Feedback in the K-user Interference Channel

Dimitris S. Papailiopoulos
Communication Sciences Institute
University of Southern California
Email: papailio@usc.edu

Changho Suh
Research Laboratory of Electronics
Massachusetts Institute of Technology
Email: chsuh@mit.edu

Alexandros G. Dimakis
Communication Sciences Institute
University of Southern California
Email: dimakis@usc.edu

***Abstract*—THIS PAPER IS ELIGIBLE FOR THE STUDENT PAPER AWARD. We consider the scalar K-user Gaussian interference channel (IC) with feedback, where channel coefficients are fixed over time and frequency and assume that trasceivers operate in full-duplex mode. We focus on two feedback models: (1) each receiver feeds back its received signal to all the transmitters and (2) functions of the received signals are fed back through a *backward* IC. We show that the feedback degrees-of-freedom (fdof) of the first model is $\frac{K}{2}$ if the global channel matrix is invertible. For the second feedback model, we show that fdof $= \frac{3}{2}$ is achievable for the 3-user IC. Then, we show how nontrivial fdof can be achieved for the K-user IC, when the global channel matrix belongs to a specific spectral family of matrices. Our achievability schemes require a *finite* number of signal-space dimensions, contrasting the asymptotic interference alignment schemes for the non-feedback case. Another consequence of feedback is that it can strictly increase the degrees-of-freedom for some classes of ICs.**

***Index Terms*—Degrees-of-freedom, Feedback, K-user Gaussian interference channel, Scalar channel**

I. Introduction

A traditional viewpoint on feedback capacity has been pessimistic over the past few decades. This is mainly due to Shannon's original result [1] which shows that feedback provides no increase in capacity for discrete memoryless point-to-point channels. Although feedback can indeed increase the capacity of multiple access channels [2], the increase in capacity for the Gaussian case is bounded by 1 bit for all channel parameters [3].

Contrary to the traditional belief on feedback capacity, recent results show that feedback is of fundamental importance when communicating over interference channels (ICs) [4], [5], [6]. Specifically, this feedback gain is shown to be *unbounded*, i.e., the gain can be arbitrarily large for certain channel parameters. Even when feedback links are modeled as rate-limited bit-pipes, it was shown in [7] that one bit of feedback is worth one bit of capacity.

In this paper, we study the role of feedback in the context of K-user ICs. Specifically we consider the K-user Gaussian IC with feedback where channel coefficients are constant over time and frequency. We focus on two feedback models: (1) the received signals of all the receivers are fed back to all of the transmitters; (2) functions of the received signals are fed back through a backward Gaussian IC. For all cases that we consider, we assume the transmitter-receiver links to be full-duplex, i.e., a transceiving user can receive and transmit at the same time. For the first idealistic model, we develop a linear achievable scheme and derive a new upper bound to show that the fdof is $\frac{K}{2}$ for all channel realizations that ensure the invertibility of the global channel matrix.

For the second, more realistic, model, we make progress on the 3-user IC. We develop a feedback strategy that builds on the interference alignment-and-cancellation idea, and achieves dof $= \frac{3}{2}$, almost surely for randomly-drawn channels. For the 4-user case, we develop a heuristic algorithm that finds feedback coding coefficients; we observe through simulations that this algorithm ensures the achievability of $\frac{4}{2}$ for a multitude of experiments. For an arbitrary K number of users, we identify a class of channels with specific spectral properties, where nontrivial fdof is achievable with a very simple feedback strategy.

Our results indicate that *i)* feedback can be used to achieve the degrees-of-freedom (dof) of an IC in a *finite* number of signaling dimensions and *ii)* feedback can increase the dof for some classes of scalar ICs where the non-feedback dof is either strictly less than $\frac{K}{2}$ or linear transceivers cannot achieve it.

Related Work: Nonfeedback strategies for the K-user Gaussian IC were studied previously [8], [9], [10], [11]. In their breakthrough paper, Cadambe and Jafar [8] showed that interference alignment (IA) [12], can achieve dof $= \frac{K}{2}$ for vector ICs with arbitrarily large channel diversity, e.g., i.i.d time-varying channels. Etkin and Ordentlich [9] showed that dof is strictly less that $\frac{K}{2}$ for some class of scalar ICs. Motahari *et al.* [10] developed a nonlinear IA technique to show that interestingly dof is $\frac{K}{2}$ with probability 1, for randomly-drawn channel coefficients. In [11], Yetis *et al.* showed that under linear transceiver assumption, dof $= 1$ for scalar ICs. On the other hand, we show that feedback enables dof $= \frac{K}{2}$ only with linear strategies that rely on a finite number of signal-space dimensions.

Feedback strategies were studied by Kramer [4], [5] and Tandon *et al.* [13], [14], assuming that feedback is available from each receiver to its corresponding transmitter. While all of these techniques result in a significant capacity increase for a multitude of interesting channel regimes, these achieve dof $= 1$, thereby showing the suboptimality of the schemes in general. On the other hand, we consider different feedback configurations to develop coding techniques that achieve the optimal dof.

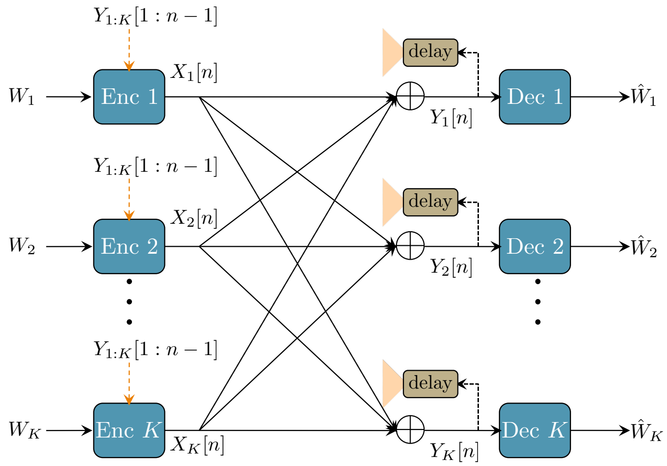

Fig. 1. The scalar K-user Gaussian IC with perfect feedback. Each transmitter gets delayed channel-output feedback from all of the receivers and the channel coefficients are fixed across time.

II. System Model

We assume two feedback models in this work and describe them in the following.

Feedback Model I: Fig. 1 describes the scalar K-user Gaussian IC with perfect feedback. The received signal of receiver k at time n is

$$Y_k[n] = \sum_{m=1}^{K} h_{km} X_m[n] + Z_k[n], \tag{1}$$

where $X_m[n]$ indicates the encoded signal of transmitter m, h_{km} is the "channel processing" coefficient from transmitter m to receiver k, and $Z_k[n]$ denotes zero-mean additive white Gaussian noise at receiver k with normalized power, i.e., $Z_k \sim \mathcal{CN}(0,1)$. We assume an average power constraint on all transmitted signals $E\{X_k[n]\} = P$. We note that the channel coefficients are assumed fixed over time.

The transmitters generate K independent and uniformly distributed messages, $W_k \in \{1, 2, \cdots, M_k\}, \forall k \in [K]$, where $[N]$ denotes the set $\{1, \cdots, N\}$. The encoded signal $X_k[n]$ of transmitter k at time n is a function of its own message and past output sequences of all the receivers

$$X_k[n] = f_k^n \left(W_k, Y_1^{n-1}, \cdots, Y_K^{n-1}\right), \tag{2}$$

where $Y_k^{n-1} \triangleq (Y_k[1], \cdots, Y_k[n-1])$ denotes the sequence of received symbols at receiver k up to time index $n-1$. A rate tuple $(R_1, \cdots, R_K)$ is said to be achievable if there exists a family of codewords, subject to the power constraint, and encoding/decoding functions such that the average decoding error probabilities go to zero as the code length N tends to infinity. The capacity region $\mathcal{C}$ is the closure of the set of all achievable rate pairs. We define the degrees-of-freedom (dof) as dof $\triangleq \lim_{P\to\infty} \max_{(R_1,\cdots,R_K)\in\mathcal{C}} \sum_{k=1}^{K} \frac{R_k}{\log P}$ and call it fdof when feedback is used to achieve it.

Feedback Model II: We consider another feedback model of practical interest where feedback is offered by the backward IC. The fed back signal received at transmitter k at time n is

$$\tilde{Y}_k[n] = \sum_{m=1}^{K} \tilde{h}_{km} \tilde{X}_m[n] + \tilde{Z}_k[n], \tag{3}$$

where $\tilde{X}_m[n]$ indicates the encoded signal of receiver m, $\tilde{h}_{km}$ indicates the backward channel coefficient from receiver m to transmitter k, and $\tilde{Z}_k[n]$ denotes the additive white Gaussian noise observed by transmitter k with normalized power such that $\tilde{Z}_k \sim \mathcal{CN}(0,1)$. Again, we assume that the average transmit power of each receiver is limited by P and that channel coefficients are fixed over time. In this model, the encoded signal $X_k[n]$ of transmitter k at time slot n is a function of the message and feedback signals

$$X_k[n] = g_k^n \left(W_k, \tilde{Y}_k^{n-1}\right). \tag{4}$$

In both models we assume full-duplex transceivers.

III. Main Results

In the following we summarize the main results of our work.

Theorem 1 (Model I): Let $\mathbf{H}$ be a global channel matrix, i.e., $[\mathbf{H}]_{mk} = h_{mk}$ and rank$(\mathbf{H}) = K$. Then, the feedback degrees-of-freedom (fdof) of the scalar K-user IC under model I is

$$\text{fdof} = \frac{K}{2}. \tag{5}$$

Proof: See Section IV. ∎

Theorem 2 (Model II): Let h_{mk} be randomly drawn from a continuous distribution and i.i.d over m and k. Then the fdof of the scalar 3-user IC under model 2 is

$$\text{fdof} = \frac{3}{2}, \quad \text{almost surely}. \tag{6}$$

Proof: The converse proof follows from that of Theorem 1. See Section V for the achievability. ∎

Theorem 3 (Model II): Let $\mathbf{H}$ be a full-rank symmetric channel matrix with L distinct eigenvalues. Then,

$$\text{fdof} = \frac{K}{L}, \tag{7}$$

is achievable.

Proof: See Subection VI-B for the achievability. ∎

Remark 1 (Linear dof*):* Assuming linear transceivers, the non-feedback dof is 1 for the scalar IC [11]. On the other hand, our feedback strategy is linear and thus shows dof gain for arbitrary scalar ICs.

Remark 2 (Finite Symbol Extensions): In sharp contrast to the symbol extended IA schemes and infinite precision real IA schemes of [8] and [10], when feedback is used then the full dof of a channel can be achieved in a *fixed* number of symbol extensions.

Remark 3 (Comparison to the non-feedback dof*):* As an example, we present a channel matrix where the non-feedback dof is probably 1 and feedback can provide gain. Note that $\mathbf{H} = [1\ -1\ -1; -1\ 1\ -1; -1\ -1\ 1]$ is invertible and hence fdof $= \frac{3}{2}$, which improves upon the non-feedback dof $= 1$. For a proof of the converse see Appendix B.

IV. Proof of Theorem 1

A. Achievability

Our achievable scheme operates on 4 time slots. In time-slot 1, each transmitter sends its own symbol and the channel output is given by

$$\mathbf{y}[1] = \mathbf{H}\mathbf{x}[1], \tag{8}$$

where $\mathbf{y}[1] \triangleq [Y_1[1], \cdots, Y_K[1]]^T$ and $\mathbf{x}[1] \triangleq [X_1[1], \cdots, X_K[1]]^T$. We ignore the noise, since we are interested in dof. Under this model, feedback is available from every receiver to every transmitter and allows each transmitter to obtain $\mathbf{y}[1]$ at time 2. Meanwhile each transmitter sends another new symbol.

The invertibility assumption on $\mathbf{H}$ enables each transmitter to compute $\mathbf{H}^{-1}\mathbf{y}[1] = \mathbf{x}[1]$ and therefore decode all of the other users' symbols. This observation naturally motivates us to employ the zero-forcing precoder in the next stage. Notice that the channel is now a virtual MIMO broadcast channel. Hence, transmitter k computes the vector $\mathbf{H}^{-1}\mathbf{x}[1]$ and sends its kth component.

The received signal at time 3 is

$$\mathbf{y}[3] = \mathbf{H}(\mathbf{H}^{-1}\mathbf{x}[1]) = \mathbf{x}[1].$$

Meanwhile, the previously-received signals (at time 2) are fed back to all of the transmitters. Again each transmitter computes $\mathbf{H}^{-1}\mathbf{x}[2]$ and then sends its corresponding component. In time-slot 4, the received signal is

$$\mathbf{y}[4] = \mathbf{H}(\mathbf{H}^{-1}\mathbf{x}[2]) = \mathbf{x}[2].$$

Hence, during 4 time-slots, $2K$ symbols are transmitted. Therefore, fdof $= \frac{K}{2}$.

Remark 4: Observe that our scheme does not require channel diversity. However, it does not either require that the channels are static, i.e., if $\mathbf{H}$ varied across time-slots, then $\frac{K}{2}$ would still be achievable through the same scheme.

B. Converse

Our converse is based on the following lemma that bounds the rate for any couple of users $i, j \in [K]$.

Lemma 1: We have that $\forall\ i, j \in [K]$ and $i \neq j$,

$$\begin{aligned} R_i + R_j &\le h(Y_i) - h(Z_i) \\ &\quad + h(Y_j, Y_{\mathcal{S}}|X_i, X_{\mathcal{S}}, V_{ij}) - h(Z_j, Z_{\mathcal{S}}), \end{aligned}$$

where $\mathcal{S} = [K]\backslash\{i, j\}$, $Y_{\mathcal{S}} \triangleq \{Y_k\}, k \in \mathcal{S}$, and $V_{ij} \triangleq h_{ij}X_j + Z_i$.

Proof: See Appendix B. ∎

Notice that

$$\begin{aligned} &h(Y_i) - h(Z_i) \le \log|K^{\mathsf{G}}_{Y_i}| \\ \Rightarrow &h(Y_j, Y_{\mathcal{S}}|X_i, X_{\mathcal{S}}, V_{ij}) - h(Z_j, Z_{\mathcal{S}}) \le \log\frac{\left|K^{\mathsf{G}}_{(Y_j, Y_{\mathcal{S}}, X_i, X_{\mathcal{S}}, V_{ij})}\right|}{\left|K^{\mathsf{G}}_{X_i, X_{\mathcal{S}}, V_{ij}}\right|}, \end{aligned}$$

where $i, j \in [K]$, with $i \neq j$, and K^{G}_X indicates the covariance matrix of a Gaussian random vector X. Straightforward computation with Lemma 1 gives $R_i + R_j \le \log P + o(\log P)$, where $\lim_{P\to\infty} \frac{o(\log P)}{\log P} = 0$. Therefore, we get

$$\sum_{k=1}^{K} R_k \le \frac{K}{2} \cdot \log P + o(\log P). \tag{9}$$

This completes the proof.

Remark 5: Alternatively, we can use Theorem 1 in [15] to prove dof $\le \frac{K}{2}$. We instead derive a new outer bound of Lemma 1 to prove the converse. Note that Lemma 1 is not limited to the degrees-of-freedom.

V. Achievability Proof of Theorem 2

In Fig. 4 we sketch our achievability scheme. We will show that 6 symbols can be transmitted during 4 time slots. During time slot 1 and 2 the transmitters send the symbols $\mathbf{x}[1]$, $\mathbf{x}[2]$; the received signal at receiver $k \in [3]$ and time slot $n \in [2]$ is

$$Y_k[n] = h_{k1}X_1[n] + h_{k2}X_2[n] + h_{k3}X_3[n].$$

In time-slot 2, while the transmitters send the second batch of symbols, each receiver feeds back an appropriately *scaled* version of the previous channel output e.g., $Y_k[1]$. The fed back signal received at transmitter k is

$$\tilde{Y}_k[2] = h_{1k}Y_1[1] + h_{2k}(b \cdot Y_2[1]) + h_{3k}(c \cdot Y_3[1]). \tag{10}$$

We design the coding coefficients b and c so that the $X_2[1]$ component of the signal is cancelled out at transmitters 1 and 3. That way, transmitter 1 can decode transmitter 3's symbol and vice versa. Expanding (10), we get

$$\begin{aligned} \tilde{Y}_k[2] = h_{k1}&(h_{11}X_1[1] + \underline{h_{12}X_2[1]} + h_{13}X_3[1]) \\ + h_{k2}&b(\underline{h_{22}X_2[1]} + h_{21}X_1[1] + h_{23}X_3[1]) \\ + h_{k3}&c(h_{33}X_3[1] + h_{31}X_1[1] + \underline{h_{32}X_2[1]}). \end{aligned}$$

We first select coefficients b and c such that the underlined terms are canceled out at transmitters 1 and 2

$$\left\{ \begin{array}{l} h_{12}h_{22}b + h_{13}h_{32}c = -h_{11}h_{12} \\ h_{32}h_{22}b + h_{33}h_{32}c = -h_{31}h_{12} \end{array} \right\} \tag{11}$$

$$\Leftrightarrow \left\{ \begin{array}{l} c = \frac{h_{12}}{h_{32}} \cdot \frac{h_{32}h_{11} - h_{31}h_{12}}{h_{12}h_{33} - h_{32}h_{13}} \\ b = -\frac{h_{13}}{h_{22}} \cdot \left(\frac{h_{32}h_{11} - h_{31}h_{12}}{h_{12}h_{33} - h_{32}h_{13}}\right) - \frac{h_{11}}{h_{22}} \end{array} \right\}. \tag{12}$$

Since the coefficients h_{21} and h_{23} are not present in the above solutions for b and c, transmitters 1 and 3 receive as feedback a nonzero linear combination of $X_1[1]$ and $X_3[1]$, almost surely, when assuming channels drawn at random and i.i.d. from continuous distributions that do not have mass around zero. This is due to the fact that the multiplying components of $X_1[1]$ and $X_3[1]$ can be rescaled and viewed as non-zero polynomials on the entries of $\mathbf{H}$. Then, their non-zero property follows from the following lemma.

Lemma 2: Let n independent random variables $X_1, X_2, \ldots, X_n$ that are drawn from continuous distributions over $\mathbb{R}$ with density functions that do not have discontinuities around 0. Then, if $P(X_1, \ldots, X_n)$ is a non-identically zero polynomial of degree d we have that $\Pr\{P(X_1, \ldots, X_n) = 0\} = 0$.

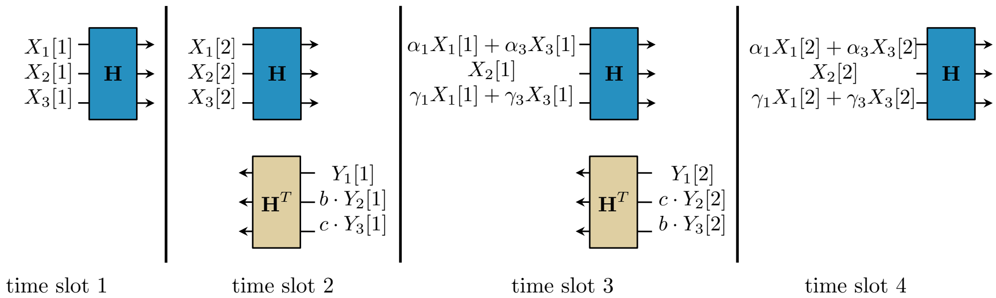

Fig. 2. Model II: Transmission Scheme for 3-user IC

Proof: The proof can be found in Appendix C. ∎

Then, assuming that transmitters 1 and 3 have perfect knowledge of the channel and the constants b and c, they can each decode the other user's symbols by simply subtracting a scaled version of $X_1[1]$ and $X_3[1]$ from their received equations. This extra knowledge is key to establishing alignment of interfering terms using an extra round of transmissions.

At time 3, transmitter 2 re-sends $X_2[1]$, while transmitter 1 sends $\alpha_1 X_1[1] + \alpha_3 X_3[1]$ and transmitter 3 sends $\gamma_1 X_1[1] + \gamma_3 X_3[1]$. These linear combinations are designed such that the resulting interference terms of $Y_k[3]$ align with the ones of $Y_k[1]$, while enforcing the useful symbols to be distinguishable. This way, each receiver will be able to decode the symbol of interest. The received signal at receiver 1 is

$$\begin{aligned} Y_1[3] &= (h_{11}\alpha_1 + h_{13}\gamma_1)X_1[1] \\ &\quad + h_{12}X_2[1] + (h_{11}\alpha_3 + h_{13}\gamma_3)X_3[1], \end{aligned}$$

and interference alignment can be achieved for α_3 and γ_3 satisfying $h_{11}\alpha_3 + h_{13}\gamma_3 = h_{13}$. At receiver 2 we have

$$\begin{aligned} Y_2[3] &= h_{22}X_2[1] \\ &\quad + (h_{21}\alpha_1 + h_{23}\gamma_1)X_1[1] + (h_{23}\gamma_3 + h_{21}\alpha_3)X_3[1]. \end{aligned}$$

The interference alignment condition for this receiver is $h_{21}\alpha_1 + h_{23}\gamma_1 = -h_{21}$ and $h_{23}\gamma_3 + h_{21}\alpha_3 = -h_{23}$. Accordingly, at receiver 3, we have

$$\begin{aligned} Y_3[3] &= (h_{33}\gamma_3 + h_{31}\alpha_3)X_3[1] \\ &\quad + (h_{31}\alpha_1 + h_{33}\gamma_1)X_1[1] + h_{32}X_2[1] \end{aligned}$$

and its alignment condition is $h_{31}\alpha_1 + h_{33}\gamma_1 = h_{31}$. Solving the 4 equations in the 4 unknowns $\alpha_1, \alpha_3, \gamma_1, \gamma_3$ yields

$$\left\{ \begin{array}{l} h_{21}\alpha_1 + h_{23}\gamma_1 = -h_{21} \\ h_{31}\alpha_1 + h_{33}\gamma_1 = h_{31} \end{array} \right\} \Leftrightarrow \left\{ \begin{array}{l} \alpha_1 = -\frac{h_{21}h_{33}+h_{23}h_{31}}{h_{21}h_{33}-h_{23}h_{31}} \\ \gamma_1 = \frac{2h_{21}h_{31}}{h_{21}h_{33}-h_{23}h_{31}} \end{array} \right\},$$

$$\left\{ \begin{array}{l} h_{11}\alpha_3 + h_{13}\gamma_3 = h_{13} \\ h_{21}\alpha_3 + h_{23}\gamma_3 = -h_{23} \end{array} \right\} \Leftrightarrow \left\{ \begin{array}{l} \alpha_3 = \frac{2h_{13}h_{23}}{h_{11}h_{23}-h_{13}h_{21}} \\ \gamma_3 = -\frac{h_{11}h_{23}+h_{13}h_{21}}{h_{11}h_{23}-h_{13}h_{21}} \end{array} \right\}.$$

The decoupling of the equations is possible due to the fact that user's 2 symbols where canceled out in the feedback phase. We now continue by rewriting the received signals as

$$\begin{aligned} Y_1[3] &= A \cdot X_1[1] + h_{12}X_2[1] + h_{13}X_3[1], \\ Y_2[3] &= h_{22}X_2[1] - h_{21}X_1[1] - h_{23}X_3[1], \\ Y_3[3] &= C \cdot X_3[1] + h_{31}X_1[1] + h_{32}X_2[1], \end{aligned} \tag{13}$$

where

$$\begin{aligned} A &= -h_{11}\frac{h_{21}h_{33} + h_{23}h_{31}}{h_{21}h_{33} - h_{23}h_{31}} + h_{13}\frac{2h_{21}h_{31}}{h_{21}h_{33} - h_{23}h_{31}}, \\ C &= -h_{33}\frac{h_{11}h_{23} + h_{13}h_{21}}{h_{11}h_{23} - h_{13}h_{21}} + h_{31}\frac{2h_{13}h_{23}}{h_{11}h_{23} - h_{13}h_{21}}. \end{aligned}$$

Using the two rounds of received signals at slots 1 and 3, each receiver can now solve for its symbol of interest

$$\begin{aligned} X_1[1] &= \frac{1}{A - h_{11}} \cdot (Y_1[3] - Y_1[1]), \\ X_2[1] &= \frac{1}{2h_{22}}(Y_2[3] + Y_2[2]), \\ X_3[1] &= \frac{1}{C - h_{33}}(Y_3[3] - Y_3[1]). \end{aligned} \tag{14}$$

The decodability requires

$$A - h_{11} \neq 0,\ 2h_{22} \neq 0,\ \text{ and } C - h_{33} \neq 0, \tag{15}$$

where each of the conditions can be restated as a nonzero polynomial in the channel coefficients. Due to Lemma 2, these polynomials are nonzero almost surely for channel elements drawn at random and i.i.d. from continuous distributions, therefore the conditions hold almost surely.

During time slot 3 the receivers send the corresponding feedback that allows, the decoding of the second batch of symbols $\mathbf{x}[2]$ in time-slot 4. Again this batch of symbols will be decodable almost surely. Hence, fdof $= \frac{3}{2}$ is achievable almost surely. This completes the proof.

Remark 6: We would like to note that the above achievability scheme works even if the channel coefficients are varying across time-slots.

VI. Model II: Beyond the Three-user Case

A. A Heuristic Algorithm

We continue our study on the $K > 3$ case, for which we develop a heuristic algorithm that finds coding coefficients for the feedback strategy. Similar to the 3-user case, we consider the 4 time transmission scheme, where at each round we assume that the transmitting end can combine and scale all the information that it has available. Let at the first round that the transmitters send the K-length vector $\mathbf{x}[1]$. Then, the received signal at time slot 1 is $\mathbf{y}[1] = \mathbf{H}\mathbf{x}[1]$. The receivers rescale the channel output and feed it back to the transmitter, which receives $\hat{\mathbf{y}}[2] = \mathbf{H}^T\mathbf{D}_1\mathbf{H}\mathbf{x}[1]$, where $\mathbf{D}_1$ indicates a $K \times K$ diagonal matrix. At timeslot 3, the transmitters combine $\mathbf{x}[1]$ with $\hat{\mathbf{y}}[2]$ and $\mathbf{x}[2]$, and the recieved vector is $\mathbf{y}[3] = \mathbf{H}\mathbf{D}_2\mathbf{H}^T\mathbf{D}_1\mathbf{H}\mathbf{x}[1] + \mathbf{H}\mathbf{D}_3\mathbf{x}[1]$, where $\mathbf{D}_i$, $i \in \{1,2,3\}$ are all $K \times K$ diagonal matrices. Hence, a sufficient condition for decodability of the first batch of symbols is the following

$$\mathbf{D}_4\mathbf{y}[1] + \mathbf{y}[3] = \mathbf{D}_5\mathbf{x}[1], \tag{16}$$

for some selection of diagonal combining matrices. We can rewrite (16) as a set of K^2 equations in the unknown diagonals

$$\mathbf{D}_4\mathbf{H} + \mathbf{H}\mathbf{D}_2\mathbf{H}^T\mathbf{D}_1\mathbf{H} + \mathbf{H}\mathbf{D}_3 = \mathbf{D}_5.$$

It is not hard to show that the total number of free variables in the above multilinear system of equations is equal to $5K-2$. We expect that this system should be feasible when the number of unknowns is at most equal to the number of equations, which is true for $K \leq 4$. For the $K = 3$ case, we obtained a closed form solution solving linear equations. However, for the $K = 4$ case, solving for the unknown diagonals requires solving intricate multilinear equations, which we cannot do efficiently.

To tackle this intrinsic nonlinearity of finding fdof achieving diagonals, we first reformulate the problem as a nonconvex optimization, which we subsequently relax to a convex one. Ideally, we would like to solve the the following optimization

$$\begin{aligned} &\min_{\mathbf{D}_i} \left\| \mathbf{D}_4\mathbf{H} + \mathbf{H}\mathbf{D}_2\mathbf{H}^T\mathbf{D}_1\mathbf{H} + \mathbf{H}\mathbf{D}_3 - \mathbf{D}_5 \right\|_F \\ &\text{s.t. } \operatorname{rank}(\mathbf{D}_5) = K, \end{aligned}$$

where the rank constraint enforces that all K symbols are distinguishable (i.e. nonzero) at the receivers. Due to the rank constraints and the nonconvex cost function the above optimization does not seem tractable. To solve it, we relax it to an alternating convex optimization problem, where we replace the nonconvex rank constraints with affine ones. Let $g(\{\mathbf{D}_i\}) = \left\| \mathbf{D}_4\mathbf{H} + \mathbf{H}\mathbf{D}_2\mathbf{H}^T\mathbf{D}_1\mathbf{H} + \mathbf{H}\mathbf{D}_3 - \mathbf{D}_5 \right\|_{\mathrm{F}}$, where $\|\mathbf{A}\|_{\mathrm{F}}$ indicates the Frobenius norm of matrix $\mathbf{A}$. Then, we use the following (randomized) heuristic.

$\mathcal{A}[n]$:

draw at random $\mathbf{D}_1$
for N iterations or until $g(\{\mathbf{D}_i\}) < \epsilon$
$\{\mathbf{D}_i\}_{i\neq 3} \leftarrow \min_{\mathbf{D}_1,\mathbf{D}_2,\mathbf{D}_4,\mathbf{D}_5} g(\{\mathbf{D}_i\})$ s.t.: $[\mathbf{D}_5]_{1,1} = 1$
$\{\mathbf{D}_i\}_{i\neq 2} \leftarrow \min_{\mathbf{D}_1,\mathbf{D}_3,\mathbf{D}_4,\mathbf{D}_5} g(\{\mathbf{D}_i\})$ s.t.: $[\mathbf{D}_5]_{1,1} = 1$
if ($n > N_1$ and $g(\{\mathbf{D}_i\}) > \epsilon_1$) OR
($n > N_2$ and $g(\{\mathbf{D}_i\}) > \epsilon_2$) re-draw $\mathbf{D}_1$

We simulate the above heuristic for the 4-user case. We set $\epsilon = 10^{-7}$, $N = 10^4$, $N_1 = 30$, $\epsilon_1 = 0.1$, $N_2 = 200$ and $\epsilon_2 = 10^{-4}$. We generate 100 channel realizations. We assume channel coefficients are fixed and are drawn i.i.d. zero-mean Gaussian with variance 1. All simulation results show that the heuristic convergence to the optimal solution, thus ensuring $\frac{1}{2}$ signal dimension per user. On the other hand, for $K > 4$, simulations demonstrate the divergence of the algorithm for all of the 100 experiments. The difficulty comes from the fact that this heuristic formulation approximately solves $\mathcal{O}(K^2)$ multilinear equations in $\mathcal{O}(K)$ unknowns. In general, we would not expect such an approach to work for arbitrary K, unless coding coefficients were carefully designed to induce significant dependencies among these equations.

B. A Spectral Family of K-user ICs with non-trivial fdof

Here we present a class of K-user ICs where nontrivial fdof can be achieved in a finite number of symbol extensions. Let

$$\mathbf{H} = \mathbf{H}^T \text{ and } \mathbf{H} \text{ has } L \text{ distinct eigenvalues.} \tag{17}$$

We refer to our achievable scheme as the *ping pong* scheme which works as follows. Our achievable scheme uses $2L$ number of time-slots to transmit $2K$ symbols. The two batches of symbols along with their corresponding fedback symbols are pipelined as in the previous 4-time-slot achievability schemes. At time-slot 1 the transmitters send $\mathbf{x}[1]$, while at time-slot 2 they send $\mathbf{x}[2]$, and they receive as feedback $\hat{\mathbf{y}}[2] = \mathbf{H}^T\mathbf{H}\mathbf{x}[1] = \mathbf{H}^2\mathbf{x}[1]$. Then, at time slot $n \in \{2, \ldots, 2L\}$ the received signals at the receiver side is

$$\mathbf{y}[n] = \mathbf{H}\hat{\mathbf{y}}[n-1] \text{ and } \hat{\mathbf{y}}[n] = \mathbf{H}\mathbf{y}[n-1]. \tag{18}$$

This means that after $2L$ timesolts the receiver end will have $\{\mathbf{H}\mathbf{x}[1], \mathbf{H}^3\mathbf{x}[1], \ldots, \mathbf{H}^{2L-1}\mathbf{x}[1]\}$ and $\{\mathbf{H}\mathbf{x}[2], \mathbf{H}^3\mathbf{x}[2], \ldots, \mathbf{H}^{2L-1}\mathbf{x}[2]\}$. Using the eigenvalue multiplicity, one can easily show that there exist constants $c_1, \ldots, c_L$ such that $\sum_{i=1}^{L} c_i\mathbf{H}^{2i-1}\mathbf{x}[1] = \mathbf{x}[1]$ and $\sum_{i=1}^{L} c_i\mathbf{H}^{2i-1}\mathbf{x}[2] = \mathbf{x}[2]$ since

$$\begin{aligned} &\sum_{i=1}^{L} c_i \left(\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^H\right)^{2i-1} = \mathbf{I}_k \Leftrightarrow \sum_{i=1}^{L} c_i\mathbf{Q}\mathbf{\Lambda}^{2i-1}\mathbf{Q}^H = \mathbf{I}_K \\ &\Leftrightarrow \sum_{i=1}^{L} c_i\mathbf{\Lambda}^{2i-1} = \mathbf{I}_K \Leftrightarrow \sum_{i=1}^{L} c_i\lambda_j^{2i-1} = 1, \forall j \in [L] \end{aligned}$$

and $\sum_{i=1}^{L} c_i\lambda_j^{2i-1} = 1, \forall j \in [L]$ is a (Vandermonde) system of L equations in L unknowns, which admits a solution. Finding the c_i coefficients is therefore sufficient for decoding the $2K$ transmitted symbols in $2L$ symbol extensions, thus achieving fdof $= \frac{K}{L}$.

VII. Conclusion

We have shown that the degrees-of-freedom is $\frac{K}{2}$ for the scalar K-user Gaussian IC with feedback where each receiver feeds back it received signal to all of the transmitters. Consequently we showed that feedback can increase degrees-of-freedom (a) for some class of ICs; and (b) for arbitrary scalar ICs assuming linear transceivers. We next explored a more realistic feedback scenario where feedback is provided through the backward IC. For the 3-user case, we developed a simple feedback strategy that achieves fdof $= \frac{3}{2}$.

Our future work is along several new directions: (1) Extending to an arbitrary number K of users under a realistic feedback model (Model II); (2) Extending to MIMO channels; (3) Exploring Model I further with the finer capacity notion.

Acknowledgment

We gratefully acknowledge Dr. Viveck R. Cadambe for pointing out [15], which can provide an alternative converse proof of Theorem 1.

APPENDIX A
PROOF OF LEMMA 1

Proof: By symmetry, it suffices to prove the bound of R_1+R_2. Let $W_{\mathcal{S}} \triangleq \{W_3, \cdots, W_K\}$ and $Y_{\mathcal{S}} \triangleq \{Y_3, \cdots, Y_K\}$.

Starting with Fano's inequality, we get:

$$
\begin{aligned}
&N(R_1 + R_2 - \epsilon_N)\\
&\le I(W_1; Y_1^N, W_2, W_{\mathcal{S}}) + I(W_2; Y_2^N, W_{\mathcal{S}})\\
&\overset{(a)}{=} I(W_1; Y_1^N|W_2, W_{\mathcal{S}}) + I(W_2; Y_2^N|W_{\mathcal{S}})\\
&\overset{(b)}{=} h(Y_1^N|W_2, W_{\mathcal{S}}) - h(Y_1^N|W_1, W_2, W_{\mathcal{S}}) + h(Y_2^N|W_{\mathcal{S}})\\
&- \left[h(Y_1^N, Y_2^N, Y_{\mathcal{S}}^N|W_2, W_{\mathcal{S}}) - h(Y_1^N, Y_{\mathcal{S}}^N|W_2, W_{\mathcal{S}}, Y_2^N)\right]\\
&= h(Y_2^N|W_{\mathcal{S}}) + h(Y_1^N, Y_{\mathcal{S}}^N|W_2, W_{\mathcal{S}}, Y_2^N)\\
&- h(Y_1^N|W_1, W_2, W_{\mathcal{S}}) - h(Y_2^N, Y_{\mathcal{S}}^N|Y_1^N, W_2, W_{\mathcal{S}})\\
&\overset{(c)}{\le} h(Y_2^N|W_{\mathcal{S}}) + h(Y_1^N, Y_{\mathcal{S}}^N|W_2, W_{\mathcal{S}}, Y_2^N)\\
&- \sum_n \sum_k h(Z_k[n])\\
&\overset{(d)}{\le} \sum h(Y_2[n]) - \sum_n \sum_k h(Z_k[n])\\
&+ \sum h(Y_1[n], Y_{\mathcal{S}}[n]|X_2[n], X_{\mathcal{S}}[n], V_{21}[n])
\end{aligned}
$$

where (a) follows from the independence of $(W_1, \cdots, W_K)$; (b) follows from a chain rule; (c) follows from Claim 1 (see below); (d) follow from the fact that conditioning reduces entropy, and the fact that $X_k[n]$ is a function of $(W_k, Y_1^{n-1}, Y_2^{n-1}, Y_{\mathcal{S}}^{n-1})$, and the fact that $V_{21}[n] \triangleq h_{21}X_1[n] + Z_2[n]$. If $(R_1, \cdots, R_K)$ is achievable, then $\epsilon \to 0$ as N tends to infinity. Therefore, we get

$$
\begin{aligned}
R_2 + R_1 \le\ & h(Y_2) - h(Z_2)\\
&+ h(Y_1, Y_{\mathcal{S}}|X_2, X_{\mathcal{S}}, V_{21}) - h(Z_1, Z_{\mathcal{S}}).
\end{aligned}
$$

■

This complete the proof.

Claim 1:

$$
\begin{aligned}
&h(Y_1^N|W_1, W_2, W_{\mathcal{S}}) \ge \sum h(Z_1[n])\\
&h(Y_2^N, Y_{\mathcal{S}}^N|Y_1^N, W_2, W_{\mathcal{S}}) \ge \sum_n \left[h(Z_2[n]) + h(Z_{\mathcal{S}}[n])\right]
\end{aligned}
$$

Proof: First consider

$$
\begin{aligned}
&h(Y_1^N|W_1, W_2, W_{\mathcal{S}})\\
&\ge h(Y_1^N|W_1, W_2, W_{\mathcal{S}}, Y_2^N, Y_{\mathcal{S}}^N)\\
&\overset{(a)}{=} \sum h(Y_1[n]|W_1, W_2, W_{\mathcal{S}}, Y_2^N, Y_{\mathcal{S}}^N, Y_1^{n-1}, X_1^n, X_2^n, X_{\mathcal{S}}^n)\\
&\overset{(b)}{=} \sum h(Z_1[n])
\end{aligned}
$$

where (a) follows from the fact that $X_k[n]$ is a function of $(W_k, Y_1^{n-1}, Y_2^{n-1}, Y_{\mathcal{S}}^{n-1})$ and (b) follows from the fact that channel is memoryless. ■

Next consider

$$
\begin{aligned}
&h(Y_2^N, Y_{\mathcal{S}}^N|Y_1^N, W_2, W_{\mathcal{S}})\\
&\ge h(Y_2^N, Y_{\mathcal{S}}^N|Y_1^N, W_2, W_{\mathcal{S}}, W_1)\\
&\overset{(a)}{\ge} \sum h(Z_2[n], Z_{\mathcal{S}}[n])
\end{aligned}
$$

where (a) follows from the fact that $X_k[n]$ is a function of $(W_k, Y_1^{n-1}, Y_2^{n-1}, Y_{\mathcal{S}}^{n-1})$ and the fact that channel is memoryless.

APPENDIX B
PROOF OF $\mathsf{dof} = 1$

Starting with Fano's inequality and data processing inequality [16], we get:

$$
\begin{aligned}
&N(R_1 + R_2 + R_3 - \epsilon_N)\\
&\le I(X_1^N; Y_1^N) + I(X_2^N; Y_2^N) + I(X_2^N; Y_2^N)\\
&\overset{(a)}{\le} I(X_1^N; Y_1^N) + I(X_2^N; Y_2^N|X_1^N)\\
&\quad + I(X_3^N; Y_3^N|X_1^N, X_2^N)\\
&= h(Y_1^N) - h(Z_3^N) + [-h(Y_1^N|X_1^N) + h(Y_2^N|X_1^N)]\\
&\quad + \left\{-h(Y_2^N|X_1^N, X_2^N) + h(Y_3^N|X_1^N, X_2^N)\right\}\\
&\overset{(b)}{\le} \sum \{h(Y_1[n]) - h(Z_3[n])\}\\
&\quad + \left\{-h(-X_2^N + X_3^N + Z_1^N) + h(X_2^N - X_3^N + Z_2^N)\right\}\\
&\quad + \left\{-h(-X_3^N + Z_2^N) + h(X_3^N + Z_3^N)\right\}\\
&\overset{(c)}{=} \sum \{h(Y_1[n]) - h(Z_3[n])\}\\
&\quad + \left\{-h(X_2^N - X_3^N - Z_1^N) + h(X_2^N - X_3^N + Z_2^N)\right\}\\
&\quad + \left\{-h(X_3^N - Z_2^N) + h(X_3^N + Z_3^N)\right\}\\
&\overset{(d)}{=} \sum \{h(Y_1[n]) - h(Z_3[n])\}\\
&\le N\{\log P + o(\log P)\}
\end{aligned}
$$

where (a) follows from the fact that adding information increases mutual information and the fact that (X_1^N, X_2^N, X_3^N) are independent; (b) follows from the fact that conditioning reduces entropy and the fact that noise is i.i.d; (c) follows from the fact that $h(X) = h(-X)$; and (d) follows from the fact that Z_i^N and $-Z_j^N$ are identically distributed $\forall i, j$. Therefore, the $\mathsf{dof} = 1$.

APPENDIX C
PROOF OF LEMMA 2

Proof:

The proof is by induction.

For $n = 1$ it is easy to see that a polynomial on a single variable X_i can have at most d roots, hence the probability that it evalueate to zero when we draw X_i is zero. The induction hypothesis is that the lemma holds for any non-zero polynomial on $n-1$ variables $X_1, \ldots, X_{n-1}$ that are drawn from distributions where 0 occurs with probability 0.

We assume that we can decompose $P(X_1, \ldots, X_n)$ in the following form

$$P(X_1, \ldots, X_n) = \sum_{i=0}^{d} X_n^i P_i(X_1, \ldots, X_{n-1}) \tag{19}$$

where the degree of each polynomial P_i is at most $d - i$, since we assumed that P has degree d, and *at least* one P_i has to be nonzero. Let without loss of generality assume that P_j is nonzero, for some $j \in \{0, \ldots, d\}$. If we randomly pick $X_1 = x_1, \ldots, X_{n-1} = x_{n-1}$, then by the induction hypothesis we have

$$\Pr\{P_j(X_1, \ldots, X_{n-1}) = 0\} = 0. \tag{20}$$

If $P_j(x_1, \ldots, x_{n-1}) \neq 0$ then $P(x_1, \ldots, x_{n-1}, X_N)$ is of degree i. Hence,

$$\Pr\{P(x_1, \ldots, x_{n-1}, X_N) = 0 | P_i(x_1, \ldots, x_{n-1}) \neq 0\} = 0. \tag{21}$$

Then, we have

$$\begin{aligned}
&\Pr\{P(X_1, \ldots, X_{n-1}, X_N) = 0\} \\
=& \Pr\{P(x_1, \ldots, x_{n-1}, X_N) = 0 | P_i(x_1, \ldots, x_{n-1}) \neq 0\} \\
&\cdot \Pr\{P_i(x_1, \ldots, x_{n-1}) \neq 0\} \\
&+ \Pr\{P(x_1, \ldots, x_{n-1}, X_N) = 0 | P_i(x_1, \ldots, x_{n-1}) = 0\} \\
&\cdot \Pr\{P_i(x_1, \ldots, x_{n-1}) = 0\} \\
&\leq \Pr\{P(x_1, \ldots, x_{n-1}, X_N) = 0 | P_i(x_1, \ldots, x_{n-1}) \neq 0\} \\
&+ \Pr\{P_i(x_1, \ldots, x_{n-1}) = 0\} \\
&= 0,
\end{aligned} \tag{22}$$

which completes the proof. ■

References

[1] C. E. Shannon, "The zero error capacity of a noisy channel," *IRE Transactions on Information Theory*, Sept. 1956.

[2] N. T. Gaarder and J. K. Wolf, "The capacity region of a multiple-access discrete memoryless channel can increase with feedback," *IEEE Transactions on Information Theory*, Jan. 1975.

[3] L. H. Ozarow, "The capacity of the white Gaussian multiple access channel with feedback," *IEEE Transactions on Information Theory*, July 1984.

[4] G. Kramer, "Feedback strategies for white Gaussian interference networks," *IEEE Transactions on Information Theory*, vol. 48, pp. 1423–1438, June 2002.

[5] G. Kramer, "Correction to "Feedback strategies for white Gaussian interference networks", and a capacity theorem for Gaussian interference channels with feedback," *IEEE Transactions on Information Theory*, vol. 50, June 2004.

[6] C. Suh and D. Tse, "Feedback capacity of the Gaussian interference channel to within 2 bits," *IEEE Transactions on Information Theory*, vol. 57, pp. 2667–2685, May 2011.

[7] A. Vahid, C. Suh, and S. Avestimehr, "Interference channels with rate-limited feedback," *IEEE Transactions on Information Theory (to appear)*, 2012.

[8] V. R. Cadambe and S. A. Jafar, "Interference alignment and the degrees of freedom for the K user interference channel," *IEEE Transactions on Information Theory*, vol. 54, pp. 3425–3441, Aug. 2008.

[9] R. Etkin and E. Ordentlich, "The degrees-of-freedom of the K-user Gaussian interference channel is discontinuous at rational channel coefficients," *IEEE Transactions on Information Theory*, vol. 55, pp. 4932–4946, Nov. 2009.

[10] A. S. Motahari, S. O. Gharan, M. A. Maddah-Ali, and A. K. Khandani, "Real interference alignment," *submitted to the IEEE Transactions on Information Theory (arXiv:1001.3403)*, Jan. 2010.

[11] C. M. Yetis, T. Gou, S. A. Jafar, and A. H. Kayran, "On feasibility of interference alignment in MIMO interference networks," *IEEE Transactions on Signal Processing*, vol. 58, pp. 4771–4782, Sept. 2010.

[12] M. Maddah-Ali, A. Motahari, and A. Khandani, "Communication over MIMO X channels: Interference alignment, decomposition, and performance analysis," *IEEE Transactions on Information Theory*, vol. 54, pp. 3457–3470, Aug. 2008.

[13] R. Tandon, S. Mohajer, and H. V. Poor *submitted to the IEEE Transactions on Information Theory (arXiv:1109.1507)*.

[14] S. Mohajer, R. Tandon, and H. V. Poor *submitted to the IEEE Transactions on Information Theory (arXiv:1110.6487)*.

[15] V. R. Cadambe and S. A. Jafar, "Degrees of freedom of wireless networks with relays, feedback, cooperation, and full duplex operation," *IEEE Transactions on Information Theory*, vol. 55, pp. 2334–2344, May 2009.

[16] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* New York Wiley, 2th ed., July 2006.