

Settling Conjectures on the Collapse of Degrees of Freedom under Finite Precision CSIT

Arash Gholami Davoodi and Syed A. Jafar
 Center for Pervasive Communications and Computing (CPCC)
 University of California Irvine, Irvine, CA 92697
 Email: {gholamid, syed}@uci.edu

Abstract—A conjecture made by Lapidoth, Shamai and Wigger 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 2 antennas and each user is equipped with 1 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 degrees of freedom of wireless networks—including interference and X networks—under channel uncertainty at the transmitter(s). In this work 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. Reminiscent of Korner and Marton’s work on the images of a set, 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, while remaining resolvable at the desired receiver. We are also able to generalize the result to arbitrary number of users, including the K user interference channel. Remarkably, for the K user interference channel, this work and the earlier work by Cadambe and Jafar reveal two contrasting sides of the same coin. Both works close a gap between the best previously known DoF inner bound of 1 and the best previously known DoF outer bound of $K/2$. However, while Cadambe and Jafar do so in the optimistic direction, showing that $K/2$ is optimal under perfect CSIT, here we close the gap in the pessimistic direction, showing that 1 DoF is optimal under finite precision CSIT.

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, 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 [3], [4]. On the other hand, the idea of blind interference alignment introduced by Jafar in [5] 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 [6] 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 [7], and by Maddah-Ali in [8], 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 [9] — 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. Overview of Contribution

The main contribution of this work is to prove the conjecture of Lapidath, 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 [10], 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 $P \rightarrow \infty$. 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 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 finite precision 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.

In the full paper posted on ArXiv [11] the result is further generalized to the K user MISO BC, which also settles the DoF of the K user interference channel under finite precision CSIT. Remarkably, the best known outer bounds for K user interference channel under finite precision CSIT prior to this work was $\frac{K}{2}$ (same as with perfect CSIT). Thus, this work on finite precision CSIT and the work of Cadambe and Jafar in [12] 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.

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 $\limsup_{x \rightarrow \infty} \frac{|f(x)|}{|g(x)|} < \infty$. $f(x) = o(g(x))$ denotes that $\limsup_{x \rightarrow \infty} \frac{|f(x)|}{|g(x)|} = 0$. $f(x) = \Theta(g(x))$ denotes that there exists a positive finite constant, M , such that $\frac{1}{M}g(x) \leq f(x) \leq Mg(x)$, $\forall x$. We use $\mathbb{P}(\cdot)$ to denote the probability function $\text{Prob}(\cdot)$. The index set $\{1, 2, \dots, n\}$ is represented compactly as $[1 : n]$ or simply $[n]$ when it would cause no confusion. $X^{[s]}$ represents $\{X(t) : t \in [s]\}$. For example, $X^{[n]} = \{X(1), X(2), \dots, X(n)\}$. With some abuse of notation we use $\{X\}$ to denote the set of values that can be taken by the random variable X . The cardinality of a set A is denoted as $|A|$.

II. 2 USER MISO BC WITH PERFECT CSIT FOR 1 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. We will use the following canonical form of the channel model.

A. Canonical Form

Without loss of generality, for the purpose of deriving a DoF outer bound the channel model is reduced to the following form, which is preferable due to the consolidation of channel parameters (See Appendix B in [11] for justification).

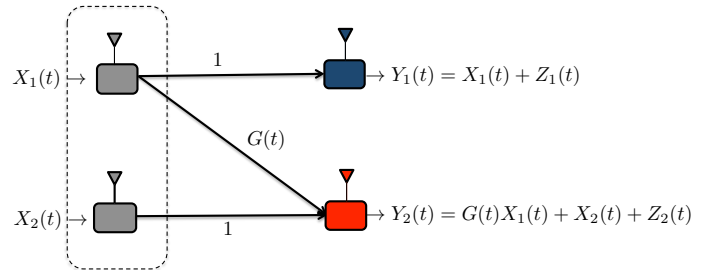


Fig. 1. 2 user MISO BC with perfect CSIT for user 1.

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

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

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

$$\frac{1}{n} \sum_{t=1}^n [(X_1(t))^2 + (X_2(t))^2] \leq P. \quad (2)$$

B. 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. For a given power constraint P , rate vector $[R_1, R_2]$ is said to be achievable if there exists a sequence of codebooks,

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 tuple (d_1, d_2) is said to be achievable if there exist $(R_1(P), R_2(P)) \in \mathcal{C}(P)$ such that $d_1 = \lim_{P \rightarrow \infty} \frac{R_1(P)}{\frac{1}{2} \log(P)}$, $d_2 = \lim_{P \rightarrow \infty} \frac{R_2(P)}{\frac{1}{2} \log(P)}$. 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 $\mathcal{D}_\Sigma = \max_{(d_1, d_2) \in \mathcal{D}} (d_1 + d_2)$.

C. Non-degenerate Channel Uncertainty

By non-degenerate channel uncertainty we mean that the probability density function (pdf) of the channel coefficients exists and is bounded. While in the full paper we allow the channel realizations to be correlated in time, for simplicity here, let us assume that conditioned on all available CSIT,¹ \mathcal{T} , the channel realizations $G(t)$ are independently generated across time, although not necessarily identically distributed. Then we only require that f_{\max} as defined below is finite.

$$f_{\max} = \max \left(1, \sup_{g(t) \in \mathcal{G}(t), t \in \mathbb{Z}_+} f_{G(t)|\mathcal{T}}(g(t)) \right) \quad (3)$$

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.

III. MAIN RESULT

Theorem 1: For the 2 user MISO BC with non-degenerate channel uncertainty, $\mathcal{D}_\Sigma \leq 1$.

The result settles the conjecture by Lapidoth et al. in [2] for all non-degenerate channel uncertainty models. It also settles the ‘‘PN’’ conjecture by Tandon et al. in [9] (see [11]). Further, generalizations to the K user case presented in the full paper [11] also settle the collapse of DoF for K user interference channel as well as the $M \times N$ user X channel.

IV. ALIGNED IMAGE SETS

The main idea we want to illustrate intuitively is a geometrical notion of aligned images of codewords—loosely related to Korner and Marton’s work on the images of a set in [10] 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,

$$\mathcal{D}_\Sigma \leq 1 + \limsup_{P \rightarrow \infty} \limsup_{n \rightarrow \infty} \frac{1}{\frac{n}{2} \log(P)} \left(h(Y_1^{[n]}|G^{[n]}) - h(Y_2^{[n]}|G^{[n]}) \right) \quad (4)$$

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, $h(Y_2^{[n]}|G^{[n]}) = h(G^{[n]}X_1^{[n]} + X_2^{[n]}|G^{[n]})$ we wish to minimize, because it represents the interference

¹For example, $G(t) = \hat{G}(t) + \tilde{G}(t)$, where $\hat{G}(t) = \mathcal{T}$ is the estimate available to the transmitter and $\tilde{G}(t)$ is the estimation error.

seen by user 2, due to the information being sent to user 1. 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 $n = 1$, 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:

$$\log |\{X_1\}| - \log |\{GX_1 + X_2\}| \quad (5)$$

averaged over G . $|\{A\}|$ means the cardinality of the set of values taken by the variable A . The codebook is the set of (X_1, X_2) values. Note that $|\{X_1\}|$, the number of distinct values of X_1 , is the number of distinct ‘‘codewords’’ as seen by user 1, who (once noise is ignored) only sees $Y_1 = X_1$, so that his ‘‘rate’’ is $\log |\{X_1\}|$. 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 .

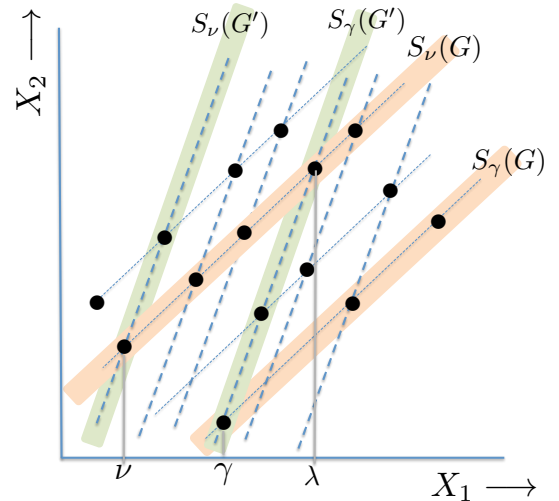


Fig. 2. Two codewords, ν and γ , and their equivalence classes, S_ν and S_γ under channel realizations G and G' .

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.

$$Gx_1 + x_2 = Gx'_1 + x'_2 \quad (6)$$

$$\Rightarrow G = -\left(\frac{x'_2 - x_2}{x'_1 - x_1}\right) \quad (7)$$

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. Thus, these lines of the same slope, $-G$, partition the set of codewords into equivalence classes, such that 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. 2. Since the X_2 values are functions of X_1 values, in the figure we label the codewords only on the X_1 axis. The codeword ν 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' .

B. Sketch of Proof

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 [13], [14], [15], [16], 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 $\bar{Y}_1 = \bar{X}_1$. 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 precisely 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, $H(\bar{X}_1^{[n]}|\bar{Y}_2^{[n]}, G^{[n]}) \leq E_{G^{[n]}, \bar{Y}_2^{[n]}} \left[\log \left[|S_{\bar{X}_1^{[n]}}(G^{[n]})| \right] \right] \leq \log \left(E_{G^{[n]}, \bar{Y}_2^{[n]}} \left[|S_{\bar{X}_1^{[n]}}(G^{[n]})| \right] \right)$, where the last step followed from 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 $\log \left(E \left[|S_{\bar{X}_1^{[n]}}(G^{[n]})| \right] \right)$ is bounded above by $\approx n \log(f_{\max} \delta) + n \log(\log(P))$. Normalizing by $\frac{n}{2} \log(P)$ and sending first n and then P to infinity sends this term to 0. Thus, combining with (4) produces the sum-DoF outer bound value 1, giving us the result of Theorem 1.

V. PROOF OF THEOREM 1

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, leading to a deterministic channel model, whose DoF will 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

$$\bar{Y}_1(t) = \bar{X}_1(t) \quad (8)$$

$$\bar{Y}_2(t) = \lfloor G(t)\bar{X}_1(t) \rfloor + \bar{X}_2(t) \quad (9)$$

and the set of inputs that satisfy the per-codeword power constraints defined as

$$\bar{\mathcal{X}}^{[n]} = \left\{ (\bar{X}_1^{[n]}, \bar{X}_2^{[n]}) \in \mathbb{Z}^{[n]} \times \mathbb{Z}^{[n]} : \right.$$

$$\left. \bar{X}_1(t), \bar{X}_2(t) \in \{0, 1, \dots, \lceil \sqrt{P} \rceil\}, \forall t \in [1 : n] \right\}$$

The 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 follows along the lines of similar proofs by Bresler and Tse in [14] and is provided in Appendix A of [11] for the sake of completeness.

2) Difference of Entropies Representing Desired Signal and Interference Dimensions

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

$$nR_1 \leq I(W_1; \bar{Y}_1^{[n]} | W_2, G^{[n]}) + o(n) \quad (10)$$

$$= H(\bar{Y}_1^{[n]} | W_2, G^{[n]}) + o(n) \quad (11)$$

$$nR_2 \leq I(W_2; \bar{Y}_2^{[n]} | G^{[n]}) + o(n) \quad (12)$$

$$= H(\lfloor G^{[n]} \bar{X}_1^{[n]} \rfloor + \bar{X}_2^{[n]} | G^{[n]}) - H(\bar{Y}_2^{[n]} | W_2, G^{[n]}) + o(n) \quad (13)$$

$$\leq \frac{n}{2} \log(P) - H(\bar{Y}_2^{[n]} | W_2, G^{[n]}) + n o(\log(P)) + o(n) \quad (14)$$

$$\Rightarrow n(R_1 + R_2) \leq \frac{n}{2} \log(P) + n o(\log(P)) + o(n) + [H(\bar{Y}_1^{[n]} | W_2, G^{[n]}) - H(\bar{Y}_2^{[n]} | W_2, G^{[n]})]$$

$$\Rightarrow \mathcal{D}_\Sigma \leq 1 + \limsup_{P \rightarrow \infty} \limsup_{n \rightarrow \infty} \quad (15)$$

$$\frac{[H(\bar{Y}_1^{[n]} | W_2, G^{[n]}) - H(\bar{Y}_2^{[n]} | W_2, G^{[n]})]}{\frac{n}{2} \log(P)}$$

$$\leq 1 + \limsup_{P \rightarrow \infty} \limsup_{n \rightarrow \infty} \max_{w_2 \in [1:2^{nR_2}]}$$

$$\frac{[H(\bar{Y}_1^{[n]} | W_2 = w_2, G^{[n]}) - H(\bar{Y}_2^{[n]} | W_2 = w_2, G^{[n]})]}{\frac{n}{2} \log(P)}$$

$$\leq 1 + \bar{\mathcal{D}}_\Delta \quad (16)$$

see [11] for more details. What remains is to bound the difference of entropy terms:

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

$$\frac{H(\bar{X}_1^{[n]} | G^{[n]}) - H(\lfloor G^{[n]} \bar{X}_1^{[n]} \rfloor + \bar{X}_2^{[n]} | G^{[n]})}{\frac{n}{2} \log(P)} \quad (17)$$

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., $\bar{X}_2^{[n]} = \mathcal{L}(\bar{X}_1^{[n]})$. In general, because the mapping may be random, \mathcal{L} is a random variable. Because conditioning cannot increase entropy,

$$\begin{aligned} & H(\lfloor G^{[n]} \bar{X}_1^{[n]} \rfloor + \mathcal{L}(\bar{X}_1^{[n]}) | G^{[n]}) \\ & \geq H(\lfloor G^{[n]} \bar{X}_1^{[n]} \rfloor + \mathcal{L}(\bar{X}_1^{[n]}) | G^{[n]}, \mathcal{L}) \quad (18) \\ & \geq \min_{L \in \{\mathcal{L}\}} H(\lfloor G^{[n]} \bar{X}_1^{[n]} \rfloor + \mathcal{L}(\bar{X}_1^{[n]}) | G^{[n]}, \mathcal{L} = L) \end{aligned}$$

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

$$\bar{X}_2^{[n]}(\bar{X}_1^{[n]}) = L_o(\bar{X}_1^{[n]}) \quad (19)$$

we have

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

$$\frac{H(\bar{X}_1^{[n]} | G^{[n]}) - H(\lfloor G^{[n]} \bar{X}_1^{[n]} \rfloor + \bar{X}_2^{[n]}(\bar{X}_1^{[n]}) | G^{[n]})}{\frac{n}{2} \log(P)} \quad (20)$$

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 $\bar{v}^{[n]} \in \{\bar{X}_1^{[n]}\}$ 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.

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

It is worthwhile to point out that the cardinality $|S_{\bar{v}^{[n]}}(G^{[n]})|$ is a function of $G^{[n]}$, which is a simple function, and therefore a measurable function [11].

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

$$\begin{aligned} & H(\bar{X}_1^{[n]} | G^{[n]}) = H(\bar{X}_1^{[n]}, S_{\bar{X}_1^{[n]}}(G^{[n]}) | G^{[n]}) \\ & = H(S_{\bar{X}_1^{[n]}}(G^{[n]}) | G^{[n]}) + H(\bar{X}_1^{[n]} | S_{\bar{X}_1^{[n]}}(G^{[n]}), G^{[n]}) \\ & = H(\lfloor G^{[n]} \bar{X}_1^{[n]} \rfloor + \bar{X}_2^{[n]}(\bar{X}_1^{[n]}) | G^{[n]}) + H(\bar{X}_1^{[n]} | S_{\bar{X}_1^{[n]}}(G^{[n]}), G^{[n]}) \quad (21) \end{aligned}$$

$$\leq H(\lfloor G^{[n]} \bar{X}_1^{[n]} \rfloor + \bar{X}_2^{[n]}(\bar{X}_1^{[n]}) | G^{[n]}) + \mathbb{E} \left[\log(|S_{\bar{X}_1^{[n]}}(G^{[n]})|) \right] \quad (22)$$

$$\leq H(\lfloor G^{[n]} \bar{X}_1^{[n]} \rfloor + \bar{X}_2^{[n]}(\bar{X}_1^{[n]}) | G^{[n]}) + \log \left(\mathbb{E} \left[|S_{\bar{X}_1^{[n]}}(G^{[n]})| \right] \right) \quad (23)$$

where (22) follows because uniform distribution maximizes entropy, and (23) follows from Jensen's inequality. Rearranging terms, we note that

$$\begin{aligned} \hat{\mathcal{D}}_\Delta & \leq \limsup_{P \rightarrow \infty} \limsup_{n \rightarrow \infty} \max_{\substack{\mathbb{P}(\bar{X}_1^{[n]}, \bar{X}_2^{[n]}) \\ (\bar{X}_1^{[n]}, \bar{X}_2^{[n]}) \in \bar{\mathcal{X}}^{[n]}} \\ & \frac{\log \left(\mathbb{E} \left[|S_{\bar{X}_1^{[n]}}(G^{[n]})| \right] \right)}{\frac{n}{2} \log(P)} \quad (24) \end{aligned}$$

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

$$\begin{aligned} \lfloor G^{[n]} \bar{x}_1^{[n]} \rfloor - \lfloor G^{[n]} \bar{v}^{[n]} \rfloor & = \bar{X}_2^{[n]}(\bar{v}^{[n]}) - \bar{X}_2^{[n]}(\bar{x}_1^{[n]}) \\ \Rightarrow G^{[n]}(\bar{x}_1^{[n]} - \bar{v}^{[n]}) & \in \bar{X}_2^{[n]}(\bar{x}_1^{[n]}) - \bar{X}_2^{[n]}(\bar{v}^{[n]}) + \Delta_{(-1,1)}^{[n]} \quad (25) \end{aligned}$$

where $-1 < \Delta_{(-1,1)}(t) < 1, \forall t \in [1 : n]$. 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 $f_{\max}^n \prod_{t=1}^n 1_{(\bar{x}_1(t) \neq \bar{v}(t))} \leq f_{\max}^n$, we can bound the probability that the images of two codewords align as follows.

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

7) Bounding the Average Size of Aligned Image Sets

$$\begin{aligned} \mathbb{E} & \left[|S_{\bar{v}^{[n]}}(G^{[n]})| \right] \\ &= \sum_{\bar{x}_1^{[n]} \in \{\bar{X}_1^{[n]}\}} \mathbb{P}(\bar{x}_1^{[n]} \in S_{\bar{v}^{[n]}}(G^{[n]})) \end{aligned} \quad (26)$$

$$= 1 + \sum_{\substack{\bar{x}_1^{[n]} \in \{\bar{X}_1^{[n]}\} \\ \bar{x}_1^{[n]} \neq \bar{v}^{[n]}}} \mathbb{P}(\bar{x}_1^{[n]} \in S_{\bar{v}^{[n]}}(G^{[n]})) \quad (27)$$

$$\leq 1 + (2f_{\max})^n \sum_{\substack{\bar{x}_1^{[n]} \in \{\bar{X}_1^{[n]}\} \\ \bar{x}_1^{[n]} \neq \bar{v}^{[n]}}} \prod_{t: \bar{x}_1(t) \neq \bar{v}(t)} \frac{1}{|\bar{x}_1(t) - \bar{v}(t)|} \quad (28)$$

$$\leq 1 + (2f_{\max})^n \prod_{t=1}^n \left(2 \log(\sqrt{P}) + 3 \right) \quad (29)$$

where (28) follows by recognizing that $|\bar{x}_1(t) - \bar{v}(t)|$ can only take integer values between 1 and $\lceil \sqrt{P} \rceil$, and writing the sum of products as the product of sums.

8) Combining the Bounds to Complete the Proof

Combining (24) and (29) we have

$$\begin{aligned} \hat{\mathcal{D}}_{\Delta} & \leq \limsup_{P \rightarrow \infty} \limsup_{n \rightarrow \infty} \\ & \frac{\log \left(1 + (2f_{\max})^n \left(2 \log(\sqrt{P}) + 3 \right)^n \right)}{\frac{n}{2} \log(P)} \\ & = 0 \end{aligned} \quad (30)$$

Finally combining (30) with (16) and (20) we have the desired outer bound, $\mathcal{D}_{\Sigma} \leq 1$ for the 2 user MISO BC with non-degenerate channel uncertainty. \square

VI. 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 [12], [17]. However, as evident from the conjecture of Lapidath, 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.

VII. ACKNOWLEDGMENTS

This work was supported in part by NSF CCF 0963925 and by NSF CCF 1319104.

REFERENCES

- [1] S. Jafar, "Interference alignment: A new look at signal dimensions in a communication network," in *Foundations and Trends in Communication and Information Theory*, 2011, pp. 1–136.
- [2] A. Lapidath, S. Shamai, and M. Wigger, "On the capacity of fading MIMO broadcast channels with imperfect transmitter side-information," in *Proceedings of 43rd Annual Allerton Conference on Communications, Control and Computing*, Sep. 28–30, 2005.
- [3] C. Huang, S. A. Jafar, S. Shamai, and S. Vishwanath, "On Degrees of Freedom Region of MIMO Networks without Channel State Information at Transmitters," *IEEE Transactions on Information Theory*, no. 2, pp. 849–857, Feb. 2012.
- [4] C. S. Vaze and M. K. Varanasi, "The degree-of-freedom regions of MIMO broadcast, interference, and cognitive radio channels with no CSIT," *IEEE Transactions on Information Theory*, vol. 58, no. 8, pp. 5354–5374, 2012.
- [5] S. A. Jafar, "Blind Interference Alignment," *IEEE Journal of Selected Topics in Signal Processing*, vol. 6, no. 3, pp. 216–227, June 2012.
- [6] H. Weingarten, S. Shamai, and G. Kramer, "On the compound MIMO broadcast channel," in *Proceedings of Annual Information Theory and Applications Workshop UCSD*, Jan 2007.
- [7] T. Gou, S. Jafar, and C. Wang, "On the degrees of freedom of finite state compound wireless networks," *IEEE Transactions on Information Theory*, vol. 57, no. 6, pp. 3268–3308, June 2011.
- [8] M. A. Maddah-Ali, "The degrees of freedom of the compound MIMO broadcast channels with finite states," *CoRR*, vol. abs/0909.5006, 2009. [Online]. Available: <http://arxiv.org/abs/0909.5006>
- [9] R. Tandon, S. A. Jafar, S. Shamai, and H. V. Poor, "On the synergistic benefits of alternating CSIT for the MISO BC," *IEEE Transactions on Information Theory*, vol. 59, no. 7, pp. 4106–4128, July 2013.
- [10] J. Korner and K. Marton, "Images of a set via two different channels and their role in multiuser communication," *IEEE Trans. Inform. Theory*, vol. 23, pp. 751–761, Nov. 1977.
- [11] A. G. Davoodi and S. A. Jafar, "Aligned Image Sets under Channel Uncertainty: Settling a Conjecture by Lapidath, Shamai and Wigger on the Collapse of Degrees of Freedom under Finite Precision CSIT," <http://arxiv.org/abs/1403.1541v1.pdf>, March 2014.
- [12] V. Cadambe and S. Jafar, "Interference alignment and the degrees of freedom of the K user interference channel," *IEEE Transactions on Information Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.
- [13] A. S. Avestimehr, S. Diggavi, and D. Tse, "Wireless network information flow: A deterministic approach," *IEEE Trans. on Information Theory*, vol. 57, pp. 1872–1905, 2011.
- [14] G. Bresler and D. Tse, "The two-user Gaussian interference channel: a deterministic view," *European Transactions in Telecommunications*, vol. 19, no. 4, pp. 333–354, June 2008.
- [15] I. Shomorony and S. Avestimehr, "On the role of deterministic models in k x k x k wireless networks," in *IEEE Information Theory Workshop*, 2012.
- [16] R. Etkin and E. Ordentlich, "The degrees-of-freedom of the K-User Gaussian interference channel is discontinuous at rational channel coefficients," *IEEE Trans. on Information Theory*, vol. 55, pp. 4932–4946, Nov. 2009.
- [17] V. Cadambe and S. Jafar, "Interference alignment and the degrees of freedom of wireless X networks," *IEEE Trans. on Information Theory*, no. 9, pp. 3893–3908, Sep 2009.