministic Channel Model

The deterministic channel model has inputs $\bar{X}_1(t)$, $\bar{X}_2(t) \in \mathbb{Z}$ and outputs $\bar{Y}_1(t), \bar{Y}_2(t) \in \mathbb{Z}$, defined as

$$\bar{Y}_1(t) = \bar{X}_1(t) \tag{16}$$
$$\bar{Y}_2(t) = \lfloor G(t)\bar{X}_1(t)\rfloor + \bar{X}_2(t) \tag{17}$$

with the discretization of inputs subject to the power constraint

$$\bar{X}_1(t), \bar{X}_2(t) \in \{0, 1, \cdots, \lceil\sqrt{P}\rceil\}, \ \forall t \in \mathbb{N} \tag{18}$$

and the set of inputs that satisfy the power constraints defined as

$$\bar{\mathcal{X}}^{[n]} = \{(\bar{X}_1^{[n]}, \bar{X}_2^{[n]}) \in \mathbb{Z}^{[n]} \times \mathbb{Z}^{[n]} : \bar{X}_1(t), \bar{X}_2(t)$$
$$\in \{0, 1, \cdots, \lceil\sqrt{P}\rceil\}, \quad \forall t \in [1:n]\} \tag{19}$$

The bounded density assumptions on the unknown channel coefficients sequence $G^{[n]}$ are the same as before.

Lemma 1: The DoF of the canonical channel model are bounded above by the DoF of the deterministic channel model. The proof of Lemma 1 appears in Appendix and follows along the lines of similar proofs by Bresler and Tse in [22].

2) Difference of Entropies Representing Desired Signal and Interference Dimensions

Starting from Fano's inequality, we proceed as follows.

$$nR_1 \leq I(W_1; \bar{Y}_1^{[n]}|W_2, G^{[n]}) + o(n) \tag{20}$$
$$= H(\bar{Y}_1^{[n]}|W_2, G^{[n]}) + o(n) \tag{21}$$
$$nR_2 \leq I(W_2; \bar{Y}_2^{[n]}|G^{[n]}) + o(n) \tag{22}$$
$$= H(\lfloor G^{[n]}\bar{X}_1^{[n]}\rfloor + \bar{X}_2^{[n]}|G^{[n]})$$
$$\quad - H(\bar{Y}_2^{[n]}|W_2, G^{[n]}) + o(n) \tag{23}$$
$$\leq \frac{n}{2}\log(P) - H(\bar{Y}_2^{[n]}|W_2, G^{[n]})$$
$$\quad + n\ o(\log(P)) + o(n) \tag{24}$$
$$\Rightarrow n(R_1 + R_2)$$
$$\leq \frac{n}{2}\log(P) + [H(\bar{Y}_1^{[n]}|W_2, G^{[n]})$$
$$\quad - H(\bar{Y}_2^{[n]}|W_2, G^{[n]})] + n\ o(\log(P)) + o(n)$$
$$\Rightarrow \mathcal{D}_\Sigma$$
$$\leq 1 + \limsup_{P\to\infty}\limsup_{n\to\infty}$$
$$\quad \times \frac{[H(\bar{Y}_1^{[n]}|W_2, G^{[n]}) - H(\bar{Y}_2^{[n]}|W_2, G^{[n]})]}{\frac{n}{2}\log(P)} \tag{25}$$
$$\leq 1 + \limsup_{P\to\infty}\limsup_{n\to\infty}\max_{w_2\in[1:2^{nR_2}]}$$
$$\quad \times \frac{H(\bar{X}_1^{[n]}|G^{[n]}) - H(\lfloor G^{[n]}\bar{X}_1^{[n]}\rfloor + \bar{X}_2^{[n]}|G^{[n]})}{\frac{n}{2}\log(P)}$$
$$= 1 + \bar{\mathcal{D}}_\Delta \tag{26}$$

so that what remains is to bound the difference of entropy terms:

$$\bar{\mathcal{D}}_\Delta \triangleq \limsup_{P\to\infty}\limsup_{n\to\infty}\max_{\substack{\mathbb{P}(\bar{X}_1^{[n]},\bar{X}_2^{[n]})\\(\bar{X}_1^{[n]},\bar{X}_2^{[n]})\in\bar{\mathcal{X}}^{[n]}}}$$
$$\times \frac{H(\bar{X}_1^{[n]}|G^{[n]}) - H(\lfloor G^{[n]}\bar{X}_1^{[n]}\rfloor + \bar{X}_2^{[n]}|G^{[n]})}{\frac{n}{2}\log(P)} \tag{27}$$

Note that in (24) we bounded $H(\lfloor G^{[n]}\bar{X}_1^{[n]}\rfloor + \bar{X}_2^{[n]}|G^{[n]})$ as follows.

$$H(\lfloor G^{[n]}\bar{X}_1^{[n]}\rfloor + \bar{X}_2^{[n]}|G^{[n]})$$

$$\leq \sum_{t=1}^{n} H(\lfloor G(t)\bar{X}_1(t)\rfloor + \bar{X}_2(t)|G(t)) \qquad (28)$$

$$\leq \sum_{t=1}^{n} \mathrm{E}\log\left((|G(t)|+1)(\lceil\sqrt{P}\rceil+1)\right) \qquad (29)$$

$$= \frac{n}{2}\log(P) + n\, o(\log(P)) \qquad (30)$$

where (29) follows from the observation that for a given $G(t)$ value, $\lfloor G^{[n]}\bar{X}_1^{[n]}\rfloor + \bar{X}_2^{[n]}$ can take at most $(1 + G(t))(1+\lceil\sqrt{P}\rceil)$ values (all integers) and the entropy of a variable that can take finitely many values is at most the log of the number of values.

3) **Functional Dependence** $\bar{X}_2^{[n]}(\bar{X}_1^{[n]})$

Next we show that one can assume that $\bar{X}_2^{[n]}$ is a function of $\bar{X}_1^{[n]}$. Given the sets of codeword vectors $\{\bar{X}_1^{[n]}\}$, $\{\bar{X}_2^{[n]}\}$, define $\mathcal{L}$ as the mapping from $\bar{X}_1^{[n]}$ to $\bar{X}_2^{[n]}$, i.e.,

$$\bar{X}_2^{[n]} = \mathcal{L}(\bar{X}_1^{[n]}) \qquad (31)$$

In general, because the mapping may be random, $\mathcal{L}$ is a random variable. Because conditioning cannot increase entropy,

$$H\left(\lfloor G^{[n]}\bar{X}_1^{[n]}\rfloor + \mathcal{L}(\bar{X}_1^{[n]})|G^{[n]}\right)$$

$$\geq H\left(\lfloor G^{[n]}\bar{X}_1^{[n]}\rfloor + \mathcal{L}(\bar{X}_1^{[n]})\Big| G^{[n]}, \mathcal{L}\right)$$

$$\geq \min_{L\in\{\mathcal{L}\}} H\left(\lfloor G^{[n]}\bar{X}_1^{[n]}\rfloor + \mathcal{L}(\bar{X}_1^{[n]})\Big| G^{[n]}, \mathcal{L}=L\right)$$

$$\qquad (32)$$

Let $L_o \in \mathcal{L}$ be the mapping that minimizes the entropy term. Then, choosing

$$\bar{X}_2^{[n]}(\bar{X}_1^{[n]}) = L_o(\bar{X}_1^{[n]}) \qquad (33)$$

we have

$$\bar{\mathcal{D}}_\Delta \leq \hat{\mathcal{D}}_\Delta$$

$$\triangleq \limsup_{P\to\infty}\limsup_{n\to\infty} \max_{\substack{\mathbb{P}(\bar{X}_1^{[n]}),\bar{X}_2^{[n]}(\bar{X}_1^{[n]}) \\ (\bar{X}_1^{[n]},\bar{X}_2^{[n]})\in\bar{\mathcal{X}}^{[n]}}}$$

$$\times \frac{H(\bar{X}_1^{[n]}|G^{[n]}) - H(\lfloor G^{[n]}\bar{X}_1^{[n]}\rfloor + \bar{X}_2^{[n]}(\bar{X}_1^{[n]})|G^{[n]})}{\frac{n}{2}\log(P)}$$

$$\qquad (34)$$

because the choice of the mapping function does not affect the positive entropy term, and it minimizes the negative entropy term. Henceforth, because $\bar{X}_2^{[n]}$ is a function of $\bar{X}_1^{[n]}$, we will refer to codewords only through $\bar{X}_1^{[n]}$ values.

4) **Definition of Aligned Image Sets**

The aligned image set containing the codeword $\bar{v}^{[n]} \in \mathrm{supp}(\bar{X}_1^{[n]})$ for channel realization $G^{[n]}$ is defined

as the set of all codewords that cast the same image as $\bar{v}^{[n]}$ at user 2.

$$S_{\bar{v}^{[n]}}(G^{[n]})$$

$$\triangleq \{\bar{x}_1^{[n]} \in \mathrm{supp}(\bar{X}_1^{[n]}) : \lfloor G^{[n]}\bar{x}_1^{[n]}\rfloor + \bar{X}_2^{[n]}(\bar{x}_1^{[n]})$$

$$= \lfloor G^{[n]}\bar{v}^{[n]}\rfloor + \bar{X}_2^{[n]}(\bar{v}^{[n]})\} \qquad (35)$$

Since we will need the average (over $G^{[n]}$) of the cardinality of an aligned image set, $\mathrm{E}|S_{\bar{v}^{[n]}}(G^{[n]})|$, it is worthwhile to point out that the cardinality $|S_{\bar{v}^{[n]}}(G^{[n]})|$ as a function of $G^{[n]}$, is a bounded simple function, and therefore measurable.[2] It is bounded because its values are restricted to natural numbers not greater than $(1 + \lceil\sqrt{P}\rceil)^{2n}$. To verify that it is a simple function, it suffices to show that the sets where some specific codewords like $v_1^{[n]}$ and $v_2^{[n]}$ align, are measurable sets.

$$S_{12} = \{G^{[n]} : \lfloor G^{[n]}\bar{v}_1^{[n]}\rfloor + \bar{X}_2^{[n]}(\bar{v}_1^{[n]})$$

$$= \lfloor G^{[n]}\bar{v}_2^{[n]}\rfloor + \bar{X}_2^{[n]}(\bar{v}_2^{[n]})\} \qquad (36)$$

$$= \bigcup_{A^{[n]}\in\mathbb{Z}^{[n]}} \{\{G^{[n]} : \lfloor G^{[n]}\bar{v}_1^{[n]}\rfloor = A^{[n]}\}$$

$$\bigcap \{G^{[n]} : \lfloor G^{[n]}\bar{v}_2^{[n]}\rfloor = B^{[n]}\}\} \qquad (37)$$

where $B^{[n]} = A^{[n]} + \bar{X}_2^{[n]}(\bar{v}_1^{[n]}) - \bar{X}_2^{[n]}(\bar{v}_2^{[n]})$. So, S_{12} is a countable union of intersections of open intervals and closed intervals in $\mathbb{R}^{[n]}$, which makes it a measurable set.[3] Thus, $|S_{\bar{v}^{[n]}}(G^{[n]})|$ is a simple function.

5) **Bounding Difference of Entropies, $\hat{\mathcal{D}}_\Delta$, in Terms of Size of Aligned Image Sets**

$$H(\bar{X}_1^{[n]}|G^{[n]}) = H(\bar{X}_1^{[n]}, S_{\bar{X}_1^{[n]}}(G^{[n]})|G^{[n]})$$

$$= H(S_{\bar{X}_1^{[n]}}(G^{[n]})|G^{[n]})$$

$$\quad + H(\bar{X}_1^{[n]}|S_{\bar{X}_1^{[n]}}(G^{[n]}), G^{[n]})$$

$$= H(\lfloor G^{[n]}\bar{X}_1^{[n]}\rfloor + \bar{X}_2^{[n]}(\bar{X}_1^{[n]})|G^{[n]})$$

$$\quad + H(\bar{X}_1^{[n]}|S_{\bar{X}_1^{[n]}}(G^{[n]}), G^{[n]})$$

$$\leq H(\lfloor G^{[n]}\bar{X}_1^{[n]}\rfloor + \bar{X}_2^{[n]}(\bar{X}_1^{[n]})|G^{[n]})$$

$$\quad + \mathrm{E}\left[\log(|S_{\bar{X}_1^{[n]}}(G^{[n]})|)\right] \qquad (38)$$

$$\leq H(\lfloor G^{[n]}\bar{X}_1^{[n]}\rfloor + \bar{X}_2^{[n]}(\bar{X}_1^{[n]})|G^{[n]})$$

$$\quad + \log\left(\mathrm{E}\left[|S_{\bar{X}_1^{[n]}}(G^{[n]})|\right]\right) \qquad (39)$$

where (38) follows because uniform distribution maximizes entropy, and (39) follows from Jensen's inequality. Rearranging terms, we note that

$$\hat{\mathcal{D}}_\Delta \leq \limsup_{P\to\infty}\limsup_{n\to\infty} \max_{\substack{\mathbb{P}(\bar{X}_1^{[n]}),\bar{X}_2^{[n]}(\bar{X}_1^{[n]}) \\ (\bar{X}_1^{[n]},\bar{X}_2^{[n]})\in\bar{\mathcal{X}}^{[n]}}}$$

$$\times \frac{\log\left(\mathrm{E}\left[|S_{\bar{X}_1^{[n]}}(G^{[n]})|\right]\right)}{\frac{n}{2}\log(P)} \qquad (40)$$

---

[2] A simple function is a finite sum of indicator functions of measurable sets [24].

[3] Note that $\lfloor Gv \rfloor = m$, corresponds to $G \in [m/v, (m+1)/v)$, which is the intersection of an open interval and a closed interval. Also recall that a closed or open subset of $\mathbb{R}^{[n]}$ is Lebesgue measurable, and countable unions and intersections of Lebesgue measurable sets are Lebesgue measurable.

6) **Bounding the Probability of Image Alignment**

Given two codewords $\bar{x}_1^{[n]}$ and $\bar{v}^{[n]}$, let us bound the probability that their images align at user 2. Note that for $\bar{x}_1^{[n]} \in S_{\bar{v}^{[n]}}(G^{[n]})$ we must have

$$\lfloor G^{[n]}\bar{x}_1^{[n]}\rfloor - \lfloor G^{[n]}\bar{v}^{[n]}\rfloor = \bar{X}_2^{[n]}(\bar{v}^{[n]}) - \bar{X}_2^{[n]}(\bar{x}_1^{[n]})$$
$$\Rightarrow G^{[n]}(\bar{x}_1^{[n]}-\bar{v}^{[n]}) \in \bar{X}_2^{[n]}(\bar{x}_1^{[n]})$$
$$-\bar{X}_2^{[n]}(\bar{v}^{[n]}) + \Delta_{(-1,1)}^{[n]} \quad (41)$$

where $\Delta_{(-1,1)}(t) \in (-1,1), \forall t \in [1:n]$. Thus, for all $t \in [1:n]$ such that $\bar{x}_1(t) \neq \bar{v}(t)$, the value of $G(t)$ must lie within an interval of length no more than $\frac{2}{|\bar{x}_1(t)-\bar{v}(t)|}$. Since the maximum value of the joint probability density function of $\{G(t) : \text{ such that } \bar{x}_1(t) \neq \bar{v}(t), t \in [1 : n]\}$ is bounded by $f_{\max}^{\sum_{t=1}^n 1(\bar{x}_1(t)\neq\bar{v}(t))} \leq f_{\max}^{[n]}$, we can bound the probability that the images of two codewords align as follows.

$$\mathbb{P}(\bar{x}_1^{[n]} \in S_{\bar{v}^{[n]}}(G^{[n]}))$$
$$\leq f_{\max}^n \prod_{t:\bar{x}_1(t)\neq\bar{v}(t)} \frac{2}{|\bar{x}_1(t)-\bar{v}(t)|} \quad (42)$$

7) **Bounding the Average Size of Aligned Image Sets**

$$\mathrm{E}\left[\left|S_{\bar{v}^{[n]}}(G^{[n]})\right|\right]$$
$$= \sum_{\bar{x}_1^{[n]}\in\{\bar{X}_1^{[n]}\}} \mathbb{P}\left(\bar{x}_1^{[n]} \in S_{\bar{v}^{[n]}}(G^{[n]})\right) \quad (43)$$
$$= 1 + \sum_{\substack{\bar{x}_1^{[n]}\in\{\bar{X}_1^{[n]}\}\\ \bar{x}_1^{[n]}\neq\bar{v}^{[n]}}} \mathbb{P}\left(\bar{x}_1^{[n]} \in S_{\bar{v}^{[n]}}(G^{[n]})\right) \quad (44)$$
$$\leq 1 + (2f_{\max})^n$$
$$\times \sum_{\substack{\bar{x}_1^{[n]}\in\{\bar{X}_1^{[n]}\}\\ \bar{x}_1^{[n]}\neq\bar{v}^{[n]}}} \prod_{t:\bar{x}_1(t)\neq\bar{v}(t)} \frac{1}{|\bar{x}_1(t)-\bar{v}(t)|} \quad (45)$$
$$\leq 1 + (2f_{\max})^n \prod_{t=1}^n \left(1+\sum_{\Delta_{\bar{x}}=1}^{\lceil\sqrt{P}\rceil}\frac{2}{\Delta_{\bar{x}}}\right) \quad (46)$$
$$\leq 1 + (2f_{\max})^n \prod_{t=1}^n \left(2\log(\sqrt{P})+3\right) \quad (47)$$

Since this is true for all $\bar{v}^{[n]} \in \{\bar{X}_1^{[n]}\}$

$$\mathrm{E}\left[\left|S_{\bar{X}_1^{[n]}}(G^{[n]})\right|\right]$$
$$\leq 1 + (2f_{\max})^n \left(2\log(\sqrt{P})+3\right)^n \quad (48)$$

8) **Combining the Bounds to Complete the Proof**

Combining (40) and (48) we have

$$\hat{\mathcal{D}}_\Delta \leq \limsup_{P\to\infty}\limsup_{n\to\infty}$$
$$\times \frac{\log\left(1+(2f_{\max})^n\left(2\log(\sqrt{P})+3\right)^n\right)}{\frac{n}{2}\log(P)}$$
$$= \limsup_{P\to\infty}\frac{\log(f_{\max})}{\frac{1}{2}\log(P)} + \limsup_{P\to\infty}\frac{\log(\log(P))}{\frac{1}{2}\log(P)}$$
$$\leq \alpha \quad (49)$$

where (49) follows because $f_{\max} = O(P^{\frac{\alpha}{2}})$. Finally combining (49) with (26) and (34) we have the desired outer bound

$$\mathcal{D}_\Sigma \leq 1 + \alpha \quad (50)$$

$\square$

VI. PROOF OF THEOREM 1 FOR K USERS

The generalization of the proof to the K user setting is, for the most part, straightforward based on the 2 user case studied earlier. To avoid repetition our presentation will only briefly summarize the aspects that follow directly and use detailed exposition for only those aspects that require special attention. We divide the proof into a similar set of steps for ease of reference with the 2 user case.

1) **Deterministic Channel Model**

As in the 2 user case, the deterministic channel model is described as:

$$\bar{Y}_k = \sum_{i=1}^{k-1}\lfloor G_{ki}(t)\bar{X}_i(t)\rfloor + \bar{X}_k(t) \quad (51)$$

where the integer inputs satisfy the following per-symbol power constraint

$$\bar{X}_k(t) \in \{0, 1, \cdots, \lceil\sqrt{P}\rceil\}, \quad \forall k \in [1:K] \quad (52)$$

As before, let us define $\bar{\mathcal{X}}^{[n]}$ as the set of codewords that satisfy the power constraint. We have the following bound.

Lemma 2: The DoF of the canonical model are bounded above by the DoF of the deterministic model. We omit the proof of Lemma 2 since it is a straightforward extension of the 2 user proof which was already presented in much detail.

2) **Difference of Entropy Terms**

For the kth user we bound the rate as

$$nR_k \leq I(W_k; \bar{Y}_k^{[n]}|G^n, W_{k+1}, W_{k+2}, \cdots, W_K) + o(n)$$
$$\leq H(\bar{Y}_k^{[n]}|G^n, W_{k+1}, \cdots, W_K)$$
$$-H(\bar{Y}_k^{[n]}|G^n, W_k, W_{k+1}, \cdots, W_K) + o(n) \quad (53)$$

where G^n includes all channel realizations. Adding the rate bounds we obtain

$$n\sum_{k=1}^K R_k \leq \frac{n}{2}\log(P) + \sum_{k=2}^K \Big(H(\bar{Y}_{k-1}^{[n]}|G^n, W_k, \cdots, W_K)$$
$$-H(\bar{Y}_k^{[n]}|G^n, W_k, \cdots, W_K)\Big)$$
$$+n\, o(\log(P)) + o(n) \quad (54)$$
$$\leq \frac{n}{2}\log(P) + \sum_{k=2}^K \Big[\max_{w_i\in\{W_i\}, i\in[k:K]}$$
$$\times \Big(H(\bar{Y}_{k-1}^{[n]}|G^n, W_i = w_i, \forall i \in [k:K])$$
$$-H(\bar{Y}_k^{[n]}|G^n, W_i = w_i, \forall i \in [k:K])\Big)\Big]$$
$$+n\, o(\log(P)) + o(n) \quad (55)$$

In DoF terms,

$$\bar{\mathcal{D}}_\Sigma \leq 1 + \sum_{k \in [2:K]} \bar{\mathcal{D}}_{\Delta,k} \tag{56}$$

So we need to bound each of the following difference of entropy terms, $\forall k \in [2:K]$

$$\bar{\mathcal{D}}_{\Delta,k} \triangleq \limsup_{P\to\infty} \limsup_{n\to\infty} \max_{\substack{\mathbb{P}(\bar{X}_1^{[n]},\cdots,\bar{X}_K^{[n]}) \\ (\bar{X}_1^{[n]},\cdots,\bar{X}_K^{[n]}) \in \bar{\mathcal{X}}^{[n]}}}$$
$$\times \frac{H(\bar{Y}_{k-1}^{[n]}|G^{[n]}) - H(\bar{Y}_k^{[n]}|G^{[n]})}{\frac{n}{2}\log(P)} \tag{57}$$

We will bound these terms one at a time. The remainder of the proof will show that $\bar{\mathcal{D}}_{\Delta,k} \leq \alpha_k$.

3) **Functional Dependence** $\bar{Y}_k^{[n]}(\bar{Y}_{k-1}^{[n]}, G^{[n]})$
For a given channel realization for user $k-1$, $G_{k-1}^{[n]}$, there are multiple vectors $(\bar{X}_1^{[n]}, \bar{X}_2^{[n]}, \cdots, \bar{X}_k^{[n]})$ that cast the same image in $\bar{Y}_{k-1}^{[n]}$. Thus, given the channel for user $k-1$, the mapping from $\bar{Y}_{k-1}^{[n]}$ to one of these vectors $(\bar{X}_1^{[n]}, \bar{X}_2^{[n]}, \cdots, \bar{X}_k^{[n]})$ is random. Let us denote it by $\mathcal{L}$, i.e.,

$$(\bar{X}_1^{[n]}, \bar{X}_2^{[n]}, \cdots, \bar{X}_k^{[n]}) = \mathcal{L}(\bar{Y}_{k-1}^{[n]}, G_{k-1}^{[n]}) \tag{58}$$

Now note that

$$H(\bar{Y}_k^{[n]}|G^n) \geq H(\bar{Y}_k^{[n]}|G^n, \mathcal{L}) \tag{59}$$
$$\geq \min_{L \in \{\mathcal{L}\}} H(\bar{Y}_k^{[n]}|G^n, \mathcal{L}=L) \tag{60}$$

Let a minimizing mapping be L_o. Fix this as the deterministic mapping,

$$(\bar{X}_1^{[n]}, \bar{X}_2^{[n]}, \cdots, \bar{X}_k^{[n]}) = L_o(\bar{Y}_{k-1}^{[n]}, G_{k-1}^{[n]}) \tag{61}$$

This implicitly allows the transmitter to have full knowledge of the channel vector of user $k-1$. We note that the choice of mapping does not affect the positive entropy term $H(\bar{Y}_{k-1}^{[n]}|G^{[n]})$ but it minimizes $H(\bar{Y}_k^{[n]}|G^{[n]})$, so that we can bound $\bar{\mathcal{D}}_{\Delta,k}$ as follows.

$$\bar{\mathcal{D}}_{\Delta,k} \leq \hat{\mathcal{D}}_{\Delta,k}$$
$$\triangleq \limsup_{P\to\infty} \limsup_{n\to\infty} \max_{\substack{\mathbb{P}(\bar{Y}_{k-1}^{[n]}|G^{[n]}),\bar{Y}_k^{[n]}(\bar{Y}_{k-1}^{[n]},G^{[n]}) \\ [3pt](\bar{X}_1^{[n]},\cdots,\bar{X}_K^{[n]}) \in \bar{\mathcal{X}}^{[n]}}}$$
$$\times \frac{H(\bar{Y}_{k-1}^{[n]}|G^{[n]}) - H(\bar{Y}_k^{[n]}|G^{[n]})}{\frac{n}{2}\log(P)} \tag{62}$$

Henceforth, note that $\bar{Y}_k^{[n]}$ is a function of $\bar{Y}_{k-1}^{[n]}, G^{[n]}$.

4) **Define Aligned Image Sets**
For channel realization $G^{[n]}$, define the aligned image set $S_{\bar{Y}_{k-1}^{[n]}}(G^{[n]})$ as the set of all $\bar{Y}_{k-1}^{[n]}$ that have the same image in $\bar{Y}_k^{[n]}$, i.e.,

$$S_{\bar{v}^{[n]}}(G^{[n]}) \triangleq \{\bar{y}_{k-1}^{[n]} \in \{\bar{Y}_{k-1}^{[n]}\} : \bar{Y}_k^{[n]}(v^{[n]}, G^{[n]})$$
$$= \bar{Y}_k^{[n]}(\bar{y}_{k-1}^{[n]}, G^{[n]})\} \tag{63}$$

5) **Bounding Difference of Entropies in Terms of Size of Aligned Image Sets**

$$H(\bar{Y}_{k-1}^{[n]}|G^{[n]})$$
$$= H(\bar{Y}_{k-1}^{[n]}, \bar{Y}_k^{[n]}|G^{[n]}) \tag{64}$$
$$= H(\bar{Y}_k^{[n]}|G^{[n]}) + H(\bar{Y}_{k-1}^{[n]}|\bar{Y}_k^{[n]}, G^{[n]}) \tag{65}$$
$$= H(\bar{Y}_k^{[n]}|G^{[n]}) + H(S_{\bar{Y}_{k-1}^{[n]}}(G^{[n]})|G^{[n]}) \tag{66}$$
$$\leq H(\bar{Y}_k^{[n]}|G^{[n]}) + \mathrm{E}\left[\log\left|S_{\bar{Y}_{k-1}^{[n]}}(G^{[n]})\right|\right] \tag{67}$$
$$\leq H(\bar{Y}_k^{[n]}|G^{[n]}) + \log\mathrm{E}\left[\left|S_{\bar{Y}_{k-1}^{[n]}}(G^{[n]})\right|\right] \tag{68}$$

where we used Jensen's inequality in (68). Rearranging terms, we note that

$$\hat{\mathcal{D}}_{\Delta,k} \leq \limsup_{P\to\infty} \limsup_{n\to\infty} \max_{\substack{\mathbb{P}(\bar{Y}_{k-1}^{[n]},G^{[n]}),\bar{Y}_k^{[n]}(\bar{Y}_{k-1}^{[n]},G^{[n]}) \\ (\bar{X}_1^{[n]},\cdots,\bar{X}_K^{[n]}) \in \bar{\mathcal{X}}^{[n]}}}$$
$$\times \frac{\log\left(\mathrm{E}\left[\left|S_{\bar{Y}_{k-1}^{[n]}}(G^{[n]})\right|\right]\right)}{\frac{n}{2}\log(P)} \tag{69}$$

6) **Bounding the Probability of Image Alignment**
Given the channel, $G_{k-1}^{[n]}$, of user $k-1$ and two realizations of $\bar{Y}_{k-1}^{[n]}$, say $\bar{y}^{[n]}$ and $\bar{y}'^{[n]}$, which map to $\bar{X}_j^{[n]}(\bar{y}^{[n]}, G_{k-1}^{[n]}) = \bar{x}_j^{[n]}$ and $\bar{X}_j^{[n]}(\bar{y}'^{[n]}, G_{k-1}^{[n]}) = \bar{x}'_j^{[n]}$, $\forall j \in [1:k]$, let us bound the probability that they produce the same image $\bar{Y}_k^{[n]}$. For notational compactness let us define $G_{kk}(t) = 1, \forall k \in [1:K], \forall t \in [1:n]$. Note that for $\bar{y}'^{[n]} \in S_{\bar{y}^{[n]}}(G^{[n]})$ we must have, $\forall t \in [1:n]$

$$\sum_{j \in [1:k]} \lfloor G_{kj}(t)\bar{x}'_j(t)\rfloor = \sum_{j \in [1:k]} \lfloor G_{kj}(t)\bar{x}_j(t)\rfloor$$

So we have,

$$\lfloor G_{kj^*(t)}(t)\bar{x}'_{j^*(t)}(t)\rfloor - \lfloor G_{kj^*(t)}(t)\bar{x}_{j^*(t)}(t)\rfloor$$
$$= \sum_{j \in [k:k], j \neq j^*(t)} \left(\lfloor G_{kj}(t)\bar{x}_j(t)\rfloor - \lfloor G_{kj}(t)\bar{x}'_j(t)\rfloor\right)$$
$$\Rightarrow G_{kj^*(t)}(t)\left(\bar{x}'_{j^*(t)}(t) - \bar{x}_{j^*(t)}(t)\right)$$
$$\in \sum_{j \in [1:k], j \neq j^*(t)} \left(\lfloor G_{kj}(t)\bar{x}_j(t)\rfloor - \lfloor G_{kj}(t)\bar{x}'_j(t)\rfloor\right)$$
$$+ \Delta_{(-1,1)} \tag{70}$$

where $\Delta_{(-1,1)} \in (-1,1)$, and we define

$$j^*(t) \triangleq \arg\max_{j \in [1:k-1]} |\bar{x}'_j(t) - \bar{x}_j(t)| \tag{71}$$

Thus, for all $t \in [1:n]$ such that $\bar{x}'_{j^*(t)}(t) \neq \bar{x}_{j^*(t)}(t)$, the value of $G_{kj^*(t)}(t)$ must lie within an interval of length no more than $\frac{2}{|\bar{x}'_{j^*(t)}-\bar{x}_{j^*(t)}|}$. Therefore, the probability that the images due to $\bar{y}^{[n]}$ and $\bar{y}'^{[n]}$ align at user k, is bounded as follows.

$$\mathbb{P}\left(\bar{y}'^{[n]} \in S_{\bar{y}^{[n]}}(G^{[n]})\right) \leq f_{\max,k}^n \prod_{t:\bar{x}'_{j^*(t)}(t) \neq \bar{x}_{j^*(t)}(t)}$$
$$\times \frac{2}{|\bar{x}'_{j^*(t)}(t) - \bar{x}_{j^*(t)}(t)|}$$

It will be useful to express the bound in terms of $\bar{y}'(t), \bar{y}(t)$. To this end, let us proceed as follows.

$$\bar{y}'(t) - \bar{y}(t)$$

$$= \sum_{j=1}^{k-1} \left(\lfloor G_{k-1,j}(t)\bar{x}'_j(t) \rfloor - \lfloor G_{k-1,j}(t)\bar{x}_j(t) \rfloor \right)$$

$$\in \sum_{j=1}^{k-1} \left(\lfloor G_{k-1,j}(t) \left(\bar{x}'_j(t) - \bar{x}_j(t) \right) \rfloor \right) + (-K, K)$$

$$|\bar{y}'(t) - \bar{y}(t)|$$

$$\leq |\bar{x}'_{j^*(t)}(t) - \bar{x}_{j^*(t)}(t)| \sum_{j=1}^{k-1} |G_{k-1,j}(t)| + K$$

$$\Rightarrow \frac{1}{|\bar{x}'_{j^*(t)}(t) - \bar{x}_{j^*(t)}(t)|} \leq \frac{\sum_{j=1}^{k-1} |G_{k-1,j}(t)|}{|\bar{y}'(t) - \bar{y}(t)| - K}$$

whenever $|\bar{y}'(t) - \bar{y}(t)| > K$. Therefore,

$$\mathbb{P}\left(\bar{y}'^{[n]} \in S_{\bar{y}^{[n]}}(G^{[n]}) \right)$$

$$\leq \bar{g}^n (f_{\max,k})^n \prod_{t:|\bar{y}'(t)-\bar{y}(t)|>K} \frac{1}{|\bar{y}'(t) - \bar{y}(t)| - K}$$

where

$$\bar{g}^n \triangleq \max\left(1, \prod_{t:\bar{x}'_{j^*(t)}(t) \neq \bar{x}_{j^*(t)}(t)} 2\sum_{j=1}^{k-1} |G_{k-1,j}(t)| \right)$$

7) **Bounding the Average Size of Aligned Image Sets**

$$\mathrm{E}\left[S_{\bar{y}}^{[n]}(G^n) \right]$$

$$= \sum_{\bar{y}'^{[n]} \in \{\bar{Y}_{k-1}^{[n]}\}} \mathbb{P}\left(\bar{y}'^{[n]} \in S_{\bar{y}^{[n]}}(G^n) \right) \tag{72}$$

$$\leq \bar{g}^n (f_{\max,k})^n \prod_{t=1}^{n} \left(\sum_{\bar{y}'(t):|\bar{y}'(t)-\bar{y}(t)| \leq K} 1 \right.$$

$$\left. + \sum_{\bar{y}'(t):K < |\bar{y}'(t)-\bar{y}(t)| \leq Q_y(t)} \frac{1}{|\bar{y}'(t) - \bar{y}(t)| - K} \right) \tag{73}$$

$$\leq \bar{g}^n (f_{\max,k})^n \left(\log(\sqrt{P}) + o(\log(P)) \right)^n \tag{74}$$

where $Q_y(t) \leq \lceil \sqrt{P} \rceil \sum_{j \in [1:k-1]} (|G_{k-1,j}(t)| + K)$.

8) **Combining the Bounds to Complete the Proof**
Combining (74) and (69) we have

$$\hat{\mathcal{D}}_{\Delta,k} \leq \limsup_{P \to \infty} \limsup_{n \to \infty}$$

$$\times \frac{\log\left((\bar{g} f_{\max,k})^n (\log(\sqrt{P}) + o(\log(P)))^n \right)}{\frac{n}{2} \log(P)}$$

$$\leq \alpha_k \tag{75}$$

Finally, combining (56), (62) and (75) we have the result,

$$\bar{\mathcal{D}}_\Sigma \leq 1 + \alpha_2 + \cdots + \alpha_K \tag{76}$$

VII. EXTENSION TO COMPLEX CHANNELS

The channel model for the complex setting is the identical to the real setting described in Section II-E, except that all symbols are complex and instead of (6), the DoF are defined as

$$d_1 = \lim_{P \to \infty} \frac{R_1(P)}{\log(P)}, \quad d_2 = \lim_{P \to \infty} \frac{R_2(P)}{\log(P)}$$

The deterministic channel model is described similar to the real setting as;

$$\bar{Y}_k = \sum_{i=1}^{k-1} \lfloor G_{ki}(t)\bar{X}_i(t) \rfloor + \bar{X}_k(t) \tag{77}$$

where the real and imaginary parts of the inputs, i.e. $\bar{X}_{kR}(t)$ and $\bar{X}_{kI}(t)$ are integers and satisfy the following per-symbol power constraint

$$\bar{X}_{kR}(t) \in \{0, 1, \cdots, \lceil \sqrt{P} \rceil\}, \quad \forall k \in [1:K] \tag{78}$$

$$\bar{X}_{kI}(t) \in \{0, 1, \cdots, \lceil \sqrt{P} \rceil\}, \quad \forall k \in [1:K] \tag{79}$$

Similar to the real setting, $\forall k \in [2:K]$, define the set of channel coefficient variables

$$\{G_k^{[n]}\} = \{G_{ki,R}(t) : i \in [1:k-1], t \in [n]\}$$

$$\cup \{G_{ki,I}(t) : i \in [1:k-1], t \in [n]\} \tag{80}$$

where $G_{ki,R}(t), G_{ki,I}(t)$ are the real and imaginary parts of $G_{ki}(t)$, respectively. Define $\{G^{[n]}\} = \bigcup_{k \in [1:K]} G_k^{[n]}$. Similar to the real setting, $\forall n \in \mathbb{N}$, and

$$\forall \{G_1\} \subset \{G^{[n]}\}, \quad \{G_2\} \subset \{G^{[n]}\},$$

$$\{G_1\} \cap \{G_2\} = \varnothing, \quad |\{G_1\}| < \infty, |\{G_2\}| < \infty$$

the conditional PDF

$$f_{G_1|\mathcal{T},G_2}(g_1|\tau, g_2) \leq f_{\max}^{|\{G_1\}|}.$$

Furthermore, similar to the K user real setting where we allow $f_{\max,k}(P)$ to scale as $O((\sqrt{P})^{\alpha_k})$ for some $\alpha_k \in [0, 1]$, here also we allow the density functions of each user to scale at different rates, representing different amounts of CSIT feedback, so that for the unknown channel coefficients associated with user k we have peak density constraint, $f_{\max,k}(P) = O(\sqrt{P^{\alpha_k}})$.

The generalization of the proof to the complex channel coefficients setting is, for the most part, straightforward based on the K user case studied earlier. To avoid repetition, here we focus only on the differences.

In DoF terms,

$$\bar{\mathcal{D}}_\Sigma \leq 1 + \sum_{k \in [2:K]} \bar{\mathcal{D}}_{\Delta,k} \tag{81}$$

We need to bound each of the following difference of entropy terms, $\forall k \in [2:K]$

$$\bar{\mathcal{D}}_{\Delta,k} \triangleq \limsup_{P \to \infty} \limsup_{n \to \infty} \max_{\substack{\mathbb{P}(\bar{X}_1^{[n]}, \cdots, \bar{X}_K^{[n]}) \\ (\bar{X}_1^{[n]}, \cdots, \bar{X}_K^{[n]}) \in \bar{\mathcal{X}}^{[n]}}}$$

$$\times \frac{H(\bar{Y}_{k-1}^{[n]}|G^{[n]}) - H(\bar{Y}_k^{[n]}|G^{[n]})}{n \log(P)} \tag{82}$$

We will bound these terms one at a time. The remainder of the proof will show that $\bar{\mathcal{D}}_{\Delta,k} \leq \alpha_k$. Note that, the functional dependence part which claims that $\bar{Y}_k^{[n]}$ is a function of $\bar{Y}_{k-1}^{[n]}$, $G^{[n]}$, aligned image sets and bounding difference of entropies in terms of size of aligned image sets are also the same as the real setting.

$$\hat{\mathcal{D}}_{\Delta,k} \leq \limsup_{P \to \infty} \limsup_{n \to \infty} \max_{\substack{\mathbb{P}(\bar{Y}_{k-1}^{[n]}, G^{[n]}), \bar{Y}_k^{[n]}(\bar{Y}_{k-1}^{[n]}, G^{[n]}) \\ (\bar{X}_1^{[n]}, \cdots, \bar{X}_K^{[n]}) \in \bar{\mathcal{X}}^{[n]}}}$$

$$\times \frac{\log\left(\mathrm{E}\left[\left|S_{\bar{Y}_{k-1}^{[n]}}(G^{[n]})\right|\right]\right)}{n \log(P)} \tag{83}$$

Now we bound the probability of alignment of images. For notational compactness let us define $G_{kk}(t) = 1, \forall k \in [1:K], \forall t \in [1:n]$. Define

$$j^*(t) \triangleq \arg\max_{j \in [1:k-1]} \max(|\bar{x}'_{j(t),R} - \bar{x}_{j(t),R}|,$$

$$|\bar{x}'_{j(t),I} - \bar{x}_{j(t),I}|). \tag{84}$$

Without loss of generality assume, $|\bar{x}'_{j^*(t),R} - \bar{x}_{j^*(t),R}| \geq |\bar{x}'_{j^*(t),I} - \bar{x}_{j^*(t),I}|$. Note that for $\bar{y}'^{[n]} \in S_{\bar{y}^{[n]}}(G^{[n]})$ we must have, $\sum_{j \in [1:k]} \lfloor G_{kj}(t)\bar{x}_j(t) \rfloor = \sum_{j \in [1:k]} \lfloor G_{kj}(t)\bar{x}'_j(t) \rfloor$ for all $t \in [1:n]$.

$$\Re\left\{ \sum_{j \in [1:k]} \lfloor G_{kj}(t)\bar{x}_j(t) \rfloor - \sum_{j \in [1:k]} \lfloor G_{kj}(t)\bar{x}'_j(t) \rfloor \right\}$$

$$= \sum_{j \in [1:k]} \lfloor G_{kjR}(t)\bar{x}_{j(t),R} - G_{kjI}(t)\bar{x}_{j(t),I} \rfloor$$

$$- \sum_{j \in [1:k]} \lfloor G_{kjR}(t)\bar{x}'_{j(t),R} - G_{kjI}(t)\bar{x}'_{j(t),I} \rfloor \tag{85}$$

$$= \lfloor G_{kj^*R}(t)\bar{x}_{j^*(t),R} - G_{kj^*I}(t)\bar{x}_{j^*(t),I} \rfloor$$

$$- \lfloor G_{kj^*R}(t)\bar{x}'_{j^*(t),R} - G_{kj^*I}(t)\bar{x}'_{j^*(t),I} \rfloor$$

$$+ \sum_{j \in [1:k], j \neq j^*(t)} (\lfloor G_{kjR}(t)\bar{x}_{j(t),R} - G_{kjI}(t)\bar{x}_{j(t),I} \rfloor$$

$$- \lfloor G_{kjR}(t)\bar{x}'_{j(t),R} - G_{kjI}(t)\bar{x}'_{j(t),I} \rfloor)$$

$$= G_{kj^*R}(t)(\bar{x}_{j^*(t),R}$$

$$- \bar{x}'_{j^*(t),R}) - G_{kj^*I}(t)(\bar{x}_{j^*(t),I} - \bar{x}'_{j^*(t),I})$$

$$+ \sum_{j \in [1:k], j \neq j^*(t)} (\lfloor G_{kjR}(t)\bar{x}_{j(t),R} - G_{kjI}(t)\bar{x}_{j(t),I} \rfloor$$

$$- \lfloor G_{kjR}(t)\bar{x}'_{j(t),R} - G_{kjI}(t)\bar{x}'_{j(t),I} \rfloor) + \Delta_{(-1,1)} \tag{86}$$

where $\Delta_{(-1,1)} \in (-1, 1)$. Similarly we have,

$$\Im\left\{ \sum_{j \in [1:k]} \lfloor G_{kj}(t)\bar{x}_j(t) \rfloor - \sum_{j \in [1:k]} \lfloor G_{kj}(t)\bar{x}'_j(t) \rfloor \right\}$$

$$= \sum_{j \in [1:k]} \lfloor G_{kjR}(t)\bar{x}_{j(t),I} + G_{kjI}(t)\bar{x}_{j(t),R} \rfloor$$

$$- \sum_{j \in [1:k]} \lfloor G_{kjR}(t)\bar{x}'_{j(t),I} + G_{kjI}(t)\bar{x}'_{j(t),R} \rfloor$$

$$= \lfloor G_{kj^*R}(t)\bar{x}_{j^*(t),I} + G_{kj^*I}(t)\bar{x}_{j^*(t),R} \rfloor$$

$$- \lfloor G_{kj^*R}(t)\bar{x}'_{j^*(t),I} + G_{kj^*I}(t)\bar{x}'_{j^*(t),R} \rfloor$$

$$+ \sum_{j \in [1:k], j \neq j^*(t)} (\lfloor G_{kjR}(t)\bar{x}_{j(t),I} + G_{kjI}(t)\bar{x}_{j(t),R} \rfloor$$

$$- \lfloor G_{kjR}(t)\bar{x}'_{j(t),I} + G_{kjI}(t)\bar{x}'_{j(t),R} \rfloor)$$

$$= G_{kj^*R}(t)(\bar{x}_{j^*(t),I} - \bar{x}'_{j^*(t),I})$$

$$+ G_{kj^*I}(t)(\bar{x}_{j^*(t),R} - \bar{x}'_{j^*(t),R})$$

$$+ \sum_{j \in [1:k], j \neq j^*(t)} (\lfloor G_{kjR}(t)\bar{x}_{j(t),I} + G_{kjI}(t)\bar{x}_{j(t),R} \rfloor$$

$$- \lfloor G_{kjR}(t)\bar{x}'_{j(t),I} + G_{kjI}(t)\bar{x}'_{j(t),R} \rfloor) + \Delta_{(-1,1)} \tag{87}$$

Thus, for all $t \in [1:n]$ such that $\bar{x}'_{j^*(t),R}(t) \neq \bar{x}_{j^*(t),R}(t)$ the value of $G_{kj^*(t)R}(t)$ and $G_{kj^*(t)I}(t)$ must lie within an interval of length no more than $\frac{4}{\left|\bar{x}'_{j^*(t),R} - \bar{x}_{j^*(t),R}\right|}$. Therefore, the probability that the images due to $\bar{y}^{[n]}$ and $\bar{y}'^{[n]}$ align at user k, is bounded as follows.

$$\mathbb{P}\left(\bar{y}'^{[n]} \in S_{\bar{y}^{[n]}}(G^{[n]})\right)$$

$$\leq f_{\max,k}^{2n} \prod_{t : \bar{x}'_{j^*(t),R}(t) \neq \bar{x}_{j^*(t),R}(t)} \frac{16}{\left|\bar{x}'_{j^*(t),R} - \bar{x}_{j^*(t),R}\right|^2} \tag{88}$$

It will be useful to express the bound in terms of $\bar{y}'_R(t)$, $\bar{y}_R(t)$, $\bar{y}'_I(t)$ and, $\bar{y}_I(t)$. To this end, let us proceed as follows.

$$\bar{y}'_R(t) - \bar{y}_R(t) \tag{89}$$

$$= \sum_{j=1}^{k-1} \left(\lfloor G_{k-1,j,R}(t)\bar{x}'_{j(t),R} - G_{k-1,j,I}(t)\bar{x}'_{j(t),I} \rfloor\right.$$

$$\left. - \lfloor G_{k-1,j,R}(t)\bar{x}_{j(t),R} - G_{k-1,j,I}(t)\bar{x}_{j(t),I} \rfloor\right)$$

$$= \sum_{j=1}^{k-1} \left(G_{k-1,j,R}(t)\left(\bar{x}'_{j(t),R} - \bar{x}_{j(t),R}\right)\right.$$

$$\left. - G_{k-1,j,I}(t)\left(\bar{x}'_{j(t),I} - \bar{x}_{j(t),I}\right)\right) + \Delta_{(-k,k)} \tag{90}$$

$$\bar{y}'_I(t) - \bar{y}_I(t) \tag{91}$$

$$= \sum_{j=1}^{k-1} \left(\lfloor G_{k-1,j,R}(t)\bar{x}'_{j(t),I} + G_{k-1,j,I}(t)\bar{x}'_{j(t),R} \rfloor\right.$$

$$\left. - \lfloor G_{k-1,j,R}(t)\bar{x}_{j(t),I} + G_{k-1,j,I}(t)\bar{x}_{j(t),R} \rfloor\right)$$

$$= \sum_{j=1}^{k-1} \left(G_{k-1,j,R}(t)\left(\bar{x}'_{j(t),I} - \bar{x}_{j(t),I}\right)\right.$$

$$\left. + G_{k-1,j,I}(t)\left(\bar{x}'_{j(t),R} - \bar{x}_{j(t),R}\right)\right) + \Delta_{(-k,k)} \tag{92}$$

So,

$$\max(|\bar{y}'_R(t) - \bar{y}_R(t)|, |\bar{y}'_I(t) - \bar{y}_I(t)|)$$

$$\leq |\bar{x}'_{j^*(t),R}(t) - \bar{x}_{j^*(t),R}(t)|$$

$$\times \left\{|\sum_{j=1}^{k-1} |G_{k-1,j,R}(t)| + \sum_{j=1}^{k-1} |G_{k-1,j,I}(t)|\right\} + K$$

$$\Rightarrow \frac{1}{|\bar{x}'_{j^*(t),R}(t) - \bar{x}_{j^*(t),R}(t)|^2}$$

$$\leq \frac{\left\{|\sum_{j=1}^{k-1} |G_{k-1,j,R}(t)| + \sum_{j=1}^{k-1} |G_{k-1,j,I}(t)|\right\}^2}{\left\{|\bar{y}'_R(t) - \bar{y}_R(t)| - K\right\} \left\{|\bar{y}'_I(t) - \bar{y}_I(t)| - K\right\}}$$

whenever $\min(|\bar{y}'_R(t) - \bar{y}_R(t)|, |\bar{y}'_I(t) - \bar{y}_I(t)|) > K$. Therefore,

$$\mathbb{P}\left(\bar{y}'^{[n]} \in S_{\bar{y}^{[n]}}(G^{[n]})\right)$$

$$\leq \bar{g}^{2n}(f_{\max,k})^{2n} \prod_{t:|\bar{y}'(t)-\bar{y}(t)|>K}$$

$$\times \frac{1}{\{|\bar{y}'_R(t) - \bar{y}_R(t)| - K\}\{|\bar{y}'_I(t) - \bar{y}_I(t)| - K\}}$$

where

$$\bar{g}^{2n} \triangleq \max\left(1, \prod_{t:\bar{x}'_{j^*(t),R}(t)\neq\bar{x}_{j^*(t),R}(t)} 16\left\{|\sum_{j=1}^{k-1}|G_{k-1,j,R}(t)| + \sum_{j=1}^{k-1}|G_{k-1,j,I}(t)|\right\}^2\right)$$

$$\mathrm{E}\left[S_{\bar{y}}^{[n]}(G^n)\right] \tag{93}$$

$$= \sum_{\bar{y}'^{[n]}\in\{\bar{Y}_{k-1}^{[n]}\}} \mathbb{P}\left(\bar{y}'^{[n]} \in S_{\bar{y}^{[n]}}(G^n)\right) \tag{94}$$

$$\leq \bar{g}^{2n}(f_{\max,k})^{2n} \prod_{t=1}^{n}\left(\sum_{\bar{y}'_R(t):|\bar{y}'_R(t)-\bar{y}_R(t)|\leq K} 1 \right.$$

$$+ \sum_{\bar{y}'_R(t):K<|\bar{y}'_R(t)-\bar{y}_R(t)|\leq Q_{y,R}(t)}$$

$$\left.\times \frac{1}{|\bar{y}'_R(t) - \bar{y}_R(t)| - K}\right)$$

$$\times \prod_{t=1}^{n}\left(\sum_{\bar{y}'_I(t):|\bar{y}'_I(t)-\bar{y}_I(t)|\leq K}\right.$$

$$+1\sum_{\bar{y}'_I(t):K<|\bar{y}'_I(t)-\bar{y}_I(t)|\leq Q_{y,I}(t)}$$

$$\left.\times \frac{1}{|\bar{y}'_I(t) - \bar{y}_I(t)| - K}\right)$$

$$\leq \bar{g}^{2n}(f_{\max,k})^{2n}\left(\log(\sqrt{P}) + o(\log(P))\right)^{2n} \tag{95}$$

where $\max(Q_{y,R}(t), Q_{y,I}(t)) \leq \lceil\sqrt{P}\rceil\sum_{j\in[1:k-1]}(|G_{k-1,j,R}(t)| + |G_{k-1,j,I}(t)| + K)$. Now, combining (95) and (83) we have

$$\hat{\mathcal{D}}_{\Delta,k} \leq \limsup_{P\to\infty}\limsup_{n\to\infty}$$

$$\times \frac{\log\left((\bar{g}f_{\max,k})^{2n}(\log(\sqrt{P}) + o(\log(P)))^{2n}\right)}{n\log(P)}$$

$$\leq \alpha_k \tag{96}$$

Finally, combining (81), (83) and (96) we have the result,

$$\bar{\mathcal{D}}_{\Sigma} \leq 1 + \alpha_2 + \cdots + \alpha_K \tag{97}$$

VIII. DISCUSSION

Since CSIT is almost never available with infinite precision, the collapse of DoF under finite precision channel uncertainty is a sobering result that stands in stark contrast against the tremendous DoF gains shown to be possible with perfect channel knowledge [18], [25]. However, as evident from the conjecture of Lapidoth, Shamai and Wigger, the pessimistic outcome is not unexpected. In terms of practical implications, just like the extremely positive DoF results, the extremely negative DoF results should be taken with a grain of salt. The collapse of DoF under finite precision CSIT is very much due to the asymptotic nature of the DoF metric, and may not be directly representative of finite SNR scenarios which are of primary concern in practice. From a technical perspective, the new outer bound technique offers hope for new insights through the studies of more general forms of CSIT, such as finite precision versions of delayed [26], mixed [15], [16], topological [27], blind [6] and alternating [10] CSIT.

APPENDIX

The proof is in two parts. First we prove that we can limit the inputs and outputs to integer values without reducing the DoF. Then, we will show that the long-term (per-codeword) power constraints can be replaced with short-term (per-symbol) power constraints without reducing the DoF. The proofs follow along the lines of similar proofs by Bresler and Tse in [22], are specialized to the broadcast setting, and fill in several details that are omitted in [22].

A. Integer Inputs and Outputs

Given codebooks with real codewords, $(X_1^{[n]}, X_2^{[n]}) \in \mathbb{R}^n \times \mathbb{R}^n$ for the canonical channel model, we show that the deterministic channel model with integer inputs $\lfloor X_1^{[n]}\rfloor, \lfloor X_2^{[n]}\rfloor$, and outputs

$$\bar{\bar{Y}}_1(t) \triangleq \lfloor X_1(t)\rfloor \tag{98}$$

$$\bar{\bar{Y}}_2(t) \triangleq \lfloor G(t)\lfloor X_1(t)\rfloor\rfloor + \lfloor X_2(t)\rfloor \tag{99}$$

achieves the same DoF. Thus, removing noise and limiting the inputs to integer values, as done in the deterministic model, does not reduce DoF relative to the original canonical channel model.

Define $E_1^{[n]} = X_1^{[n]} - \lfloor X_1^{[n]}\rfloor + Z_1^{[n]}$. Taking a similar approach to [22, Lemma 5] we have,

$$I(W_1; Y_1^{[n]}|G^{[n]})$$

$$= I(W_1; \bar{\bar{Y}}_1^{[n]} + E_1^{[n]}|G^{[n]}) \tag{100}$$

$$\leq I(W_1; \bar{\bar{Y}}_1^{[n]}, E_1^{[n]}|G^{[n]}) \tag{101}$$

$$\leq I(W_1; \bar{\bar{Y}}_1^{[n]}|G^{[n]}) + I(W_1; E_1^{[n]}|G^{[n]}, \bar{\bar{Y}}_1^{[n]}) \tag{102}$$

$$\leq I(W_1; \bar{\bar{Y}}_1^{[n]}|G^{[n]}) + h(E_1^{[n]}|G^{[n]})$$

$$-h(E_1^{[n]}|G^{[n]}, \bar{\bar{Y}}_1^{[n]}, X_1^{[n]}) \tag{103}$$

$$\leq I(W_1; \bar{\bar{Y}}_1^{[n]}|G^{[n]}) + h(E_1^{[n]}|G^{[n]}) - h(Z_1^{[n]}) \tag{104}$$

$$\leq I(W_1; \bar{\bar{Y}}_1^{[n]}|G^{[n]}) + \frac{n}{2}\log(2) \tag{105}$$

so that the difference between $I(W_1; Y_1^{[n]}|G^{[n]})$ and $I(W_1; \bar{\bar{Y}}_1^{[n]}|G^{[n]})$ approaches 0 when normalized by $\frac{n}{2}\log(P)$, as first n and then P is sent to infinity.

Similarly, by defining $E_2^{[n]} = G^n X_1^{[n]} - \lfloor G^n \lfloor X_1^{[n]} \rfloor \rfloor + X_2^{[n]} - \lfloor X_2^{[n]} \rfloor + Z_2^{[n]}$, we have,

$$I(W_2; Y_2^{[n]}|G^{[n]})$$
$$= I(W_2; \bar{\bar{Y}}_2^{[n]} + E_2^{[n]}|G^{[n]}) \tag{106}$$
$$\leq I(W_2; \bar{\bar{Y}}_2^{[n]}|G^{[n]}) + h(E_2^{[n]}|G^{[n]}) - h(Z_2^{[n]})$$
$$\leq I(W_2; \bar{\bar{Y}}_2^{[n]}|G^{[n]})$$
$$\quad + \sum_{t=1}^{n} \mathrm{E}_{G(t)} \left[\frac{1}{2} \log \left((G(t)+1)^2 + 1 \right) \right] \tag{107}$$

so that the difference between $I(W_2; Y_2^{[n]}|G^{[n]})$ and $I(W_2; \bar{\bar{Y}}_2^{[n]}|G^{[n]})$ approaches 0 when normalized by $\frac{n}{2}\log(P)$, as first n and then P is sent to infinity.. Thus, the deterministic channel with inputs $(\lfloor X_1^{[n]} \rfloor, \lfloor X_2^{[n]} \rfloor)$, outputs $(\bar{\bar{Y}}_1^{[n]}, \bar{\bar{Y}}_2^{[n]})$, and per-codeword power constraints,

$$\frac{1}{n} \sum_{t=1}^{n} \left((\lfloor X_1(t) \rfloor)^2 + (\lfloor X_2(t) \rfloor)^2 \right) \leq P \tag{108}$$

achieves at least the same DoF as the original canonical channel model.

B. Per-Symbol Power Constraints

Given codewords $(\lfloor X_1^{[n]} \rfloor, \lfloor X_2^{[n]} \rfloor)$ for the deterministic channel with outputs $(\bar{\bar{Y}}_1^{[n]}, \bar{\bar{Y}}_2^{[n]})$, such that the codewords satisfy per-codeword power constraints, define $\forall t \in [1:n]$

$$\bar{X}_1(t) = \lfloor X_1(t) \rfloor \bmod \lceil \sqrt{P} \rceil \tag{109}$$
$$\bar{X}_2(t) = \lfloor X_2(t) \rfloor \bmod \lceil \sqrt{P} \rceil \tag{110}$$
$$\bar{Y}_1(t) = \bar{X}_1(t) \tag{111}$$
$$\bar{Y}_2(t) = \lfloor G(t)\bar{X}_1(t) \rfloor + \bar{X}_2(t) \tag{112}$$
$$\hat{X}_1(t) = \lfloor X_1(t) \rfloor - \bar{X}_1(t) \tag{113}$$
$$\hat{X}_2(t) = \lfloor X_2(t) \rfloor - \bar{X}_2(t) \tag{114}$$
$$\hat{Y}_1(t) = \bar{\bar{Y}}_1(t) - \bar{Y}_1(t) \tag{115}$$
$$\hat{Y}_2(t) = \bar{\bar{Y}}_2(t) - \bar{Y}_2(t) \tag{116}$$

Since $\bar{X}_1(t), \bar{X}_2(t) \in \{0, 1, \cdots, \lceil \sqrt{P} \rceil\}$, the new codewords $(\bar{X}_1^{[n]}, \bar{X}_2^{[n]})$ satisfy per-symbol power constraints. Now let us compare the rates achieved on the channel $(\lfloor X_1^{[n]} \rfloor, \lfloor X_2^{[n]} \rfloor) \longrightarrow (\bar{\bar{Y}}_1^{[n]}, \bar{\bar{Y}}_2^{[n]})$ to the rates achieved on the new channel $(\bar{X}_1^{[n]}, \bar{X}_2^{[n]}) \longrightarrow (\bar{Y}_1^{[n]}, \bar{Y}_2^{[n]})$.

$$I(W_1; \bar{\bar{Y}}_1^{[n]}|G^{[n]})$$
$$= I(W_1; \bar{Y}_1^{[n]} + \hat{Y}_1^{[n]}|G^{[n]}) \tag{117}$$
$$\leq I(W_1; \bar{Y}_1^{[n]}, \hat{Y}_1^{[n]}|G^{[n]}) \tag{118}$$
$$\leq I(W_1; \bar{Y}_1^{[n]}|G^{[n]}) + H(\hat{Y}_1^{[n]}|G^{[n]}) \tag{119}$$
$$\leq I(W_1; \bar{Y}_1^{[n]}|G^{[n]}) + \sum_{t=1}^{n} H(\hat{Y}_1(t)) \tag{120}$$
$$\leq I(W_1; \bar{Y}_1^{[n]}|G^{[n]}) + \sum_{t=1}^{n} H(\hat{X}_1(t)) \tag{121}$$

Similarly

$$I(W_2; \bar{\bar{Y}}_2^{[n]}|G^{[n]})$$
$$= I(W_2; \bar{Y}_2^{[n]} + \hat{Y}_2^{[n]}|G^{[n]}) \tag{122}$$
$$\leq I(W_2; \bar{Y}_2^{[n]}, \hat{Y}_2^{[n]}|G^{[n]}) \tag{123}$$
$$\leq I(W_2; \bar{Y}_2^{[n]}|G^{[n]}) + H(\hat{Y}_2^{[n]}|G^{[n]}) \tag{124}$$
$$\leq I(W_2; \bar{Y}_2^{[n]}|G^{[n]}) + \sum_{t=1}^{n} H(\hat{Y}_2(t)|G(t)) \tag{125}$$
$$= I(W_2; \bar{Y}_2^{[n]}|G^{[n]}) + \sum_{t=1}^{n} H(\lfloor G(t)\lfloor X_1(t) \rfloor \rfloor$$
$$\quad + \lfloor X_2(t) \rfloor - \lfloor G(t)\bar{X}_1(t) \rfloor - \bar{X}_2(t)|G(t)) \tag{126}$$
$$= I(W_2; \bar{Y}_2^{[n]}|G^{[n]}) + \sum_{t=1}^{n} H(\lfloor G(t)\hat{X}_1(t) \rfloor + \hat{X}_2(t)$$
$$\quad + \sum_{t=1}^{n} H(\Delta(t)|G(t)) \tag{127}$$
$$\leq I(W_2; \bar{Y}_2^{[n]}|G^{[n]}) + \sum_{t=1}^{n} H(\lfloor G(t)\hat{X}_1(t) \rfloor|G(t))$$
$$\quad + \sum_{t=1}^{n} H(\hat{X}_2(t)) + \sum_{t=1}^{n} H(\Delta(t)) \tag{128}$$
$$\leq I(W_2; \bar{Y}_2^{[n]}|G^{[n]}) + \sum_{t=1}^{n} H(\hat{X}_1(t)) + \sum_{t=1}^{n} H(\hat{X}_2(t))$$
$$\quad + n \log(3) \tag{129}$$

where $\Delta(t) \in \{-1, 0, 1\}$. To complete the proof we only need to show that $\sum_{t=1}^{n} H(\hat{X}_i(t)) \leq n \, o(\log(P))$, i.e., these terms can contribute no more than 0 in the DoF sense. We will use the fact that any integer number X can be written as $Q\lfloor \frac{X}{Q} \rfloor - Q1(x < 0) + (X \bmod Q)$ where $Q \triangleq \lceil \sqrt{P} \rceil$ is also an integer value.

$$H(\hat{X}_1(t))$$
$$= H \left(Q \left\lfloor \frac{1}{Q} \lfloor X_1(t) \rfloor \right\rfloor - Q\mathbb{I}(\lfloor X_1(t) \rfloor < 0) \right) \tag{130}$$
$$= H \left(\left\lfloor \frac{1}{Q} \lfloor X_1(t) \rfloor \right\rfloor - \mathbb{I}(\lfloor X_1(t) \rfloor < 0) \right) \tag{131}$$
$$\leq H \left(\left\lfloor \frac{1}{Q} \lfloor X_1(t) \rfloor \right\rfloor \right) + H \left(\mathbb{I}(\lfloor X_1(t) \rfloor < 0) \right) \tag{132}$$
$$\leq H \left(\left\lfloor \frac{1}{Q} \lfloor X_1(t) \rfloor \right\rfloor \right) + 1 \tag{133}$$

Define

$$p(t) = \frac{\mathrm{E}(\lfloor X_1(t) \rfloor)^2}{nP} \tag{134}$$

where the expectation is over messages, i.e., the choice of codewords, so that

$$\sum_{t=1}^{n} p(t) \leq 1 \tag{135}$$

Let $\rho_k \triangleq \mathbb{P}(\lfloor X_1(t) \rfloor = k)$, so that

$$\sum_{k=-\infty}^{\infty} k^2 \rho_k \leq p(t)nP \tag{136}$$

Next let us bound the probability $\tilde{\rho}_m$ of $\lfloor X_1(t) \rfloor$ falling within the mth quantization interval.

$$\tilde{\rho}_m \triangleq \sum_{k=mQ}^{(m+1)Q-1} \rho_k, \quad \forall m \in \mathbb{Z} \tag{137}$$

$$\tilde{\rho}_m \leq \frac{1}{\min(m^2, (m+1)^2)Q^2}$$
$$\times \sum_{k=mQ}^{(m+1)Q-1} k^2 \rho_k, \quad \forall m \in \mathbb{Z}/\{0,-1\} \tag{138}$$

$$\leq \frac{np(t)}{\min(m^2, (m+1)^2)}, \quad \forall m \in \mathbb{Z}/\{0,-1\} \tag{139}$$

$$\tilde{\rho}_m^* \triangleq \frac{np(t)}{\min(m^2, (m+1)^2)} \tag{140}$$

In the following derivation we will make use of the fact that $\tilde{\rho}_m \log\left(\frac{1}{\tilde{\rho}_m}\right)$ is an increasing function of $\tilde{\rho}_m$ when $\tilde{\rho}_m < 1/e$, so for these $\tilde{\rho}_m$ values one can replace $\tilde{\rho}_m$ with $\tilde{\rho}_m^*$ to obtain an outer bound. Further, we will use the fact that the maximum value of $\tilde{\rho}_m \log\left(\frac{1}{\tilde{\rho}_m}\right)$ is $\frac{1}{e \ln(2)}$.

We bound the entropy term in (133) as follows

$$H\left(\left\lfloor \frac{1}{Q} \lfloor X_1(t) \rfloor \right\rfloor\right)$$

$$= \sum_{m=-\infty}^{\infty} \tilde{\rho}_m \log\left(\frac{1}{\tilde{\rho}_m}\right) \tag{141}$$

$$= \sum_{m \in \{0,-1,1\}} \tilde{\rho}_m \log\left(\frac{1}{\tilde{\rho}_m}\right) + \sum_{\substack{m:\tilde{\rho}_m^* > \frac{1}{e} \\ m \notin \{0,-1,1\}}} \tilde{\rho}_m \log\left(\frac{1}{\tilde{\rho}_m}\right)$$

$$+ \sum_{\substack{m:\tilde{\rho}_m^* \leq \frac{1}{e} \\ m \notin \{0,-1,1\}}} \tilde{\rho}_m \log\left(\frac{1}{\tilde{\rho}_m}\right) \tag{142}$$

$$= \sum_{m \in \{0,-1,1\}} \frac{1}{e \ln(2)} + \sum_{\substack{m:\tilde{\rho}_m^* > \frac{1}{e} \\ m \notin \{0,-1,1\}}} \frac{1}{e \ln(2)}$$

$$+ \sum_{\substack{m:\tilde{\rho}_m^* \leq \frac{1}{e} \\ m \notin \{0,-1,1\}}} \tilde{\rho}_m^* \log\left(\frac{1}{\tilde{\rho}_m^*}\right) \tag{143}$$

$$\leq \frac{3}{e \ln(2)} + \frac{4\sqrt{np(t)}}{e \ln(2)} + 2 \sum_{m=2}^{\infty} \frac{np(t)}{m^2} \log\left(\frac{m^2}{np(t)}\right) \tag{144}$$

$$\leq \frac{3 + 4\sqrt{np(t)}}{e \ln(2)} + 4np(t) \sum_{m=2}^{\infty} \frac{\log(m)}{m^2}$$

$$+ 2np(t) \log\left(\frac{1}{np(t)}\right) \sum_{m=2}^{\infty} \frac{1}{m^2} \tag{145}$$

$$\leq \frac{3 + 4\sqrt{np(t)}}{e \ln(2)} + 6np(t)$$

$$+ 2np(t) \log\left(\frac{1}{np(t)}\right) \tag{146}$$

$$\leq \frac{3 + 4\max(1, np(t))}{e \ln(2)} + 6np(t) + \frac{2}{e \ln(2)} \tag{147}$$

where in (145) we used the facts that $\sum_{m=2}^{\infty} \frac{\log(m)}{m^2} \leq 1.5$ and $\sum_{m=2}^{\infty} \frac{1}{m^2} \leq 1$. In (147) we used the property that the maximum value of the function $x \log\left(\frac{1}{x}\right)$ is $\frac{1}{e \ln(2)}$.

Substituting into (133), summing over all $t \in [1:n]$, and using (135) we have

$$\sum_{t=1}^{n} H(\hat{X}_1(t))$$

$$\leq n + \frac{5n}{e \ln(2)} + \frac{4}{e \ln(2)} \sum_{t=1}^{n}(1 + np(t)) + 6n \sum_{t=1}^{n} p(t)$$

$$\leq n\left(1 + \frac{13}{e \ln(2)} + 6\right) \tag{148}$$

$$\leq n \, o(\log(P)) \tag{149}$$

The same arguments show that $\sum_{t=1}^{n} H(\hat{X}_2(t)) \leq n \, o(\log(P))$ as well. Thus, we conclude that the deterministic channel with per-symbol power constraints, achieves at least the same DoF as the deterministic channel with per-codeword power constraints.

Define

$$X_1(t) = \tilde{G}_{11}(t)\tilde{X}_1(t) + \tilde{G}_{12}(t)\tilde{X}_2(t) \tag{150}$$

$$X_2(t) = \left(\frac{\det(\tilde{\mathbf{G}}(t))}{\tilde{G}_{11}(t)}\right)\tilde{X}_2(t) \tag{151}$$

$$G(t) = \frac{\tilde{G}_{21}(t)}{\tilde{G}_{11}(t)} \tag{152}$$

$$P = (2M^2 + M^4)\tilde{P} \tag{153}$$

Note that $P = \Theta(\tilde{P})$, and $G(t)$ is bounded away from zero and infinity, as $\frac{1}{M^2} < |G(t)| < M^2, \forall t \in \mathbb{N}$.

$$(X_1(t))^2 + (X_2(t))^2$$

$$\leq \left((\tilde{G}_{11}(t))^2 + (\tilde{G}_{12}(t))^2 + \left(\frac{\det(\tilde{\mathbf{G}}(t))}{\tilde{G}_{11}(t)}\right)^2\right)$$

$$\times((\tilde{X}_1(t))^2 + (\tilde{X}_2(t))^2) \tag{154}$$

$$\leq (2M^2 + M^4)((\tilde{X}_1(t))^2 + (\tilde{X}_2(t))^2) \tag{155}$$

In addition to the channel vector for user 1, let us allow the CSIT to include the determinant of the channel matrix. This cannot reduce capacity, so it can only make the outer bound stronger.

$$\tilde{G}_{11}(t), \tilde{G}_{12}(t), \det(\mathbf{G}(t)) \in \mathcal{T}, \quad \forall t \in \mathbb{N} \tag{156}$$

Note that for continuous distributions, $G(t)$ is not a function of $\mathcal{T}$. With the available CSIT, suppose the transmitter sets:

$$\tilde{X}_1(t) = \frac{1}{\tilde{G}_{11}(t)}X_1(t) - \left(\frac{\tilde{G}_{12}(t)}{\det(\tilde{\mathbf{G}}(t))}\right)X_2(t) \tag{157}$$

$$\tilde{X}_2(t) = \left(\frac{\tilde{G}_{11}(t)}{\det(\tilde{\mathbf{G}}(t))}\right)X_2(t) \tag{158}$$

Substituting into (1) we obtain the canonical channel model (4). Noting that the transformation from $\tilde{X}_1(t), \tilde{X}_2(t)$ to $X_1(t), X_2(t)$ is invertible and that the new power constraint (5) allows all feasible $\tilde{X}_1(t), \tilde{X}_2(t)$, it is evident that the capacity of the channel in its canonical form cannot be smaller than

that of the original channel. Thus, the canonical channel transformation is valid for our DoF outer bound. □

Remark: It is not necessary to provide side information of the determinant of the channel matrix to the transmitter. One could also normalize the desired channel coefficient values to unity by scaling the received signals at the receivers, which would only scale the noise variance by a bounded amount that is inconsequential for DoF. We choose to provide the determinant as side information to the transmitter, because for a pessimistic outer bound that shows the collapse of DoF, including more CSIT only makes the result stronger. It shows that even this additional CSIT cannot prevent the collapse of DoF.

REFERENCES

[1] S. A. Jafar, "Interference alignment: A new look at signal dimensions in a communication network," *Found. Trends Commun. Inf. Theory*, vol. 7, no. 1, pp. 1–136, 2011.

[2] A. Lapidoth, S. Shamai (Shitz), and M. Wigger, "On the capacity of fading MIMO broadcast channels with imperfect transmitter side-information," in *Proc. 43rd Annu. Allerton Conf. Commun., Control Comput.*, Sep. 2005.

[3] A. Lapidoth, S. Shamai (Shitz), and M. Wigger, "Degrees of freedom in MIMO broadcast with finite precision CSI," in *Proc. ITA Open Problems Session*, Feb. 2006.

[4] C. Huang, S. A. Jafar, S. Shamai (Shitz), and S. Vishwanath, "On degrees of freedom region of MIMO networks without channel state information at transmitters," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 849–857, Feb. 2012.

[5] C. S. Vaze and M. K. Varanasi, "The degree-of-freedom regions of MIMO broadcast, interference, and cognitive radio channels with no CSIT," *IEEE Trans. Inf. Theory*, vol. 58, no. 8, pp. 5354–5374, Aug. 2012.

[6] S. A. Jafar, "Blind interference alignment," *IEEE J. Sel. Topics Signal Process.*, vol. 6, no. 3, pp. 216–227, Jun. 2012.

[7] H. Weingarten, S. Shamai (Shitz), and G. Kramer, "On the compound MIMO broadcast channel," in *Proc. Annu. Inf. Theory Appl. Workshop UCSD*, Jan. 2007, pp. 1–7.

[8] T. Gou, S. A. Jafar, and C. Wang, "On the degrees of freedom of finite state compound wireless networks," *IEEE Trans. Inf. Theory*, vol. 57, no. 6, pp. 3268–3308, Jun. 2011.

[9] M. A. MaddahAli, "On the degrees of freedom of the compound MISO broadcast channels with finite states," in *Proc IEEE Int. Symp. Inf. Theory*, Jun. 2010, pp. 2273–2277.

[10] R. Tandon, S. A. Jafar, S. Shamai (Shitz), and H. V. Poor, "On the synergistic benefits of alternating CSIT for the MISO broadcast channel," *IEEE Trans. Inf. Theory*, vol. 59, no. 7, pp. 4106–4128, Jul. 2013.

[11] C. Hao, B. Rassouli, and B. Clerckx. (Oct. 2013). "Degrees-of-freedom region of MISO-OFDMA broadcast channel with imperfect CSIT." [Online]. Available: http://arxiv.org/abs/1310.6669

[12] T. Liu and P. Viswanath, "An extremal inequality motivated by multiterminal information-theoretic problems," *IEEE Trans. Inf. Theory*, vol. 53, no. 5, pp. 1839–1851, May 2007.

[13] B. Rassouli, C. Hao, and B. Clerckx. (2013). "DoF analysis of the K-user MISO broadcast channel with alternating CSIT." [Online]. Available: http://arxiv.org/abs/1311.6647

[14] C. Hao and B. Clerckx. (2013). "MISO broadcast channel with imperfect and (un)matched CSIT in the frequency domain: DoF region and transmission strategies." [Online]. Available: http://arxiv.org/abs/1310.6657

[15] T. Gou and S. A. Jafar, "Optimal use of current and outdated channel state information: Degrees of freedom of the MISO BC with mixed CSIT," *IEEE Commun. Lett.*, vol. 16, no. 7, pp. 1084–1087, Jul. 2012.

[16] S. Yang, M. Kobayashi, D. Gesbert, and X. Yi, "Degrees of freedom of time correlated MISO broadcast channel with delayed CSIT," *IEEE Trans. Inf. Theory*, vol. 59, no. 1, pp. 315–328, Jan. 2013.

[17] J. Korner and K. Marton, "Images of a set via two channels and their role in multi-user communication," *IEEE Trans. Inf. Theory*, vol. 23, no. 6, pp. 751–761, Nov. 1977.

[18] V. R. Cadambe and S. A. Jafar, "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.

[19] G. Caire, N. Jindal, and S. Shamai (Shitz), "On the required accuracy of transmitter channel state information in multiple antenna broadcast channels," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, Nov. 2007, pp. 287–291.

[20] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, "Multiuser MIMO achievable rates with downlink training and channel state feedback," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2845–2896, Jun. 2010.

[21] A. S. Avestimehr, S. N. Diggavi, and D. N. C. Tse, "Wireless network information flow: A deterministic approach," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 1872–1905, Apr. 2011.

[22] G. Bresler and D. Tse, "The two-user Gaussian interference channel: A deterministic view," *Eur. Trans. Telecommun.*, vol. 19, no. 4, pp. 333–354, Jun. 2008.

[23] 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.

[24] E. M. Stein and R. Shakarchi, *Real Analysis: Measure Theory, Integration, and Hilbert Spaces*. Princeton, NJ, USA: Princeton Univ. Press, 2009.

[25] V. R. Cadambe and S. A. Jafar, "Interference alignment and the degrees of freedom of wireless X networks," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 3893–3908, Sep. 2009.

[26] M. A. Maddah-Ali and D. Tse, "Completely stale transmitter channel state information is still very useful," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4418–4431, Jul. 2012.

[27] S. A. Jafar, "Topological interference management through index coding," *IEEE Trans. Inf. Theory*, vol. 60, no. 1, pp. 529–568, Jan. 2014.

Arash Gholami Davoodi is a Ph.D. student at the Department of Electrical Engineering and Computer Science at the University of California Irvine, Irvine, CA USA under the supervision of Prof. Syed Ali Jafar. He received his B.Sc. and M.Sc. degrees in Electrical Engineering (Communication Systems) in 2009 and 2011 from Sharif University of Technology, Tehran, Iran. His research is focused on multiuser information theory, wireless communications and network coding. Arash received a Best Conference Paper Award at IEEE GLOBECOM 2014.

Syed Ali Jafar (S'99–M'04–SM'09–F'14) received his B. Tech. from IIT Delhi, India, in 1997, M.S. from Caltech, USA, in 1999, and Ph.D. from Stanford, USA, in 2003, all in Electrical Engineering. His industry experience includes positions at Lucent Bell Labs, Qualcomm Inc. and Hughes Software Systems. He is a Professor in the Department of Electrical Engineering and Computer Science at the University of California Irvine, Irvine, CA USA. His research interests include multiuser information theory, wireless communications and network coding. Dr. Jafar received the Blavatnik National Laureate in Physical Sciences and Engineering in 2015, the UCI Academic Senate Distinguished Mid-Career Faculty Award for Research in 2015, the School of Engineering Mid-Career Excellence in Research Award in 2015, the School of Engineering Maseeh Outstanding Research Award in 2010, the ONR Young Investigator Award in 2008, and the NSF CAREER award in 2006. His co-authored papers received the IEEE Information Theory Society Best Paper Award in 2009, IEEE Communications Society Best Tutorial Paper Award in 2013, IEEE Communications Society Heinrich Hertz Award in 2015, IEEE Signal Processing Society Young Author Best Paper Award (to student co-authors) in 2015, an IEEE GLOBECOM Best Paper Award in 2012 and an IEEE GLOBECOM Best Paper Award in 2014. Dr. Jafar received the UC Irvine EECS Professor of the Year award five times, in 2006, 2009, 2011, 2012, and 2014, from the Engineering Students Council and the Teaching Excellence Award in 2012 from the School of Engineering. He was a University of Canterbury Erskine Fellow in 2010 and an IEEE Communications Society Distinguished Lecturer for 2013-2014. Dr. Jafar was recognized as a Thomson Reuters Highly Cited Researcher and included by Sciencewatch among The World's Most Influential Scientific Minds in 2014 and again in 2015. He served as Associate Editor for IEEE TRANSACTIONS ON COMMUNICATIONS 2004-2009, for IEEE COMMUNICATIONS LETTERS 2008-2009 and for IEEE Transactions on Information Theory 2009- 2012. Dr. Jafar was elevated to IEEE Fellow, Class of 2014, for contributions to analyzing the capacity of wireless communication networks.