On the Feedback Capacity of the Fully Connected K-User Interference Channel

Article *in* IEEE Transactions on Information Theory · October 2011

DOI: 10.1109/TIT.2012.2235524 · Source: arXiv

3 authors:

Soheil Mohajer
University of California, Berkeley

Ravi Tandon
The University of Arizona

H. Vincent Poor
Princeton University

On the Feedback Capacity of the Fully Connected K-User Interference Channel

Soheil Mohajer, Ravi Tandon, and H. Vincent Poor
Department of Electrical Engineering
Princeton University
Princeton, NJ 08544, USA
E-mail: {smohajer, rtandon, poor}@princeton.edu

Abstract

The symmetric K user interference channel with fully connected topology is considered, in which (a) each receiver suffers interference from all other $K-1$ transmitters, and (b) each transmitter has causal and noiseless feedback from its respective receiver. The number of generalized degrees of freedom (DoF) is characterized in terms of α, where the interference-to-noise ratio (INR) is given by $\mathsf{INR} = \mathsf{SNR}^{\alpha}$. It is shown that the per-user DoF of this network is the same as that of the 2-user interference channel with feedback, except for $\alpha = 1$, for which existence of feedback does not help in terms of DoF. The coding scheme proposed for this network, termed cooperative interference alignment, is based on two key ingredients, namely, interference alignment and interference decoding. Moreover, an approximate characterization is provided for the symmetric feedback capacity of the network, when the SNR and INR are far apart from each other.

I. Introduction

Wireless networks with multiple pairs of transceivers are quite common in modern communications, notable examples being wireless local area networks (WLANs) and cellular networks. Multiple independent flows of information share a common medium in such multiple unicast wireless networks. The broadcast and superposition nature of the wireless medium introduces complex signal interactions between multiple competing flows. In contrast to the point-to-point wireless channel, where a noisy version of a single transmitted signal is received at a given receiver, a combination of various wireless signals are observed at receivers in multiple unicast systems. In such scenarios, each decoder has to to deal with all interfering signals in order to decode its intended message. Managing such interfering signals in a multi-user network is a long standing and fundamental problem in wireless communication.

The work was supported in part by the Air Force Office of Scientific Research under MURI Grant FA9550-09-1-0643, and in part by the DTRA under Grant HDTRA-07-1-0037.

The simplest example in this category is the 2-user interference channel [1], in which two transmitters with independent messages are attempting to communicate with their respective receivers over the wireless transmission medium. Even for this simple 2-user network, the complete information-theoretic characterization has been open for several decades. To study more general networks, there is a clear need for a deep understanding and perhaps develop novel interference management techniques.

Although the exact characterization of the capacity region of the 2-user Gaussian interference channel is still unknown, several inner and outer bounds are known. These bounds are very useful in the sense of providing an approximate characterization when there exists a guarantee on the gap between them. This approach has resulted in an approximate characterization, within one bit, by Etkin, Tse, and Wang in [2] as well as Telatar and Tse in [3]. This characterization includes upper bounds for the capacity of the network, as well as encoding/decoding strategies based on Han-Kobayashi scheme [1], which perform close to optimal. Moreover, it has been shown that the gap between the fundamental information-theoretic bounds and what can be achieved using the proposed schemes is provably small. Therefore, the capacity can be approximated within a narrow range, although the exact region is still unknown.

A similar approximate characterization (with a larger gap) for this problem is developed in [4], in which both coding scheme and bounding techniques are devised by studying the problem under the *deterministic* model. This framework, introduced by Avestimehr, Diggavi, and Tse in [5], focuses on complex signal interactions in a wireless network by ignoring the randomness of the noise. Recently, it has been successfully applied to several problems, providing valuable insights for the more practically relevant Gaussian problems.

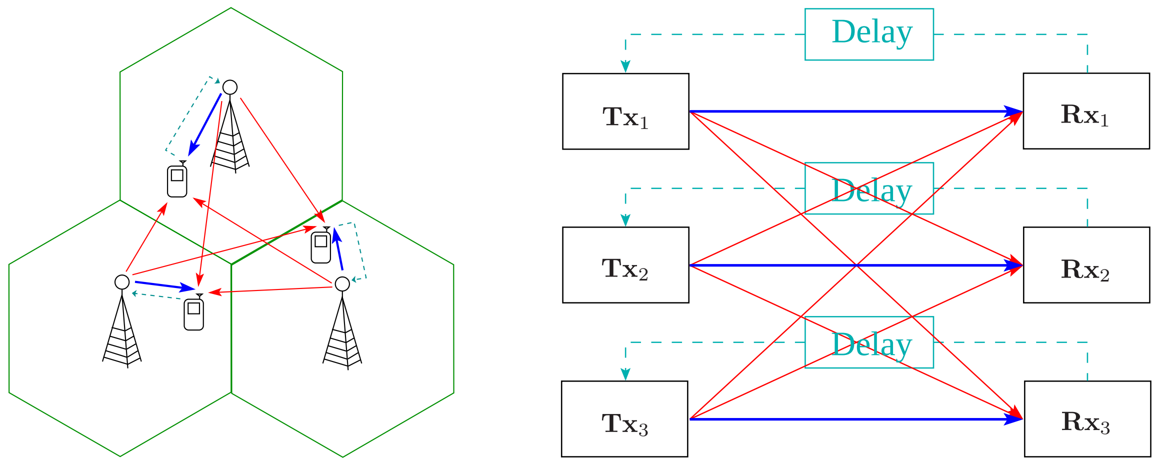
Several interference management techniques have been proposed for operating over more complex interference networks. Completely or partially decoding and removing interference (interference suppression) when it is strong and treating it as noise when it is weak are perhaps the most widely used schemes. More sophisticated schemes such as interference alignment [6], [7], and interference neutralization [8], [9] have been proposed recently. However, it still remains to be seen whether the capacity of general interference networks can be achieved with any combination of these techniques.

It has been shown that feedback does not increase the capacity of point-to-point discrete memoryless channels [10]. However, feedback is beneficial in improving the capacity regions of more complex networks (see [11] and references therein). The effects of feedback on the capacity region of the interference channel have been studied in several papers. Feedback coding schemes for K-user Gaussian interference networks have been developed by Kramer in [12]. Outer bounds for the 2-user interference channel with generalized feedback have been derived in [13] and [14]. The entire feedback capacity region of the 2 user Gaussian interference channel has been characterized within a 2 bit gap by Suh and Tse in [15]. Perhaps, the most interesting part of the result of [15] is the multiplicative gain provided by feedback at high signal-to-noise ratio (SNR). The gap between the capacity of the channel with and without feedback can be arbitrarily large for certain channel parameters. The key technique here is to use the feedback links to create an artificial path from each transmitter to its respective receiver through the other nodes in the network. For instance, the message intended for $\mathbf{Rx}_1$, can be sent either through the direct link $\mathbf{Tx}_1 \rightarrow \mathbf{Rx}_1$, or the cyclic path $\mathbf{Tx}_1 \to \mathbf{Rx}_2 \to \mathbf{Tx}_2 \to \mathbf{Rx}_1$. In particular, the advantage of such artificial paths can be clearly understood when the cross links are much stronger than that the direct links (e.g., the strong interference regime). This observation becomes very natural by studying the problem under the deterministic framework.

The first extension of [15] to a multi-user setting is the K-user cyclic interference channel with feedback, where each receiver's signal is interfered with only one of its neighboring transmitters, in a cyclic fashion. The effect of feedback on the capacity region of this network is addressed in [16]. It is shown that although feedback improves the symmetric capacity of the K-user interference channel, the improvement in symmetric capacity per user vanishes as K grows. The intuitive reason behind this result is that the configuration of the network allows only one cyclic path, which has to be shared between all pair of transceivers. The amount of information that can be conveyed through this path does not scale with K, and therefore the gain for each user scales inverse linearly with K.

In another extreme, each transmit signal may be corrupted by all the other signals transmitted by the other base stations. This model is appropriate for a network with densely located nodes, where everyone hears everyone else. This network, which we call *the fully connected K-user interference channel* (FC-IC), is another generalization of the 2-user interference channel. Fig. I shows the fully connected IC with feedback for $K = 3$ users. In this paper, we study the FC-IC network with feedback, and for simplicity, we consider a symmetric network topology, where all the direct links (from each transmitter to its respective receiver) have the same gain, and similarly, the gain of all cross (interfering) links are identical. The same problem without feedback has been studied by Jafar and Vishwanath in [17], where the number of symmetric degrees of freedom is characterized. In this paper, the impact of feedback is studied for the K-user FC-IC. The main contribution of this paper is to show that feedback can arbitrarily improve the performance of the network, and in contrast to the cyclic case [16], it *does scale* with the number of users in the systems. In particular, except for the intermediate interference regime where the signal-to-noise ratio is equal to the interference-to-noise ratio ($\mathsf{SNR} = \mathsf{INR}$), the effect of interference from $K - 1$ users is as if there were only one interfering transmitter in the network. This is analogous to the result of [7], where it is shown that the number of per-user degrees of freedom of the K-user fading interference channel, is the same as if there were only 2 users in the network.

In order to get the maximal benefit of feedback, we propose a novel encoding scheme, called cooperative interference alignment, which combines two well-known interference management techniques, namely, interference alignment and interference decoding. More precisely, the encoding at the transmitters is such that all the interfering signals are aligned at each receiver. However, a fundamental difference between our approach and the standard interference alignment approach is that we need to decode interference to be able to to remove it from the received signal, while the aligned interference is usually suppressed in standard approaches. A challenge here, which makes this problem fundamentally different from the 2-user inference channel, is that the interference is a combination of $(K - 1)$ interfering messages, and decoding all of them induces strict bounds on the rate of the interfering messages. However, each transmitter does not need to decode all the interfering messages individually, instead, upon receiving feedback, it only decodes the combination of them that corrupts the intended signal is of interest. To this end, we propose using a common structured code, which has the property that the summation of codewords



(a) A cellular interference network.

(b) Interference network with feedback.

Fig. 1. A cellular network with three base stations and three clients in (a), simplified and modelled as the network in (b).

of different users is still another codeword from the same codebook. Lattice codes [18] are a suitable choice to satisfy this desired property. This idea is similar to that used in [19] and [20].

The rest of this paper is organized as follows. First, we formally present the model, introduce notation, and state the problem in Section II. The main result of the paper is presented in Section III. Before proving the result for the Gaussian network, we study the problem under the deterministic model in Section IV, where we characterize the exact feedback capacity of the deterministic network. Based on the insight and intuition obtained by analysis of the deterministic network, we present the converse proof and the coding scheme for the Gaussian network in Sections V and VI, respectively. Having the approximate feedback capacity of the network, we derive the generalized degrees of freedom with feedback in Section VII, and finally, conclude the paper in Section VIII. In order to make the paper easily readable, some of the technical proofs are postponed to the appendices.

II. Problem Statement

In this work we consider a network with K pairs of transmitter/receivers. Each transmitter $\mathbf{Tx}_k$ has a message W_k that it wishes to send to its respective receiver $\mathbf{Rx}_k$. The signal transmitted by each transmitter is corrupted by the interfering signals sent by other transmitters, and received at the receiver. This can be mathematically modelled as

$$Y_k(t) = \sqrt{\mathsf{SNR}} X_k(t) + \sum_{\substack{i=1\\ i\neq k}}^{K} \sqrt{\mathsf{INR}} X_i(t) + Z_k(t), \tag{1}$$

where X_k and Y_k are the signals transmitted and received by $\mathbf{Tx}_k$ and $\mathbf{Rx}_k$, respectively, and $Z_k \sim \mathcal{N}(0,1)$ is an additive white Gaussian noise. All transmitting powers are constrained to 1, i.e., $\mathbb{E}[X_k^2] \leq 1$, for $k = 1, \ldots, K$. We assume a symmetric network, where all the cross links have the same gain (INR), and the gains of the all the direct link (SNR) are identical.

There is a perfect feedback link from each receiver to its respective transmitter. Hence, at each time instance, each transmitter generates each transmitting signal based on its own message as well as the output sequence observed at its receiver over the past time instances, i.e.,

$$X_{kt} = f_{kt}(W_k, Y_{k1}, Y_{k2}, \dots, Y_{k(t-1)}) = f_{kt}(W_k, Y_k^{t-1}), \tag{2}$$

where we use shorthand notation $Y_k^{t-1} = (Y_{k1}, Y_{k2}, \dots, Y_{k(t-1)})$ to indicate the output sequence observed at $\mathbf{Rx}_k$ up to time $t-1$.

A rate tuple $(R_1, R_2, \dots, R_K)$ is called achievable if there exists a family of codebooks with block length T with proper power and corresponding encoding/decoding functions such that the average decoding error probability tends to zero for all users as T increases. We denote the set of all achievable rate tuples by $\mathscr{R}$. In the high signal to noise ratio regime, the performance of wireless networks is measured in terms of the number of degrees of freedom, that is the pre-log factor in the expression of the capacity in terms of SNR. We consider the generalized degrees of freedom (GDoF) for this network in the presence of feedback. Since the problem is parametrized in terms of two growing factors, namely SNR and INR, we use the standard parameter α (as in [2] and [17]) to capture the growth rate of INR in terms of SNR. More formally, we define

$$\alpha = \frac{\log \mathsf{INR}}{\log \mathsf{SNR}}, \tag{3}$$

and the *per-user* generalized degrees of freedom as

$$d(\alpha) = \frac{1}{K} \limsup_{\mathsf{SNR}\to\infty} \frac{\max_{(R,\dots,R)\in\mathscr{R}} \sum_{k=1}^{K} R_k(\mathsf{SNR}, \alpha)}{\frac{1}{2}\log \mathsf{SNR}}. \tag{4}$$

It is worth mentioning that the half factor appears in the denominator since we are dealing with real signals. Our primary goal is to characterize the generalized degrees of freedom of the K-user interference channel with output feedback.

As mentioned earlier, the GDoF characterizes the performance of the network in the asymptotic SNR regime. However, in order to study practical networks, capacity is a more accurate measure to capture the performance. In order to consider such a high resolution analysis, we define the symmetric capacity of the network, that is

$$R_{\text{sym}} = \max_{(R,\dots,R)\in\mathscr{R}} R.$$

In this work we are interested in characterizing R_{sym} for the K-user interference channel with feedback. Although finding the exact symmetric capacity is extremely difficult, we make progress on this problem, and approximately characterize the capacity when the SNR and INR are not close to each other, that is when α (defined in (3)) is not equal to 1. To this end, we derive outer bounds and propose coding schemes for the network, and show that the gap between the achievable rate and the outer bound is a function *only* of K, the number of users in the network, and is independent of SNR and INR.

III. Main Results

In this section we present the main results of this paper. The first theorem characterizes the generalized degrees of freedom of the K-user FC-IC with feedback.

Theorem 1. *For the K-user fully connected interference channel (FC-IC) with output feedback, the per-user* GDoF *is given by*

$$d_{\rm FB}(\alpha) = \begin{cases} 1-\frac{\alpha}{2} & \alpha<1 \textit{ (weak interference)} \\ \frac{1}{K} & \alpha=1 \\ \frac{\alpha}{2} & \alpha>1 \textit{ (strong interference)} \end{cases} \tag{5}$$

In order to demonstrate the benefit gained by output feedback, we present the following theorem from [17], which characterizes the GDoF for the FC-IC without feedback.

Theorem 2 ([17], Theorem 3.1)**.** *The per-user* GDoF *for the K-user interference channel without feedback is given by*

$$d_{\textit{No}\ \rm FB}(\alpha) = \begin{cases} 1-\alpha & 0\le\alpha\le\frac{1}{2} \textit{ (noisy interference)} \\ \alpha & \frac{1}{2}\le\alpha\le\frac{2}{3} \textit{ (weak interference)} \\ 1-\frac{\alpha}{2} & \frac{2}{3}\alpha<1 \textit{ (moderate weak interference)} \\ \frac{1}{K} & \alpha=1 \\ \frac{\alpha}{2} & 1<\alpha\le 2 \textit{ (strong interference)} \\ 1 & \alpha>2 \textit{ (very strong interference).} \end{cases} \tag{6}$$

The generalized degrees of freedom of the K-user interference channel with/without feedback are illustrated in Figure 2. As derived in [17], the GDoF for the K-user no feedback case, is similar to that of 2-user case [2], except for $\alpha=1$. Similarly, here we show that for the channel with feedback, the GDoF for the K-user case is the same as that of the 2-user channel [15], except for $\alpha=1$. At this particular point, the whole K by K network behaves as a singular network, and the available DoF $=1$ has to be shared between K users.

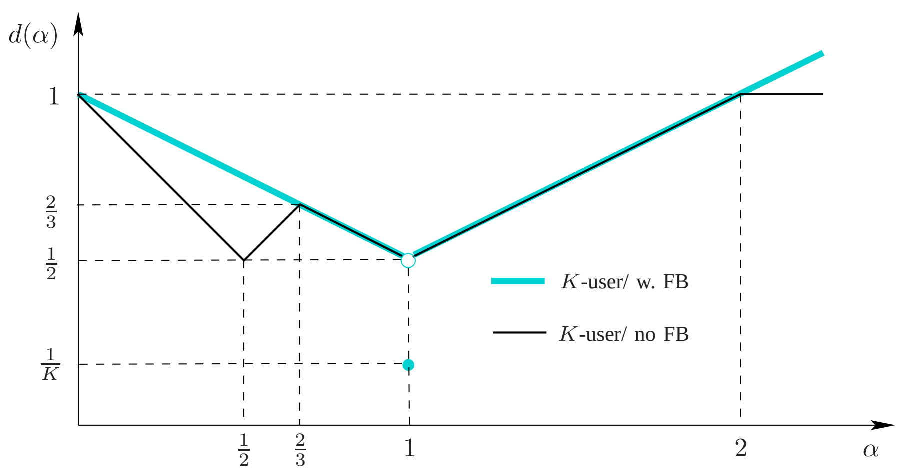


Fig. 2. The per-user generalized degrees of freedom for the K-user interference channel.

The following theorem characterizes the approximate capacity of the channel for arbitrary signal-to-noise ratio.

Theorem 3. *The symmetric[1] capacity of the K user interference channel with feedback with $\alpha \neq 1$ can be approximated by*

$$C_{\rm sym} \triangleq \frac{1}{4}\log(1+\mathsf{SNR}+\mathsf{INR}) + \frac{1}{4}\log\left(1+\frac{\mathsf{SNR}}{\mathsf{INR}}\right). \tag{7}$$

More precisely, the symmetric capacity is upper bounded by $R_{\rm sym} \leq C_{\rm sym} + \frac{K-1}{4} + \frac{1}{2}\log K$. Moreover, there exists a coding scheme that can support any rate satisfying $R_{\rm sym} \leq C_{\rm sym} - \frac{1}{4}\log 2K^3$.

IV. The Deterministic Model

In this section we study the problem of interest in a deterministic framework introduced in [5]. The key point in this model is to focus on signal interactions instead of the additive noise, and obtain insight about both coding schemes and outer bounds for the original problem.

The intuition behind this approach is that the noise is modelled by a deterministic operation on the received signal which splits the received signal into a completely useless part and a completely noiseless part. The part of the received signal below the noise level is completely useless since it is corrupted by noise. However, the part above the noise level is assumed to be not affected by noise and can be used to retrieve information.

Let p be any prime number and $\mathbb{F}$ be the finite field over the set $\{0, 1, \ldots, p-1\}$ with sum and product operations modulo p. Moreover, define

$$n = \lfloor \log_p \mathsf{SNR} \rfloor \quad \text{and} \quad m = \lfloor \log_p \mathsf{INR} \rfloor.$$

Each received signal can be mapped into a p-ary stream. Let $X_k \in \mathbb{F}^q$ and $Y_k \in \mathbb{F}^q$ be the p-ary expansion of the transmit and received signal by user k, respectively, where $q = \max\{m, n\}$. The shift linear deterministic channel model for this network can be written as

$$Y_k = D^{q-n} X_k + \sum_{i \neq k} D^{q-m} X_i, \tag{8}$$

where all the operations are performed modulo p. Here, D is the shift matrix, defined as

$$D = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix}_{q \times q}.$$

The following theorem characterizes the symmetric capacity of the deterministic network introduced above. In the rest of this section, we prove this theorem by first deriving an upper bound on the symmetric capacity, and then

[1] Indeed our result is stronger than the statement of Theorem 3, in the sense that we prove that the sum-capacity of the channel with feedback is approximately $KR_{\rm sym}$. However, since our focus in this work is on a symmetric topology, we present the result in terms of per-user symmetric capacity.

proposing coding schemes for different interference regimes. The ideas arising in this section will be later used when we focus on the Gaussian network in Sections VI and V.

Theorem 4. *The symmetric feedback capacity of the linear deterministic K-user fully connected interference channel with parameters n and m is given by*

$$R_{\text{sym}} = \begin{cases} n - \frac{m}{2} & n > m \text{ (weak interference)}, \\ \frac{n}{K} & m = n, \\ \frac{m}{2} & n < m \text{ (strong interference)}. \end{cases} \tag{9}$$

Remark 1. *From the rate expression in Theorem 4 one can easily see that the normalized feedback capacity of the channel under the linear deterministic model is given by*

$$\frac{R_{\text{sym}}}{n} = \begin{cases} 1 - \frac{1}{2}\left(\frac{m}{n}\right) & \frac{m}{n} < 1, \\ \frac{1}{K} & \frac{m}{n} = 1, \\ \frac{1}{2}\left(\frac{m}{n}\right) & \frac{m}{n} > 1, \end{cases}$$

which is analogous to the GDoF *expression in Theorem 1, by noting that m/n is analogous to α for the Gaussian setting.*

A. Encoding Scheme

In the following we present a transmission scheme that can achieve the rate claimed in Theorem 4. We first demonstrate the proposed scheme in two examples with specific parameters, through which the basic ideas and intuitions are transparent. Although generalization of the proposed coding strategy for arbitrary n and m is straight-forward, we present the scheme and its analysis in Appendix A in sake of completeness.

a) Weak Interference Regime $(m < n)$: The goal is to achieve $R_{\text{sym}} = n - \frac{m}{2}$ bits per user. We propose an encoding that operates on a block of length 2. The basic idea can be seen from Fig. 3, wherein the coding scheme is demonstrated for $n = 3$ and $m = 2$.

For these specific parameters, we have $R_{\text{sym}} = 2$. As it is shown in Fig. 3, the proposed coding scheme is able to convey four intended symbols from each transmitter to its respective receiver in two channel uses. The information symbols intended for $\mathbf{Rx}_1$ are denoted by a_1, a_2, a_3, a_4. Each transmitter sends three fresh symbols in its first channel use. Receivers get one interference-free symbol, and two more equations, including their intended symbol as well as interference. The output signals are sent to the transmitters over the feedback link, in order to be used for the next transmission. In the second channel use, each transmitter forwards the interfering parts of its received feedback on its top two levels. The lowest level will be used to transmit the remaining fresh symbol.

Now, consider the received signals at $\mathbf{Rx}_1$ in two channel uses. It has received 6 linearly independent equations, involving 8 variables, which seems to be unsolvable at first glance. However, we do not need to decode b_1, b_2, c_1, and c_2, individually. Instead, we can solve the system of linear of equations in a_1, a_2, a_3, a_4, (b_1+c_1), and (b_2+c_2), which can be solved for the intended variables. Hence, a per-user rate of 2 symbols/channel-use is achievable with feedback.

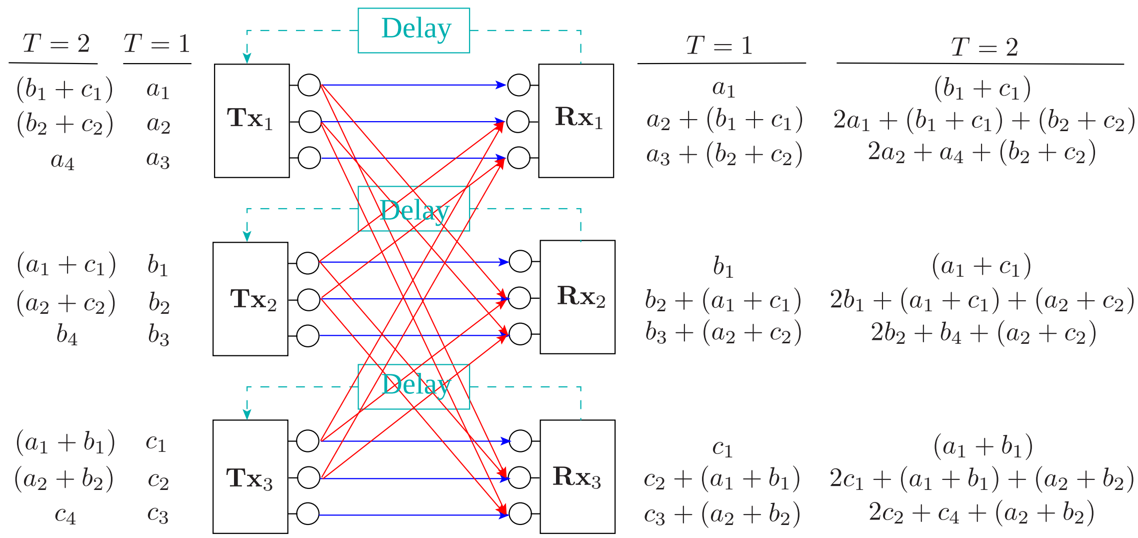


Fig. 3. Coding scheme for the linear deterministic model in the weak interference regime, for $K=3$, $n=3$, and $m=2$.

b) Strong Interference Regime $(m>n)$*:* In this section we present an encoding scheme which can support a symmetric rate of $R_{\text{sym}} = \frac{m}{2}$. Again we focus on specific parameters, $n=2$ and $m=3$, which implies $R_{\text{sym}} = 3/2$.

As shown in Fig. 4, the proposed coding strategy delivers three intended symbols to each receiver in two channel uses. In the first channel use, each transmitter sends its fresh symbols to its respective receiver. However, due to the strong interference, receivers are not able to decode any part of their intended symbols, and can only send their received signals to their respective transmitters through the feedback links. Each transmitter then removes its own contribution from the received signal, and forwards the remaining over the second channel use. Similar to the weak interference regime, at the end of the transmission each receiver has 6 equations, involving three intended symbols (a_1, a_2 and a_3 for $\mathbf{Rx}_1$), and three interfering symbols (b_1+c_1, b_2+c_2, and b_3+c_3 for $\mathbf{Rx}_1$), which can be solved. Note that the system of linear equations might not be linearly independent, depending of p, the field size. In particular, for these specific parameters, operating in the binary field ($p=2$), the coefficient of a_3 becomes zero, and therefore a_3 cannot be decoded from the received equations. However, p is an arbitrary parameter, which can be carefully chosen to provide a full-rank coefficient matrix. Therefore, a per-user rate of $3/2$ symbols/channel-use is achieved with feedback.

c) Moderate Interference Regime $(m=n)$*:* As discussed in the outer bound argument, the capacity curve is discontinuous at $m=n$. A trivial encoding scheme to achieve rate $R_{\text{sym}} = n/K$ is to perform time-sharing over K blocks: in block k only $\mathbf{Tx}_k$ transmits its message at rate $R_k = n$ while all the transmitters keep silent. Note that this coding scheme does not get any benefit from the feedback link.

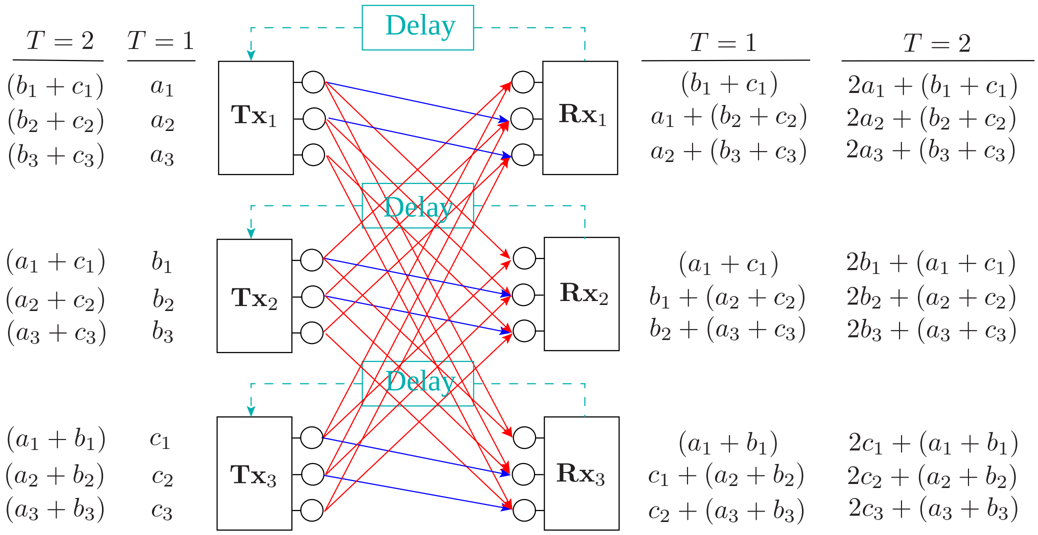


Fig. 4. Coding scheme for the linear deterministic model in the strong interference regime, for $K=3$, $n=2$, and $m=3$.

B. Outer Bound

In this section we derive an outer bound on the symmetric feedback capacity of the fully-connected interference channel. Assume there exists an encoding scheme with block length T, which can reliably convey messages of each transmitter to its intended receiver. We begin with the following chain of inequalities:

$$
\begin{aligned}
H(W_1)+H(W_2) &= H(W_1,W_2|W_3,\ldots,W_K) \qquad (10)\\
&\leq H(W_1,W_2,Y_1^T,Y_2^T|W_3,\ldots,W_K)\\
&= H(Y_2^T|W_3,\ldots,W_K)+H(W_2|W_3,\ldots,W_K,Y_2^T)+H(Y_1^T|W_2,W_3,\ldots,W_K,Y_2^T)\\
&\quad +H(W_1|W_2,W_3,\ldots,W_K,Y_1^T,Y_2^T)\\
&\leq H(Y_2^T)+H(W_2|Y_2^T)+H(Y_1^T|W_2,W_3,\ldots,W_K,Y_2^T)+H(W_1|Y_1^T)\\
&\leq T[\max(m,n)+2\epsilon_T]+H(Y_1^T|W_2,W_3,\ldots,W_K,Y_2^T), \qquad (11)
\end{aligned}
$$

where (10) holds since messages are assumed to be independent, and (11) is due to Fano's inequality, in which $\epsilon_T \to 0$, as T grows. We can continue with bounding the remaining term in (11) as

$$
\begin{aligned}
H(Y_1^T|W_2,W_3,\dots,W_K,Y_2^T) &\le H(Y_1^T,Y_3^T,\dots,Y_K^T|W_2,W_3,\dots,W_K,Y_2^T) \\
&= \sum_{t=1}^{T} H(Y_{1t},Y_{3t},\dots,Y_{Kt}|W_2,W_3,\dots,W_K,Y_2^T,Y_1^{t-1},Y_3^{t-1},\dots,Y_K^{t-1}) \\
&= \sum_{t=1}^{T} H(Y_{1t},Y_{3t},\dots,Y_{Kt}|W_2,W_3,\dots,W_K,Y_2^T,Y_1^{t-1},Y_3^{t-1},\dots,Y_K^{t-1},X_{2t},X_{3t},\dots,X_{Kt}) \qquad (12) \\
&\le \sum_{t=1}^{T} H\Big(D^{q-n}X_{1t}+\sum_{i\neq 1}D^{q-m}X_{it}, D^{q-n}X_{3t}+\sum_{i\neq 3}D^{q-m}X_{it},\dots,D^{q-n}X_{Kt}+\sum_{i\neq K}D^{q-m}X_{it}\Big| \\
&\qquad\qquad Y_{2t},X_{2t},X_{3t},\dots,X_{Kt}\Big) \\
&\le \sum_{t=1}^{T} H(D^{q-n}X_{1t},D^{q-m}X_{1t}|Y_{2t}-D^{q-n}X_{2t}-D^{q-m}\sum_{j>2}X_{jt}) \\
&= \sum_{t=1}^{T} H(D^{q-n}X_{1t},D^{q-m}X_{1t}|D^{q-m}X_{1t}) \\
&= T(n-m)^+, \qquad (13)
\end{aligned}
$$

where (12) is due to the fact that $X_{jt} = f_{jt}(W_j, Y_j^{t-1})$. Replacing (13) in (11) we arrive at

$$R_1 + R_2 \le \frac{1}{T}[H(W_1)+H(W_2)] \le \max(m,n) + (n-m)^+ + 2\epsilon_T = \max(m, 2n-m) + 2\epsilon_T. \qquad (14)$$

Finally, since we are interested in symmetric rate characterization, we can set $R_1 = R_2$, which yields

$$R_{\text{sym}} \le \max\left(\frac{m}{2}, n-\frac{m}{2}\right) + \epsilon_T. \qquad (15)$$

Letting $T \to \infty$ and $\epsilon_T \to 0$, we obtain the upper bound as claimed in Theorem 4.

The capacity behavior of the network has a discontinuity at $m = n$, where the symmetric achievable rate scales inverse linearly with K. The reason behind this phenomenon is very apparent by focusing on the deterministic model. This study reveals that when $m = n$ the received signals at all the receivers are *exactly* the same. Therefore, each receiver should be able to decode all the messages, and hence its decoding capability is shared between all the signals, which results in $R_{\text{sym}} = n/K$. More formally, we can write

$$
\begin{aligned}
T\sum_{k=1}^{K} R_k = H(W_1,W_2,\dots,W_K) &\le I(W_1,W_2,\dots,W_K;Y_1^T,Y_2^T,\dots,Y_K^Y) + KT\epsilon \\
&= I(W_1,W_2,\dots,W_K;Y_1^T) + KT\epsilon \qquad (16) \\
&\le H(Y_1^T) + KT\epsilon \le Tn + KT\epsilon, \qquad (17)
\end{aligned}
$$

where (16) is due to the fact that $Y_1^T = Y_2^T = \cdots = Y_K^T$. Dividing (17) by KT and setting $R_1 = \cdots = R_K = R_{\text{sym}}$, we arrive at $R_{\text{sym}} \le n/K$.

V. The Gaussian Network: A Coding Scheme

The encoding scheme we propose for this problem is similar to that of the 2-user case. It is shown in [15] that for the 2-user feedback interference channel, depending on the interference regime (value of α), it is (approximately) optimum to decode the interfering message. Due to existence of the feedback, decoding the interference is not only useful for its removal and consequent decoding of the desired message (akin to the strong interference regime without feedback), but also helps for decoding a part of the intended message that is conveyed through the feedback path. In the 2-user case, at the end of the transmission block, each receiver not only decodes its own message completely, but also partially decodes the message of the other receiver.

A fundamental difference here is that in the K-user problem, there are multiple interfering messages that can be heard at each receiver. Partial decoding of all interfering messages would dramatically decrease the maximum rate of the desired message. Our approach to deal with this is to consider the total interference received from all other users as a single message and decode it, without resorting to resolving the individual component of the interference. There are two key conditions to be fulfilled that allow us to perform such decoding, namely, (i) interfering signals should be *aligned*, and (ii) the summation of interfering signals should belong to a message set of proper size which can be decoded at each receiver. Here, the first condition is satisfied since the network is symmetric (all the interfering links have the same gain), and therefore all the interfering messages are received at the same power level. In order to satisfy the second condition, we can use a common *lattice code* in all transmitters, instead of random Gaussian codebooks. The structure of a lattice codebook and its closeness with respect to summation, imply that the summation of aligned interfering codewords observed at each receiver is still a codeword from the same codebook. This allows us to perform decoding by searching over the single codebook, instead of the Cartesian product of all codebooks. Due to the fact that the aligned interference is decoded, we call this coding scheme *cooperative interference alignment*.

We use the following lemma in our analysis of the proposed coding scheme. The proof of this can be found in Appendix B

Lemma 1. *Let $\mathcal{C} = \Lambda_c \cap \mathcal{V}_q$ be a good channel code with rate R, where $\mathcal{V}_q$ is the Voronoi cell of the coarse lattice Λ_q, and Λ_c is the fine lattice with $\Lambda_q \subseteq \Lambda_c$. Moreover, the average power of the codewords is 1, that is $\frac{1}{n}\sigma^2(\Lambda_q) = 1$. Consider a lattice codeword $\mathbf{c} \in \mathcal{C}$ and a random dither vector $\mathbf{d}$, and the random object $\mathbf{s} = [\mathbf{c} - \mathbf{d}] \mod \Lambda_q$. Then $\mathbf{c}$ can be decoded from $\mathbf{y} = \alpha \mathbf{s} + \mathbf{z}$ provided that*

$$R \leq \frac{1}{2} \log \left(1 + \frac{\alpha^2}{\beta^2}\right), \tag{18}$$

where $\beta^2 = \mathbb{E}[\mathbf{z}^2]$.

In the rest of this section, we prove the direct part of Theorem 3. The analysis of two cases, namely weak and strong interference regimes, is performed separately.

A. Weak Interference Regime $\alpha < 1$

We consider three messages w_{k0}, w_{k1}, and w_{k2}, for transmitter $\mathbf{Tx}_k$ which will be conveyed to receiver $\mathbf{Rx}_k$ over two blocks. All similar sub-messages from different users have the same rates, which are denoted by R_{k0}, R_{k1}, and R_{k2}. Encoding of w_{k1} and w_{k2} is performed using usual random Gaussian codebooks with block length T and average power 1, which results in codewords $\mathbf{s}_{k1}$ and $\mathbf{s}_{k2}$.

In order to encode w_{k0}, we use a common lattice code which is shared between all transmitters. Each transmitter maps its sub-message to a lattice codeword $\mathbf{s}_{k0}$. Let Λ_q be a good quantization lattice with $\frac{1}{T}\sigma^2(\Lambda_q) = 1$, and Λ_c be a good fine lattice good for channel coding, with $\Lambda_q \subseteq \Lambda_c$. We denote the Voronoi cell of the lattices by $\mathcal{V}_q$ and $\mathcal{V}_c$, respectively. It is well-known that $\mathcal{C} = \Lambda_c \cap \mathcal{V}_q$ is a good channel codebook [18], which is a closed set with respect to summation under the "$\mod \Lambda_q$" operation. We also use $[\mathbf{x}]_q$ to denote $\mathbf{x} \mod \Lambda_q$. Each sub-message w_{k0} is mapped to a lattice codeword $\mathbf{c}_{k0} = f\left(w_{k0}\right)$. We denote by $\mathbf{c}_0$ the codeword $[\mathbf{c}_{10} + \mathbf{c}_{20} + \cdots + \mathbf{c}_{K0}]_q$, and define an artificial message w_0 the message corresponding to this codeword, that is

$$w_0 = f^{-1}\big([f(w_{10}) + f(w_{20}) + \cdots f(w_{K0})]_q\big). \tag{19}$$

Once the lattice codeword is found, the encoder at $\mathbf{Tx}_k$ computes $\mathbf{s}_{k0} = [\mathbf{c}_{k0} - \mathbf{d}_k]_q$, where $\{\mathbf{d}_k : \ k = 1, \ldots, K\}$ are random dither vectors, with $\mathbf{d}_k \sim \mathrm{Unif}(\mathcal{V}_q)$ and known at all the transmitters and receivers. Finally, the signal transmitted by $\mathbf{Tx}_k$ in the first block (of length T) is formed as

$$\mathbf{x}_{k1} = \sqrt{\frac{\mathsf{INR}-1}{\mathsf{INR}}}\mathbf{s}_{k0} + \sqrt{\frac{1}{\mathsf{INR}}}\mathbf{s}_{k1}. \tag{20}$$

Therefore, the signal received at $\mathbf{Rx}_k$ can be written as

$$\mathbf{y}_{k1} = \sqrt{\mathsf{SNR}}\mathbf{x}_{k1} + \sqrt{\mathsf{INR}}\sum_{i\neq k}\mathbf{x}_{i1} + \mathbf{z}_{k1} \tag{21}$$

$$= \sqrt{\frac{\mathsf{SNR}}{\mathsf{INR}}(\mathsf{INR}-1)}\mathbf{s}_{k0} + \sqrt{\frac{\mathsf{SNR}}{\mathsf{INR}}}\mathbf{s}_{k1} + \sqrt{\mathsf{INR}-1}\sum_{i\neq k}\mathbf{s}_{i0} + \sum_{i\neq k}\mathbf{s}_{i1} + \mathbf{z}_{k1} \tag{22}$$

This received signal is sent to the transmitter $\mathbf{Tx}_k$ over the feedback link. Knowing $\mathbf{x}_{k1}$ and $\mathbf{y}_{k1}$, the transmitter can compute

$$\tilde{\mathbf{y}}_k = \mathbf{y}_{k1} - (\sqrt{\mathsf{SNR}} - \sqrt{\mathsf{INR}})\mathbf{x}_{k1} = \sqrt{\mathsf{INR}}\sum_{i=1}^{K}\mathbf{x}_{i1} + \mathbf{z}_{k1} = \sqrt{\mathsf{INR}-1}\sum_{i=1}^{K}\mathbf{s}_{i0} + \sum_{i=1}^{K}\mathbf{s}_{i1} + \mathbf{z}_{k1}.$$

Using Lemma 1, $\mathbf{c}_0$ can be decoded from $\tilde{\mathbf{y}}_k$ at $\mathbf{Tx}_k$ provided that

$$R_0 \leq \frac{1}{2}\log\left(1 + \frac{\mathsf{INR}-1}{K+1}\right). \tag{23}$$

Note that at this point $\mathbf{Rx}_k$ cannot decode $\mathbf{c}_0$.

Once $\mathbf{c}_0$ is decoded, each transmitter creates $\mathbf{s}_0 = [\mathbf{c}_0 - \mathbf{d}_0] \mod \Lambda_q$, where $\mathbf{d}_0$ is a common random dither vector known at all transmitters/receivers. In the second block, $\mathbf{Tx}_k$ transmits

$$\mathbf{x}_{k2} = \sqrt{\frac{\mathsf{INR}-1}{\mathsf{INR}}}\mathbf{s}_0 + \sqrt{\frac{1}{\mathsf{INR}}}\mathbf{s}_{k2}. \tag{24}$$

The signal received at $\mathbf{Rx}_k$ in the second block can be written as

$$\mathbf{y}_{k2} = \sqrt{\mathsf{SNR}}\mathbf{x}_{k2} + \sqrt{\mathsf{INR}}\sum_{i\neq k}\mathbf{x}_{i2} + \mathbf{z}_{k2} \tag{25}$$

$$= \sqrt{\frac{\mathsf{SNR}}{\mathsf{INR}}(\mathsf{INR}-1)}\mathbf{s}_0 + \sqrt{\frac{\mathsf{SNR}}{\mathsf{INR}}}\mathbf{s}_{k2} + \sqrt{\mathsf{INR}-1}\sum_{i\neq k}\mathbf{s}_0 + \sum_{i\neq k}\mathbf{s}_{i2} + \mathbf{z}_{k2} \tag{26}$$

$$= \left(\sqrt{\frac{\mathsf{SNR}}{\mathsf{INR}}} + K - 1\right)\sqrt{\mathsf{INR}-1}\mathbf{s}_0 + \sqrt{\frac{\mathsf{SNR}}{\mathsf{INR}}}\mathbf{s}_{k2} + + \sum_{i\neq k}\mathbf{s}_{i2} + \mathbf{z}_{k2}. \tag{27}$$

Receiver $\mathbf{Rx}_k$ first decodes $\mathbf{c}_0$ treating everything else as noise. This is possible as long as

$$R_0 \leq \frac{1}{2}\log\left(1 + \frac{(\mathsf{INR}-1)\left(\sqrt{\frac{\mathsf{SNR}}{\mathsf{INR}}} + K - 1\right)^2}{\frac{\mathsf{SNR}}{\mathsf{INR}} + K}\right). \tag{28}$$

After decoding and removing $\mathbf{c}_0$ from the received signal, $\mathbf{Rx}_k$ can decode the Gaussian codeword $\mathbf{s}_{k2}$, provided that

$$R_2 \leq \frac{1}{2}\log\left(1 + \frac{\mathsf{SNR}}{K\mathsf{INR}}\right). \tag{29}$$

The decoder also uses $\mathbf{c}_0$ to reconstruct $\mathbf{s}_0$ and remove it from $\mathbf{y}_{k1}$ in order to consecutively decode $\mathbf{s}_{k0}$ and $\mathbf{s}_{k1}$. It first computes

$$\mathbf{y}_{k1} + \sqrt{\mathsf{INR}-1}\left(\sum_{i=1}^{K}\mathbf{d}_k - \mathbf{s}_0 - \mathbf{d}_0\right) \tag{30}$$

$$= \left(\sqrt{\frac{\mathsf{SNR}}{\mathsf{INR}}} - 1\right)\sqrt{\mathsf{INR}-1}\mathbf{s}_{k0} + \sqrt{\frac{\mathsf{SNR}}{\mathsf{INR}}}\mathbf{s}_{k1} + \sqrt{\mathsf{INR}-1}\left[\sum_{i=1}^{K}(\mathbf{s}_{i0} + \mathbf{d}_i) - (\mathbf{s}_0 + \mathbf{d}_0)\right] + \sum_{i\neq k}\mathbf{s}_{i1} + \mathbf{z}_{k1}. \tag{31}$$

Note that the term inside brackets equals zero when taking modulo Λ_q. Codewords $\mathbf{s}_{k0}$ and $\mathbf{s}_{k1}$ can be decoded provided that

$$R_0 \leq \frac{1}{2}\log\left(1 + \frac{(\mathsf{INR}-1)\left(\sqrt{\frac{\mathsf{SNR}}{\mathsf{INR}}} - 1\right)^2}{\frac{\mathsf{SNR}}{\mathsf{INR}} + K}\right), \tag{32}$$

$$R_1 \leq \frac{1}{2}\log\left(1 + \frac{\mathsf{SNR}}{K\mathsf{INR}}\right). \tag{33}$$

It only remains to choose R_0, R_1, and R_2 that satisfy all constraints in (23), (28), (29), (32), and (33). It is easy to verify that the choice of

$$R_0^\star = \min\left\{\frac{1}{2}\log\left(1 + \frac{\mathsf{INR}-1}{K+1}\right), \log\left(1 + \frac{(\mathsf{INR}-1)(\sqrt{\mathsf{SNR}} - \sqrt{\mathsf{INR}})^2}{\mathsf{SNR} + K\mathsf{INR}}\right)\right\},$$
$$R_1^\star = R_2^\star = \frac{1}{2}\log\left(1 + \frac{\mathsf{SNR}}{K\mathsf{INR}}\right) \tag{34}$$

satisfies all the constraints, and therefore

$$
\begin{aligned}
R_{\text{sym}} &= \frac{1}{2}(R_0^\star + R_1^\star + R_2^\star) \\
&= \frac{1}{2}\log\left(1+\frac{\mathsf{SNR}}{K\mathsf{INR}}\right) + \min\left\{\frac{1}{4}\log\left(1+\frac{\mathsf{INR}-1}{K+1}\right), \frac{1}{4}\log\left(1+\frac{(\mathsf{INR}-1)(\sqrt{\mathsf{SNR}}-\sqrt{\mathsf{INR}})^2}{\mathsf{SNR}+K\mathsf{INR}}\right)\right\}
\end{aligned}
$$

can be simultaneously achieved for all the K pairs of transmitters/receivers.

In the following we rephrase this achievable rate in a manner so that it can be easily compared to C_{sym} in Theorem 3. Let $\delta > 0$ be an arbitrarily positive constant, where $\alpha = 1-\delta$. The definition of α in (3) implies that $\mathsf{INR} = \mathsf{SNR}^{1-\delta}$. It is easy to verify that

$$
\left(1+\frac{(\mathsf{INR}-1)(\sqrt{\mathsf{SNR}}-\sqrt{\mathsf{INR}})^2}{\mathsf{SNR}+K\mathsf{INR}}\right)\left(\frac{\mathsf{SNR}+K\mathsf{INR}}{K\mathsf{INR}}\right) \geq \frac{1}{K}\left(1+(\sqrt{\mathsf{SNR}}-\sqrt{\mathsf{INR}})^2\right) \geq \frac{1}{4K}\left(1+\mathsf{SNR}+\mathsf{INR}\right)
$$

for $K \geq 2$. Therefore

$$
\begin{aligned}
&\frac{1}{4}\log\left(1+\frac{(\mathsf{INR}-1)(\sqrt{\mathsf{SNR}}-\sqrt{\mathsf{INR}})^2}{\mathsf{SNR}+K\mathsf{INR}}\right) + \frac{1}{2}\log\left(1+\frac{\mathsf{SNR}}{K\mathsf{INR}}\right) \\
&\quad \geq \frac{1}{4}\log\left(1+\mathsf{SNR}+\mathsf{INR}\right) + \frac{1}{4}\log\left(1+\frac{\mathsf{SNR}}{\mathsf{INR}}\right) - \frac{1}{4}\log 4K^2. \qquad (35)
\end{aligned}
$$

On the other hand, since $\mathsf{INR} < \mathsf{SNR}$, we have

$$
\left(1+\frac{\mathsf{INR}-1}{K+1}\right)\left(1+\frac{\mathsf{SNR}}{K\mathsf{INR}}\right) \geq \frac{1}{K(K+1)}\left(1+\mathsf{INR}+\mathsf{SNR}\right), \qquad (36)
$$

which implies

$$
\frac{1}{4}\log\left(1+\frac{\mathsf{INR}-1}{K+1}\right) + \frac{1}{2}\log\left(1+\frac{\mathsf{SNR}}{K\mathsf{INR}}\right) \geq \frac{1}{4}\log\left(1+\mathsf{INR}+\mathsf{SNR}\right) + \frac{1}{4}\log\left(1+\frac{\mathsf{SNR}}{\mathsf{INR}}\right) - \frac{1}{4}\log K^2(K+1) \qquad (37)
$$

Therefore, for $\alpha < 1$, the symmetric rate

$$
R_{\text{sym}} = \frac{1}{4}\log\left(1+\mathsf{INR}+\mathsf{SNR}\right) + \frac{1}{4}\log\left(1+\frac{\mathsf{SNR}}{\mathsf{INR}}\right) - \max\left\{\frac{1}{4}\log 4K^2, \frac{1}{4}\log K^2(K+1)\right\} \qquad (38)
$$

is achievable.

B. *Strong Interference Regime* ($\alpha > 1$)

The encoding scheme for the strong interference regime is slightly simpler than that for the weak interference case. In the following we propose an encoding scheme over two blocks. Each transmitter has a message w_k of rate R_0 which is mapped to a lattice codeword $\mathbf{c}_k$. Again we assume that all encoders share a common common lattice code. The transmitting sequence over the first block is obtained by adding the random dither vector, that is, $\mathbf{x}_{k1} = \mathbf{s}_k = [\mathbf{c}_k - \mathbf{d}_k]_q$. At the end of the first block receiver are not able to decode any useful information, and just forward their received signal to the encoders. The transmitter first decodes the effective interference after removing its own signal, and then creates a lattice codeword by re-adding its codeword to that. At the end of the second block, each receiver first decodes the sum interference, and then removes it from its received signal in the first block. This allows decoding of the intended message.

The received signal at $\mathbf{Rx}_k$ can be written as

$$\mathbf{y}_{k1} = \sqrt{\mathsf{SNR}}\mathbf{x}_{k1} + \sqrt{\mathsf{INR}} \sum_{i \neq k} \mathbf{x}_{i1} + \mathbf{z}_{k1}, \tag{39}$$

where

$$\mathbf{x}_{k1} = \mathbf{s}_k = [\mathbf{c}_k - \mathbf{d}_k]_q,$$

as stated above. This signal is sent back to the transmitter over the feedback. After removing $\mathbf{x}_{k1}$ from the channel output and taking the modulo operation, transmitter k has access to

$$\begin{aligned} \tilde{\mathbf{y}}_k &= \left[\frac{1}{\sqrt{\mathsf{INR}}} \left(\mathbf{y}_{k1} - \sqrt{\mathsf{SNR}}\mathbf{x}_{k1} \right) \right]_q = \left[\sum_{i \neq k} \mathbf{x}_{i1} + \frac{1}{\sqrt{\mathsf{INR}}} \mathbf{z}_{k1} \right]_q \\ &= \left[\sum_{i \neq k} [\mathbf{c}_i - \mathbf{d}_i]_q + \mathbf{z}'_{k1} \right]_q = \left[\left[\sum_{i \neq k} \mathbf{c}_i \right]_q - \left[\sum_{i \neq k} \mathbf{d}_i \right]_q + \mathbf{z}'_{k1} \right]_q, \end{aligned} \tag{40}$$

where $\mathbf{z}'_{k1} = \mathbf{z}_{k1} / \sqrt{\mathsf{INR}}$ is an additive Gaussian noise variable with variance $\mathbb{E} \parallel \mathbf{z}'_{k1} \parallel^2 \leq \mathsf{INR}^{-1}$. The transmitter wishes to decode the lattice point $[\sum_{i \neq k} \mathbf{c}_i]_q$. Note that here $[\sum_{i \neq k} \mathbf{d}_i]_q$ is a known dither vector drawn from the uniform distribution over the Voronoi cell $\mathcal{V}_q$. Using Lemma 1, we can decode $[\sum_{i \neq k} \mathbf{c}_i]_q$ provided that

$$R_0 \leq \frac{1}{2} \log(1 + \frac{1}{\mathsf{INR}^{-1}}) = \frac{1}{2} \log(1 + \mathsf{INR}). \tag{41}$$

Having $[\sum_{i \neq k} \mathbf{c}_i]_q$ decoded at the k-th transmitter, $\mathbf{Tx}_k$ sends the following signal in the second block:

$$\mathbf{x}_{k2} = \mathbf{s}_0 = \left[\left[\sum_{i \neq k} \mathbf{c}_i \right]_q + \mathbf{c}_k - \mathbf{d}_0 \right]_q = \left[\sum_{i=1}^{K} \mathbf{c}_i - \mathbf{d}_0 \right]_q = [\mathbf{c}_0 - \mathbf{d}_0]_q, \tag{42}$$

where $\mathbf{c}_0 = [\sum_i \mathbf{c}_i]_q$, and $\mathbf{d}_0$ is a random dither known at all the nodes in the network. Note that, this way all the transmitters send the same sequence simultaneously. Therefore, the received signal at receiver k is given by

$$\mathbf{y}_{k2} = \left(\sqrt{\mathsf{SNR}} + (K-1)\sqrt{\mathsf{INR}} \right) \mathbf{s}_0 + \mathbf{z}_{k2}. \tag{43}$$

Having received this, each decoder wishes to decode $\mathbf{c}_0$, which is feasible as long as

$$R_0 \leq \frac{1}{2} \log \left(1 + \mathsf{SNR} + (K-1)^2 \mathsf{INR} + 2(K-1)\sqrt{\mathsf{SNR} \cdot \mathsf{INR}} \right). \tag{44}$$

Next, $\mathbf{Rx}_k$ computes

$$\begin{aligned} &\left[\gamma \mathbf{y}_{k1} - \gamma \sqrt{\mathsf{INR}} \left[\mathbf{c}_0 - \sum_{i=1}^{K} \mathbf{d}_i \right]_q + \mathbf{d}_k \right]_q \\ &\quad = \left[\gamma \left(\sqrt{\mathsf{SNR}} - \sqrt{\mathsf{INR}} \right) [\mathbf{c}_k - \mathbf{d}_k]_q + \gamma \sqrt{\mathsf{INR}} \sum_{i=1}^{K} [\mathbf{c}_i - \mathbf{d}_i]_q + \gamma \mathbf{z}_{k1} - \gamma \sqrt{\mathsf{INR}} \left[\mathbf{c}_0 - \sum_{i=1}^{K} \mathbf{d}_i \right]_q + \mathbf{d}_k \right]_q \\ &\quad = \left[\gamma \left(\sqrt{\mathsf{SNR}} - \sqrt{\mathsf{INR}} \right) [\mathbf{c}_k - \mathbf{d}_k]_q + \gamma \mathbf{z}_{k1} + \mathbf{d}_k \right]_q, \end{aligned} \tag{45}$$

where γ is a scalar depending on the signal and noise power which plays the same role as in the proof of Lemma 1 (see Appendix B). Finally, $\mathbf{Rx}_k$ uses the right-hand side (RHS) of (45) to decode $\mathbf{c}_0$. Lemma 1 guarantees a successful decoding of w_k at $\mathbf{Rx}_k$, provided that

$$R_0 \leq \frac{1}{2}\log\left(1+\left(\sqrt{\mathsf{INR}}-\sqrt{\mathsf{SNR}}\right)^2\right). \tag{46}$$

It is easy to verify that the choice of R_0 on the RHS of (46) satisfies both (41) and (44). Therefore, since this coding scheme is performed over two blocks, a symmetric rate of

$$R_{\mathrm{sym}} = \frac{1}{2}R_0 = \frac{1}{4}\log\left(1+\left(\sqrt{\mathsf{INR}}-\sqrt{\mathsf{SNR}}\right)^2\right) \tag{47}$$

is achievable with feedback.

Similar to the weak interference case, we rephrase this achievable rate in a form to be easily comparable to the upper bound. First note that $\alpha > 1$, and $\mathsf{INR} = \mathsf{SNR}^{1+\delta}$ for some positive $\delta > 0$, where $\alpha = 1+\delta$. Hence,

$$\left(\sqrt{\mathsf{INR}}-\sqrt{\mathsf{SNR}}\right)^2 \geq \frac{1}{4}(\mathsf{SNR}+\mathsf{INR}). \tag{48}$$

On the other hand, for $\mathsf{INR} > \mathsf{SNR}$, we have $1+\mathsf{SNR}/\mathsf{INR} < 2$. Therefore,

$$\begin{aligned} R_{\mathrm{sym}} &= \frac{1}{4}\log\left(1+(\sqrt{\mathsf{SNR}}-\sqrt{\mathsf{INR}})^2\right) \\ &\geq \frac{1}{4}\log\left(1+\mathsf{SNR}+\mathsf{INR}\right)+\frac{1}{4}\log\left(1+\frac{\mathsf{SNR}}{\mathsf{INR}}\right)-\frac{1}{4}\log 8. \end{aligned} \tag{49}$$

Hence, $R_{\mathrm{sym}} = C_{\mathrm{sym}} - \frac{1}{4}\log 8$ is achievable, which completes the proof of Theorem 3.

Remark 2. *It is worth mentioning that the coding schemes proposed for both weak and strong interference regimes provide secrecy for the message of each transmitter against all receivers except its respective one. More precisely, it is easy to show that the equivocation rates are upper bounded by*

$$\frac{1}{T}I(W_k; y_j^T) \leq R - \epsilon_T, \qquad k \neq j. \tag{50}$$

The main intuition behind this is the following. Each receiver can only decode its own message, as well as the sum-lattice codeword corresponding to the message of other users. For instance, after decoding W_1, $\mathbf{Rx}_1$ *remains with a codeword that depends on* $W_2, W_3, \ldots, W_K$. *Hence,* $W_3, \ldots, W_K$ *act as a mask (encryption key) to hide* W_2 *from* $\mathbf{Rx}_1$. *Therefore, although* $\mathbf{Rx}_1$ *receives a certain amount of information about a function of all other messages, the amount of information it gets about each unintended individual message is negligible. This phenomenon is very similar to the encoding scheme used in [21] to guarantee information-secrecy. However, here this secrecy is naturally provided by the coding scheme, without any additional penalty in terms of the symmetric achievable rate of the network.*

VI. The Gaussian Network: An Upper Bound

In this section we prove the converse part of Theorem 3. To this end, we derive an upper bounds on the symmetric rate of the network. The essence of this bound is the same as the converse proof for the deterministic network. That is, in the strong interference regime, given all the messages except for two of them, the output signal of any of the respective receivers is not only sufficient to decode its own message, but can also be used to decode the other missing message. Similarly, in the weak interference regime, although one receiver cannot completely decode the message of the other transmitter, it receives enough information to partially decode that message.

We first define $\tilde{z}_{it} = z_{it} - z_{2t}$ for $i = 3, 4, \ldots, K$ and $t = 1, \ldots, T$. Then, we can write

$$\begin{aligned}
T(R_1 + R_2) &\leq H(W_1) + H(W_2) = H(W_1, W_2 | W_3, \ldots, W_K) \quad (51)\\
&= H(W_2 | W_3, \ldots, W_K) + H(W_1 | W_2, W_3, \ldots, W_K)\\
&= I(W_2; y_2^T | W_3, \ldots, W_K) + H(W_2 | y_2^T, W_3, \ldots, W_K)\\
&\quad + I(W_1; y_1^T y_2^T | W_2, W_3, \ldots, W_K) + H(W_1 | y_1^T y_2^T, W_2, W_3, \ldots, W_K)\\
&\leq I(W_2; y_2^T, \tilde{z}_3^T, \ldots, \tilde{z}_K^T | W_3, \ldots, W_K) + I(W_1; y_1^T y_2^T, \tilde{z}_3^T, \ldots, \tilde{z}_K^T | W_2, W_3, \ldots, W_K) + 2T\epsilon_T\\
&= h(y_2^T, \tilde{z}_3^T, \ldots, \tilde{z}_K^T | W_3, \ldots, W_K) - h(y_2^T, \tilde{z}_3^T, \ldots, \tilde{z}_K^T | W_2, W_3, \ldots, W_K)\\
&\quad + h(y_1^T y_2^T, \tilde{z}_3^T, \ldots, \tilde{z}_K^T | W_2, W_3, \ldots, W_K) - h(y_1^T y_2^T, \tilde{z}_3^T, \ldots, \tilde{z}_K^T | W_1, W_2, W_3, \ldots, W_K) + 2T\epsilon_T\\
&= h(y_2^T, \tilde{z}_3^T, \ldots, \tilde{z}_K^T | W_3, \ldots, W_K) + h(y_1^T | y_2^T, \tilde{z}_3^T, \ldots, \tilde{z}_K^T, W_2, W_3, \ldots, W_K)\\
&\quad - h(y_1^T y_2^T, \tilde{z}_3^T, \ldots, \tilde{z}_K^T | W_1, W_2, W_3, \ldots, W_K) + 2T\epsilon_T, \quad (52)
\end{aligned}$$

where ϵ_T vanishes as T grows. Note that we used independence of the messages in (51). We can bound each term in (52) individually. The first term can be bounded as

$$\begin{aligned}
&h(y_2^T, \tilde{z}_3^T, \ldots, \tilde{z}_K^T | W_3, \ldots, W_K) \leq h(y_2^T) + h(\tilde{z}_3^T) + \cdots + h(\tilde{z}_K^T)\\
&\quad \leq T h(y_2) + \frac{T(K-2)}{2} \log(4\pi e) \quad (53)\\
&\quad \leq \frac{T}{2} \log \left(1 + \mathsf{SNR} + (K-1)\mathsf{INR} + 2\sqrt{\mathsf{SNR} \cdot \mathsf{INR}} \sum_{j \neq 2} \rho_{2j} + 2\mathsf{INR} \sum_{\substack{i > j \\ i,j \neq 2}} \rho_{ij} \right) + \frac{T(K-1)}{2} \log(4\pi e)\\
&\quad \leq \frac{T}{2} \log \left(1 + \mathsf{SNR} + (K-1)^2 \mathsf{INR} + 2(K-1)\sqrt{\mathsf{SNR} \cdot \mathsf{INR}} \right) + \frac{T(K-1)}{2} \log(4\pi e), \quad (54)
\end{aligned}$$

where $\rho_{ij} \in [-1, 1]$ is the correlation coefficient between channel inputs x_i and x_j. In (53) we used the fact that $\mathbb{E}[\tilde{z}_i^2] = 2$.

Bounding the second term is more involved. First note that

$$\begin{aligned}
&I(y_1^T; y_3^T, \ldots, y_K^T | y_2^T, \tilde{z}_3^T, \ldots, \tilde{z}_K^T, W_2, W_3, \ldots, W_K)\\
&\quad = \sum_{t=1}^{T} I(y_1^T; y_{3t}, \ldots, y_{Kt} | y_2^T, \tilde{z}_3^T, \ldots, \tilde{z}_K^T, W_2, W_3, \ldots, W_K, y_3^{t-1}, \ldots, y_K^{t-1})\\
&\quad = \sum_{t=1}^{T} I(y_1^T; y_{3t}, \ldots, y_{Kt} | y_2^T, \tilde{z}_3^T, \ldots, \tilde{z}_K^T, W_2, W_3, \ldots, W_K, y_3^{t-1}, \ldots, y_K^{t-1}, x_{2t}, x_{3t}, \ldots, x_{Kt}) \quad (55)\\
&\quad = 0 \quad (56)
\end{aligned}$$

where (55) holds since for $j = 2, \ldots, K$, $x_{jt} = f_{jt}(W_j, y_j^{t-1})$ is a deterministic function of the message and channel output; the last equality in (56) is due to the fact that for $j = 3, \ldots, K$, we have

$$\begin{aligned}
y_{jt} &= \sqrt{\mathsf{SNR}}x_{jt} + \sqrt{\mathsf{INR}}\sum_{i\notin\{2,j\}} x_{it} + \sqrt{\mathsf{INR}}x_{2t} + z_{jt} \\
&= \left[\sqrt{\mathsf{SNR}}x_{2t} + \sqrt{\mathsf{INR}}\sum_{i\notin\{2,j\}} x_{it} + \sqrt{\mathsf{INR}}x_{jt} + z_{2t}\right] - \sqrt{\mathsf{SNR}}x_{2t} + (\sqrt{\mathsf{SNR}} - \sqrt{\mathsf{INR}})x_{jt} + (z_{jt} - z_{2t}) \\
&= y_{2t} - \sqrt{\mathsf{SNR}}x_{2t} + (\sqrt{\mathsf{SNR}} - \sqrt{\mathsf{INR}})x_{jt} + \tilde{z}_{jt}, \qquad (57)
\end{aligned}$$

which implies that y_{jt} can be deterministically recovered from $(y_{2t}, x_{2t}, x_{jt}, \tilde{z}_{jt})$. Hence, each term in (56) is zero. From (56) we can bound the second term in (52) as

$$\begin{aligned}
h(y_1^T|y_2^T, \tilde{z}_3^T, \ldots, \tilde{z}_K^T, W_2, W_3, \ldots, W_K) &= h(y_1^T|y_2^T, y_3^T, \ldots, y_K^T, \tilde{z}_3^T, \ldots, \tilde{z}_K^T, W_2, W_3, \ldots, W_K) \\
&= h(y_1^T|y_2^T, y_3^T, \ldots, y_K^T, \tilde{z}_3^T, \ldots, \tilde{z}_K^T, W_2, W_3, \ldots, W_K, x_2^T, \ldots, x_K^T) \\
&\leq h(\sqrt{\mathsf{SNR}}x_1^T - \sqrt{\mathsf{INR}}\sum_{i\neq 1} x_i^T + z_1^T | y_2^T - \sqrt{\mathsf{SNR}}x_2^T - \sqrt{\mathsf{INR}}\sum_{j>2} x_j^T, x_2^T, \ldots, x_K^T) \\
&\leq h(\sqrt{\mathsf{SNR}}x_1^T + z_1^T | \sqrt{\mathsf{INR}}x_1^T + z_2^T) \\
&\leq \frac{T}{2}\log\left(1 + \frac{\mathsf{SNR}}{1+\mathsf{INR}}\right) + \frac{T}{2}\log(2\pi e). \qquad (58)
\end{aligned}$$

Finally, we can bound the third term in (52) as follows:

$$\begin{aligned}
&h(y_1^T, y_2^T, \tilde{z}_3^T, \ldots, \tilde{z}_K^T|W_1, W_2, W_3, \ldots, W_K) \\
&= \sum_{t=1}^{T} h(y_{1t}, y_{2t}, \tilde{z}_{3t}, \ldots, \tilde{z}_{Kt}|y_1^{t-1}, y_2^{t-1}, \tilde{z}_3^{t-1}, \ldots, \tilde{z}_K^{t-1}, W_1, W_2, W_3, \ldots, W_K) \\
&\geq \sum_{t=1}^{T} h(y_{1t}, y_{2t}, \tilde{z}_{3t}, \ldots, \tilde{z}_{Kt}|y_1^{t-1}, y_2^{t-1}, \tilde{z}_3^{t-1}, \ldots, \tilde{z}_K^{t-1}, W_1, W_2, W_3, \ldots, W_K, x_{1t}, \ldots, x_{Kt}) \\
&= \sum_{t=1}^{T} h(z_{1t}, z_{2t}, \tilde{z}_{3t}, \ldots, \tilde{z}_{Kt}|y_1^{t-1}, y_2^{t-1}, \tilde{z}_3^{t-1}, \ldots, \tilde{z}_K^{t-1}, W_1, W_2, W_3, \ldots, W_K, x_{1t}, \ldots, x_{Kt}) \\
&= \sum_{t=1}^{T} h(z_{1t}, z_{2t}, \tilde{z}_{3t}, \ldots, \tilde{z}_{Kt}) \qquad (59) \\
&= \sum_{t=1}^{T} h(z_{1t}, z_{2t}, z_{3t}, \ldots, z_{Kt}) \\
&= \frac{TK}{2}\log(2\pi e), \qquad (60)
\end{aligned}$$

where (59) is due to the facts that the channels are memoryless and the noise at time t is independent of all the messages and signals and noises in the past. Substituting (54), (58) and (60) in (52), and recalling the fact that we are interested in the maximum $R_1 = R_2 = R_{\text{sym}}$, we get

$$R_{\text{sym}} \leq \frac{1}{4}\log\left(1 + \mathsf{SNR} + (K-1)^2\mathsf{INR} + 2(K-1)\sqrt{\mathsf{SNR}\cdot\mathsf{INR}}\right) + \frac{1}{4}\log\left(1 + \frac{\mathsf{SNR}}{1+\mathsf{INR}}\right) + \frac{K-1}{4}.$$

This bound can be further simplified as follows. It is easy to show that

$$\mathsf{SNR}+(K-1)^2\mathsf{INR}+2(K-1)\sqrt{\mathsf{SNR}\cdot\mathsf{INR}}=\left(\sqrt{\mathsf{SNR}}+(K-1)\sqrt{\mathsf{INR}}\right)^2\leq K^2(\mathsf{SNR}+\mathsf{INR}) \tag{61}$$

which implies

$$\begin{aligned} R_{\text{sym}} &\leq \frac{1}{4}\log(1+\mathsf{SNR}+(K-1)^2\mathsf{INR}+2(K-1)\sqrt{\mathsf{SNR}\cdot\mathsf{INR}})+\frac{1}{4}\log\left(1+\frac{\mathsf{SNR}}{1+\mathsf{INR}}\right)+\frac{K-1}{4} \\ &\leq \frac{1}{4}\log(1+\mathsf{SNR}+\mathsf{INR})+\frac{1}{4}\log\left(1+\frac{\mathsf{SNR}}{\mathsf{INR}}\right)+\frac{K-1}{4}+\frac{1}{2}\log K, \end{aligned} \tag{62}$$

which is the desired bound.

VII. The Generalized Degrees of Freedom

In this section we prove Theorem 1. The proof for $\alpha \neq 1$ is straight-forward from Theorem 3 as follows. Recall the achievable symmetric rate in Theorem 3. Hence,

$$\begin{aligned} d_{\text{FB}}(\alpha) &= \limsup_{\mathsf{SNR}\to\infty}\frac{R_{\text{sym}}(\mathsf{SNR},\alpha)}{\frac{1}{2}\log(\mathsf{SNR})} \\ &= \limsup_{\mathsf{SNR}\to\infty}\frac{\frac{1}{4}\log(1+\mathsf{SNR}+\mathsf{SNR}^{\alpha})+\frac{1}{4}\log(1+\mathsf{SNR}^{1-\alpha})}{\frac{1}{2}\log(\mathsf{SNR})} \\ &= \frac{1}{2}\max\{1,\alpha\}+\frac{(1-\alpha)^+}{2} \\ &= \begin{cases} 1-\frac{\alpha}{2} & \alpha<1 \\ \frac{\alpha}{2} & \alpha>1. \end{cases} \end{aligned}$$

The number of generalized degrees of freedom is discontinuous at $\alpha = 1$, and the proof in that case follows from a different argument. Note that $d(1) = 1/K$, or equivalently $R_{\text{sym}} = \frac{1}{2}\log(1+\mathsf{SNR})$ can be easily achieved by time-sharing between the users: during the block k, user k encodes and sends its message, while all the other transmitters keep silent.

In order to show optimality of this number of degrees of freedom for $\mathsf{INR} = \mathsf{SNR}$, we use the cut-set bound. This gives us a tighter bound, which is similar to that of the deterministic network for $m = n$. A similar intuition can explain this phenomenon: when the gain of the direct and cross links are the same, the output signals at all receivers are statistically equivalent, and given any of them, the uncertainty in the others is small. We can formally write

$$\begin{aligned}
TKR_{\text{sym}} &= T\sum_{k=1}^{K} R_k = H(W_1,\ldots,W_K)\\
&\le I(y_1^T,\ldots,y_K^T;W_1,\ldots,W_K) + KT\epsilon_T\\
&= \sum_{t=1}^{T} I(y_{1t},\ldots,y_{Kt};W_1,\ldots,W_K|y_1^{t-1},\ldots,y_K^{t-1}) + KT\epsilon_T\\
&= \sum_{t=1}^{T}\left[h(y_{1t},\ldots,y_{Kt}|y_1^{t-1},\ldots,y_K^{t-1}) - h(y_{1t},\ldots,y_{Kt}|W_1,\ldots,W_K,y_1^{t-1},\ldots,y_K^{t-1})\right] + KT\epsilon_T\\
&\le \sum_{t=1}^{T}\left[h(y_{1t},\ldots,y_{Kt}) - h(y_{1t},\ldots,y_{Kt}|W_1,\ldots,W_K,y_1^{t-1},\ldots,y_K^{t-1},x_{1t},\ldots,x_{Kt})\right] + KT\epsilon_T \qquad (63)\\
&= \sum_{t=1}^{T}\Big[h(y_{1t},y_{2t}-y_{1t},y_{3t}-y_{1t},\ldots,y_{Kt}-y_{1t})\\
&\qquad - h(z_{1t},\ldots,z_{Kt}|W_1,\ldots,W_K,y_1^{t-1},\ldots,y_K^{t-1},x_{1t},\ldots,x_{Kt})\Big] + KT\epsilon_T\\
&\le \sum_{t=1}^{T}\left[h(y_{1t}) + h(y_{2t}-y_{1t}) + \cdots + h(y_{Kt}-y_{1t})\right] - \sum_{t=1}^{T}\sum_{k=1}^{K} h(z_{kt}) + KT\epsilon_T \qquad (64)\\
&= \sum_{t=1}^{T}\left[h(y_{1t}) + h(z_{2t}-z_{1t}) + \cdots + h(z_{Kt}-z_{1t})\right] - \sum_{t=1}^{T}\sum_{k=1}^{K} h(z_{kt}) + KT\epsilon_T \qquad (65)\\
&\le \frac{T}{2}\log\left((2\pi e)\left(1+(\sqrt{\mathsf{SNR}}+(K-1)\sqrt{\mathsf{INR}})^2\right)\right) + \sum_{j=2}^{K}\frac{T}{2}\log(4\pi e) - \frac{KT}{2}\log(2\pi e) + KT\epsilon_T\\
&\le \frac{T}{2}\log\left(1+K\mathsf{SNR}\right) + \frac{(K-1)T}{2} + KT\epsilon_T, \qquad (66)
\end{aligned}$$

where (63) holds since $x_{kt} = f_{kt}(W_k, y_k^{t-1})$; in (64) we used the fact that in a memoryless channel noise terms in time t are independent of all variables in the past; and (65) follows from $y_k - y_1 = z_k - z_1$, for $k = 2,\ldots,K$. Dividing by KT, we get

$$R_{\text{sym}} \le \frac{1}{K}\log(1+K\mathsf{SNR}) + \frac{K-1}{2K},$$

which implies $d(1) \le \frac{1}{K}$. This completes the proof of the theorem.

Remark 3. *Note that the approximate capacity characterization in Theorem 3 is only valid for* $\alpha \neq 1$. *In fact this result does not cover the behavior of the capacity when* $\mathsf{INR} = \mathsf{SNR}(1+\zeta(\mathsf{SNR}))$ *with* $\zeta(\mathsf{SNR}) \to 0$ *as* $\mathsf{SNR} \to \infty$. *For such regime, the gap between our outer bound and the achievable rate is not constant. However, since in the study of the generalized degrees of freedom we only allow a specific growth for* INR *in terms of* SNR, *such regime is excluded by definition, and we have a complete characterization of the* GDoF *with feedback.*

VIII. CONCLUSION

We have studied the feedback capacity of the fully connected K-user interference channel under a symmetric topology. This is a natural extension of the feedback capacity characterization for the 2-user case in [15], in which it is shown that channel output feedback can significantly improve the performance of the 2-user interference channel. Rather surprisingly, it turns out that such an improvement can also be achieved in the K-user case, except if the intended and interfering signals have the same received power at the receivers. In particular, we have shown that the per-user feedback capacity of the K-user FC-IC is as if there were only one source of interference in the network. Compared to the network without feedback [17], this result shows that feedback can significantly improve the network capacity.

The coding scheme used to achieve the capacity of the network combines two well-known interference management techniques, namely, interference alignment and interference decoding. In fact, the messages at the transmitters are encoded such that the $K-1$ interfering signals are received aligned at each receiver. Closedness of lattice codes with respect to summation implies that the aligned received interference is a codeword that can be decoded, as in the 2-user case. Another interesting aspect of this scheme is that each message is kept secret from all receivers, except the intended one. This implies that an appropriately defined secrecy capacity of the network coincides with the capacity with no secrecy constraint.

APPENDIX A

CODING SCHEMES FOR THE DETERMINISTIC NETWORK: ARBITRARY (n, m)

A. *Weak Interference Regime* ($m < n$)

In the following, we generalize the coding scheme presented in Fig. 3 for arbitrary parameters m and n. Denote the message of user k which will be transmitted in 2 channel uses by a p-ary sequence of length $2R_{\text{sym}}$, namely, $\mathbf{S}_k = [S_k(1), \dots, S_k(2n-m)]$. Each user sends $p = n$ fresh symbols over its first channel use, i.e.,

$$X_{k1} = \begin{bmatrix} S_k(1) & S_k(2) & \cdots & S_k(n) \end{bmatrix}',$$

where X' denotes the transpose of the matrix X. The signal received at the $\mathbf{Rx}_k$ can be split into two parts, the part above the interference level which contains $(n-m)$ interference free symbols, and the lower m symbols which is a combination of the intended symbols and interference,

$$Y_{k1} = \begin{bmatrix} S_k(1) & \dots & S_k(n-m) & S_k(n-m+1) + S_{\sim k}(1) & \dots & S_k(n) + S_{\sim k}(m) \end{bmatrix}',$$

where $S_{\sim k}(j) = \sum_{i \neq k} S_i(j)$ is the summation of all p-ary symbols sent by all the base stations except $\mathbf{Tx}_k$. This received signal is sent to the transmitter via the feedback link. Transmitter $\mathbf{Tx}_k$ first removes its own signal from this feedback signal, and then forwards the remaining symbols on its top most m levels. It also transmits $(n-m)$ new fresh symbols over its lower levels:

$$X_{k2} = \begin{bmatrix} S_{\sim k}(1) & \dots & S_{\sim k}(m) & S_k(n+1) & \dots & S_k(2n-m) \end{bmatrix}'.$$

A similar operation is performed at all other transmitters, which results in a received signal at $\mathbf{Rx}_k$ of the form

$$\begin{aligned} Y_{k2} &= X_{k2} + D^{n-m}\sum_{i\neq k} X_{i2} \\ &= \begin{bmatrix} S_{\sim k}(1) \\ \vdots \\ S_{\sim k}(m) \\ S_k(n+1) \\ \vdots \\ S_k(2n-m) \end{bmatrix} + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \sum_{i\neq k} S_{\sim i}(1) \\ \vdots \\ \sum_{i\neq k} S_{\sim i}(m) \end{bmatrix} = \begin{bmatrix} S_{\sim k}(1) \\ \vdots \\ S_{\sim k}(m) \\ S_k(n+1) \\ \vdots \\ S_k(2n-m) \end{bmatrix} + (K-1) \begin{bmatrix} 0 \\ \vdots \\ 0 \\ S_k(1) \\ \vdots \\ S_k(m) \end{bmatrix} + (K-2) \begin{bmatrix} 0 \\ \vdots \\ 0 \\ S_{\sim k}(1) \\ \vdots \\ S_{\sim k}(m) \end{bmatrix}. \end{aligned} \quad (67)$$

We used the fact that $\sum_{i\neq k} S_{\sim i}(j) = (K-1)S_k(j) + (K-2)S_{\sim k}(j)$ in the last equality. Having Y_{k1} and Y_{k2}, receiver $\mathbf{Rx}_k$ wishes to decode $\mathbf{S}_k$. Note that we have a linear system with $2n$ equations and $2n$ variables (including m variables $S_{\sim k}(j)$ for $j = 1, \ldots, m$ and $2n-m$ variables including $S_k(j)$ for $j = 1, \ldots, 2n-m$), which can be uniquely solved[2]. Therefore, $\mathbf{Rx}_k$ can recover all its $2n-m$ symbols transmitted by $\mathbf{Tx}_k$, which implies a communication rate of $R_k = (2n-m)/2$. Note that the encoding operations at all transmitters are the same, and hence, a similar rate can be achieved for all pairs by applying a similar decoding.

B. Strong Interference Regime ($m > n$)

Similar to the weak interference regime, this scheme is performed over two consecutive time instances, and provides a total of m information symbols for each user. Denote the message of user k by a $\mathbf{S}_k = [S_k(1), \ldots, S_k(m)]$, which is a p-ary sequence of length m. In the first time instance, each user broadcasts its entire message,

$$X_{k1} = \begin{bmatrix} S_k(1) & \cdots & S_k(m) \end{bmatrix}',$$

which implies the received signal at $\mathbf{Rx}_k$ to be

$$Y_{k1} = D^{m-n}\mathbf{S}_k + \mathbf{S}_{\sim k} = \begin{bmatrix} S_{\sim k}(1) & \cdots & S_{\sim k}(m-n) & S_k(1) + S_{\sim k}(m-n+1) & \cdots & S_k(n) + S_{\sim k}(m) \end{bmatrix}'.$$

This output is sent to the transmitter through the feedback link. In the second time slot, the transmitter simply removes its signal and forwards the remaining, that is,

$$X_{k2} = \begin{bmatrix} S_{\sim k}(1) & \cdots & S_{\sim k}(m-n) & S_{\sim k}(m-n+1) & \cdots & S_{\sim k}(m) \end{bmatrix}'.$$

[2]It is easy to verify that the coefficient matrix is full-rank.

Hence, we have

$$Y_{k2} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ S_{\sim k}(1) \\ \vdots \\ S_{\sim k}(n) \end{bmatrix} + \begin{bmatrix} \sum_{i\neq k} S_{\sim i}(1) \\ \\ \vdots \\ \\ \\ \sum_{i\neq k} S_{\sim i}(m) \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ S_{\sim k}(1) \\ \vdots \\ S_{\sim k}(n) \end{bmatrix} + (K-1) \begin{bmatrix} S_k(1) \\ \\ \vdots \\ \\ \\ S_k(m) \end{bmatrix} + (K-2) \begin{bmatrix} S_{\sim k}(1) \\ \\ \vdots \\ \\ \\ S_{\sim k}(m) \end{bmatrix} \tag{68}$$

$$= (K-1)\mathbf{S}_k + (D^{m-n} + (K-2)I_m)\mathbf{S}_{\sim k}. \tag{69}$$

Having Y_{k1} and Y_{k2} together, $\mathbf{Rx}_k$ has a linear system with $2m$ equation and $2m$ variables (including m variables in $\mathbf{S}_k$ and m variables in $\mathbf{S}_{\sim k}$):

$$\begin{bmatrix} Y_{k1} \\ Y_{K2} \end{bmatrix} = \begin{bmatrix} D^{m-n} & I_m \\ (K-1)I_m & D^{m-n} + (K-2)I_m \end{bmatrix} \begin{bmatrix} \mathbf{S}_k \\ \mathbf{S}_{\sim k} \end{bmatrix}. \tag{70}$$

This system has a unique solution if and only if the coefficient matrix is full-rank, which holds if and only if $K \not\equiv 1(\mod q)$, which can be easily satisfied for a proper choice[3] of p. Fig. 4 pictorially demonstrates this coding scheme for 3-user case.

APPENDIX B
PROOF OF LEMMA 1

Having the dither vector available, the receiver first computes

$$\begin{aligned} \tilde{\mathbf{y}} &= [\gamma \mathbf{y} + \mathbf{d}] \mod \Lambda_q = [\mathbf{s} + \mathbf{d} + (\gamma\alpha - 1)\mathbf{s} + \gamma \mathbf{z}] \mod \Lambda_q \\ &= [\mathbf{c} + (\gamma\alpha - 1)\mathbf{s} + \gamma \mathbf{z}] \mod \Lambda_q, \end{aligned}$$

where γ is a free parameter which will be fixed later. The receiver then decodes $\mathbf{c}$ from $\tilde{\mathbf{y}}$ by treating $\mathbf{z}' = [(\gamma\alpha - 1)\mathbf{s} + \gamma \mathbf{z}] \mod \Lambda_q$ as noise. Note that

$$\begin{aligned} \frac{1}{n}\mathbb{E}[\mathbf{z}'^2] &\leq \frac{1}{n}\mathbb{E}\left[|(\gamma\alpha - 1)\mathbf{s} + \gamma \mathbf{z}|^2\right] \\ &= \frac{1}{n}(\gamma\alpha - 1)^2\sigma^2(\Lambda_q) + \gamma^2\beta^2 \\ &= \frac{\beta^2}{\alpha^2 + \beta^2} \end{aligned}$$

[3]Note that this result does not necessarily holds for all values of p and K. For instance, this approach does not give a set of independent linear equations for the 3-user case over the binary field. However, the encoding scheme for larger field size ($p > 2$) still reveals valuable insights for the Gaussian channel.

where the last equality is due to the choice of $\gamma = \alpha/(\alpha^2 + \beta^2)$. On the other hand, $\tilde{\mathbf{y}}$ is uniformly distributed over $\mathcal{V}_q$. Therefore,

$$\frac{1}{n} I(\mathbf{y};\mathbf{c}) \geq \frac{1}{n} I(\tilde{\mathbf{y}};\mathbf{c}) = \frac{1}{n} h(\tilde{\mathbf{y}}) - \frac{1}{n} h(\mathbf{z}') \tag{71}$$

$$\geq \frac{1}{2} \log \frac{1}{G(\Lambda_q)} - \frac{1}{2} \log \left(2\pi e \frac{\beta^2}{\alpha^2 \beta^2} \right) \tag{72}$$

$$= \frac{1}{2} \log \left(1 + \frac{\alpha^2}{\beta^2} \right), \tag{73}$$

where, the last equality holds since Λ_q is a good quantization lattice and $G(\Lambda_q) \to \frac{1}{2\pi e}$ as n grows. Hence, the message can be decoded as its rate does not exceed this mutual information.

REFERENCES

[1] T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Information Theory*, vol. 27, pp. 49–60, Jan. 1981.

[2] R. H. Etkin, D. Tse, and H. Wang, "Gaussian interference channel capacity to within one bit," *IEEE Trans. Information Theory*, vol. 54, no. 12, pp. 5534–5562, Dec. 2008.

[3] E. Telatar and D. Tse, "Bounds on the capacity region of a class of interference channels," in *Proc. International Symposium on Information Theory (ISIT)*, Nice, France, 2007, pp. 2871–2874.

[4] G. Bresler and D. Tse, "The two-user Gaussian interference channel: a deterministic view," *European Trans. Telecommunications*, vol. 19, pp. 333–354, June 2008.

[5] A. S. Avestimehr, S. N. Diggavi, and D. N. C. Tse, "Wireless network information flow: A deterministic approach," *IEEE Trans. Information Theory*, vol. 57, pp. 1872–1905, April 2011.

[6] M. Maddah-Ali, A. Motahari, and A. Khandani, "Communication over MIMO X channels: Interference alignment, decomposition, and performance analysis," *IEEE Trans. Information Theory*, vol. 54, no. 8, pp. 3457–3470, August 2008.

[7] V. Cadambe and S. Jafar, "Interference Alignment and Degrees of Freedom of the K-User Interference Channel," *IEEE Trans. Information Theory*, vol. 54, no. 8, pp. 3425–3441, August 2008, .

[8] S. Mohajer, S. N. Diggavi, C. Fragouli, and D. N. C. Tse, "Transmission techniques for relay-interference networks," in *Proc. the 46 th Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, USA, Sep. 2008.

[9] ——, "Approximate capacity characterization for a class of Gaussian relay-interference wireless networks," *IEEE Trans. Information Theory*, vol. 57, no. 5, pp. 2837–2864, May 2011.

[10] C. Shannon, "The zero error capacity of a noisy channel," *IRE Trans. Information Theory*, vol. 2, no. 3, pp. 8–19, 1956.

[11] A. El Gamal and Y. Kim, "Lecture notes on network information theory," *Arxiv preprint arXiv:1001.3404*, 2010.

[12] G. Kramer, "Feedback strategies for white Gaussian interference networks," *IEEE Trans. Information Theory*, vol. 48, no. 6, pp. 1423–1438, 2002.

[13] M. Gastpar and G. Kramer, "On noisy feedback for interference channels," in *Proc. Asilomar Conference on Signals, Systems and Computers (ACSSC)*, Pacific Grove, CA, USA, 2006, pp. 216–220.

[14] R. Tandon and S. Ulukus, "Dependence balance based outer bounds for Gaussian networks with cooperation and feedback," *IEEE Trans. Information Theory*, vol. 57, no. 7, pp. 4063–4086, July 2011.

[15] C. Suh and D. Tse, "Feedback capacity of the Gaussian interference channel to within 2 bits," *IEEE Trans. Information Theory*, vol. 57, no. 5, pp. 2667–2685, May 2011.

[16] R. Tandon, S. Mohajer, and H. V. Poor, "On the symmetric feedback capacity of the K-user cyclic Z-interference channel," *submitted to IEEE Trans. Information Theory*, Sep. 2011, arxiv preprint, arXiv:1109.1507.

[17] S. Jafar and S. Vishwanath, "Generalized degrees of freedom of the symmetric Gaussian K-user interference channel," *IEEE Trans. Information Theory*, vol. 56, no. 7, pp. 3297–3303, July 2010.

[18] R. Zamir, S. Shamai (Shitz), and U. Erez, “Nested linear/lattice codes for structured multiterminal binning,” *IEEE Trans. Information Theory*, vol. 48, no. 6, pp. 1250–1276, June 2002.

[19] B. Nazer and M. Gastpar, “Computation over multiple-access channels,” *IEEE Trans. Information Theory*, vol. 53, no. 10, pp. 3498–3516, 2007.

[20] A. Jafarian, J. Jose, and S. Vishwanath, “Algebraic lattice alignment for K-user interference channels,” in *Proc. the 47 th Annual Allerton Conference on Communication, Control, and Computing*. Monticello, IL, USA: IEEE, 2009, pp. 88–93.

[21] S. Mohajer, S. Diggavi, H. V. Poor, and S. Shamai (Shitz), “On the parallel relay wire-tap network,” in *Proc. the 49 th Annual Allerton Conference on Communication, Control, and Computing*, Monticello, IL, USA, Sep. 2011.