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

Soheil Mohajer, Ravi Tandon, Member, IEEE, and H. Vincent Poor, Fellow, IEEE

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 (GDoF) is characterized in terms of α , where the interferenceto-noise ratio (INR) is given by $INR = SNR^{\alpha}$. It is shown that the per-user GDoF 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 GDoF. 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 deal with all interfering signals in order to decode its intended message. Managing such

Manuscript received October 21 2011; revised May 24 2012; and accepted December 14 2012.

This paper was presented in part at IEEE International Symposium on Information Theory (ISIT), Cambridge, USA, 2012.

Soheil Mohajer was with the Department of Electrical Engineering, Princeton University, Princeton, NJ, USA. He is now with the Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, Berkeley, CA, USA (E-mail: mohajer@eecs.berkeley.edu).

Ravi Tandon was with the Department of Electrical Engineering, Princeton University, Princeton, NJ, USA. He is now with the Department of Electrical and Computer Engineering, Virginia Tech, Blacksburg, VA, USA. (E-mail: tandonr@vt.edu).

H. Vincent Poor is with the Department of Electrical Engineering, Princeton University, Princeton, NJ, USA. (E-mail: poor@princeton.edu).

The work was supported in part by the Air Force Office of Scientific Research under MURI Grant FA-9550-09-1-0643 and in part by the DTRA under Grant HDTRA-07-1-0037. The work of Soheil Mohajer is partially supported by The Swiss National Science Foundation under Grant PBELP2-133369.

Copyright (c) 2012 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org. interfering signals in a multi-user network is a long standing and fundamental problem in wireless communication.

The simplest example in this category is the 2-user interference channel [1], in which two transmitters with independent messages wish to communicate with their respective receivers over the wireless transmission medium. Even for this simple 2user network, the complete information-theoretic characterization of the capacity region 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 informationtheoretic 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.

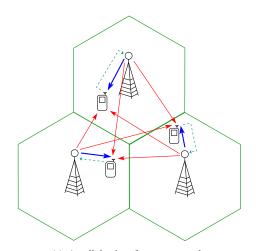
A similar approximate characterization (with a larger gap) for this problem is developed in [4], in which both coding scheme and bounding techniques are devised by studying the problem under the *deterministic* model. This framework, introduced by Avestimehr, Diggavi, and Tse in [5], focuses on complex 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] 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.

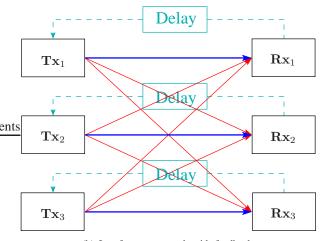
It is well known that feedback does not increase the capacity of point-to-point discrete memoryless channels [8]. However, feedback is beneficial in improving the capacity regions of more complex networks (see [9] and references therein). The effects of feedback on the capacity region of the interference channel have been studied in several papers. Feedback coding schemes for K-user Gaussian interference networks have been developed by Kramer in [10]. Outer bounds for the 2-user interference channel with generalized feedback have been derived in [11] and [12]. The effect of feedback on the capacity of the 2-user interference channel is studied in [13], where it is shown that feedback provides multiplicative gain in the capacity at high signal-to-noise ratio (SNR), when the interference links are much stronger than the direct links. 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 [14]. This includes all regimes of interference, and finite and asymptotic regimes of SNR. The gap between the capacity of the channel with and without feedback can be arbitrarily large for certain channel parameters. The key technique for the strong interference regime 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 $\mathbf{R}\mathbf{x}_1$, can be sent either through the direct link $Tx_1 \rightarrow Rx_1$, or the cyclic path $\mathbf{Tx}_1 \rightarrow \mathbf{Rx}_2 \rightarrow \mathbf{Tx}_2 \rightarrow \mathbf{Rx}_1$. In particular, the advantage of such artificial paths can be clearly understood when the cross links are much stronger than the direct links (**PS**fragereprincements interference regime). This observation becomes very natural by studying the problem under the deterministic framework.

The first extension of [14] 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 [15]. 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. A similar setting without feedback has been studied by Jafar and Vishwanath in [16], where the number of symmetric degrees of freedom is characterized. An approximate sum capacity of this network is recently found by Ordentlich et al. in [17]. In this paper,



(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).

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 [15], 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 (SNR = 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 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 of interest. To this end, we propose using a common structured code, which has the property that the summation of codewords 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. We further extend the result of the paper, and study the case of global feedback, where each transmitter receives feedback from all receivers in Section VIII, and finally, conclude the paper in Section IX. In order to make the paper easily readable, some of the technical proofs are postponed to the appendices. Parts of this work have been presented in [21].

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

$$y_k(t) = \sqrt{\mathsf{SNR}} x_k(t) + \sum_{\substack{i=1\\i \neq k}}^K \sqrt{\mathsf{INR}} x_i(t) + z_k(t), \qquad (1)$$

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.,

$$x_{kt} = g_{kt}(W_k, y_{k1}, y_{k2}, \dots, y_{k(t-1)}) = g_{kt}(W_k, y_k^{t-1}), \quad (2)$$

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 $\mathbf{R}\mathbf{x}_k$ up to time t-1.

A rate tuple (R_1, R_2, \ldots, 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 factors¹, namely SNR and INR, we use the standard parameter α (as in [2] and [16]) 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

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

¹The notion of degrees of freedom (DoF) captures the asymptotic behavior of the capacity, where the transmit power grows to infinity. However, this forces all channels to be equally strong, i.e., all the power of all received signals from different links grow at the same rate. Therefore, it is not very insightful towards finding optimal transmission schemes when some signals are significantly stronger or weaker than others. The generalized degrees of freedom which allows different rate of growth for SNR and INR is more useful metric in such scenarios. We refer the reader to [22] for a comprehensive discussion on these metrics.

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

$$d_{\rm FB}(\alpha) = \begin{cases} 1 - \frac{\alpha}{2} & \alpha < 1 \text{ (weak interference)} \\ not \text{ well-defined} & \alpha = 1 \\ \frac{\alpha}{2} & \alpha > 1 \text{ (strong interference)} \end{cases}$$

We present the proof of Theorem 1 in Section VII. Note that the theorem above does not characterize the GDoF for $\alpha = 1$. In fact for this regime, the GDoF is not well-defined and can get different values, depending on mutual growth of SNR and INR. We refer the interested reader to Section VII for a detailed discussion.

In order to demonstrate the benefit gained by output feedback, we present the following theorem from [16], which characterizes the GDoF for the FC-IC without feedback.

Theorem 2 ([16], Theorem 3.1). The per-user GDoF for the K-user interference channel without feedback is given by

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

The generalized degrees of freedom of the K-user interference channel with/without feedback are illustrated in Figure 2. As derived in [16], 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 [14], except for $\alpha = 1$. At this particular point, the GDoF can be bounded from below and above by $\frac{1}{K}$ and $\frac{1}{2}$, respectively.

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

Theorem 3. The symmetric capacity of the K user interference channel with feedback with²

$$\frac{\mathsf{INR}}{\mathsf{SNR}} \notin \left(\frac{1}{2}, 2\right)$$

can be approximated by

$$\tilde{C}_{\text{sym}} \triangleq \frac{1}{4} \log(1 + \mathsf{SNR} + \mathsf{INR}) + \frac{1}{4} \log\left(1 + \frac{\mathsf{SNR}}{1 + \mathsf{INR}}\right).$$
(5)

²A similar result can be shown when INR/SNR $\notin (1 - \delta_1, 1 + \delta_2)$ where $\delta_1, \delta_2 > 0$ are constants. In that case the gap between the achievable rate and the upper bound may depend on δ_1 and δ_2 . We refer the interested reader to the discussion at the end of Section VII on the capacity behavior at INR \approx SNR.

More precisely, the symmetric capacity is upper bounded by $C_{\text{sym}} \leq \tilde{C}_{\text{sym}} + \frac{K-1}{4} + \frac{1}{2} \log K$. Moreover, there exists a coding scheme that can support any rate satisfying $R_{\text{sym}} \leq \tilde{C}_{\text{sym}} - \frac{1}{4} \log 16K^2(K+1)$.

We will present the achievability part of Theorem 3 in Section V. The proof of the converse part can be found in Section VI.

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 $\{0, 1, \dots, p-1\}$ with sum and product operations modulo p. Moreover, define

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

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

$$Y_k = D^{q-n} X_k + \sum_{i \neq k} D^{q-m} X_i, \tag{6}$$

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 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

$$R_{\text{sym}} = \begin{cases} n - \frac{m}{2} & n > m \text{ (weak interference)}, \\ \frac{n}{K} & m = n, \\ \frac{m}{2} & n < m \text{ (strong interference)}. \end{cases}$$
(7)

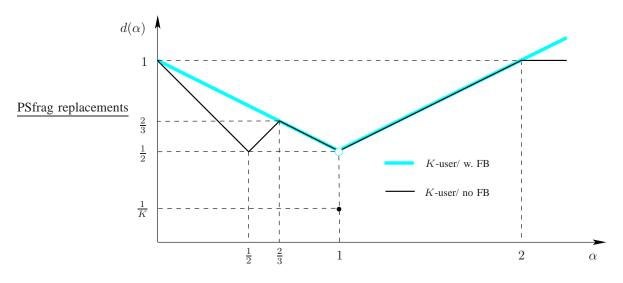


Fig. 2. The per-user generalized degrees of freedom for the K-user interference channel.

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.

We may also study a generalized version of the symmetric model introduced in (6). As we will see in the rest of this section, the symmetric topology the current model allows a simple interference alignment at the receivers. A natural generalization of this model assigns a random sign to each channels, while channel gains are kept symmetric. More precisely, in this model, called *quasi-symmetric K-user fully connected interference channel*³, the gain of all the direct links are identical and all the cross links have identical gains, but each link has a random sign which captures random phase in the Gaussian model. Note that, without loss of generality, we may assume the sign of all direct links are positive, and write the channel model as

$$Y_k = D^{q-n} X_k + \sum_i \lambda_{ki} D^{q-m} X_i, \tag{8}$$

where $\lambda_{ki} \in \{-1, +1\}$ for $i \neq k$ captures the sign of the cross link from \mathbf{Tx}_i to \mathbf{Rx}_k . The following theorem states that a similar result as Theorem 4 holds for 3-user network.

Theorem 5. The symmetric feedback capacity of the linear deterministic quasi-symmetric 3-user fully connected interference channel introduced in (8) with parameters n and m is

³We wish to thank the anonymous reviewer for suggesting this model.

given by

$$R_{\rm sym} = \begin{cases} n - \frac{m}{2} & n > m \text{ (weak interference)}, \\ \frac{n}{3} & m = n \text{ and } \Lambda + \mathbf{I} \text{ is singular}, \\ \frac{n}{2} & m = n \text{ and } \Lambda + \mathbf{I} \text{ is non-singular}, \\ \frac{m}{2} & n < m \text{ (strong interference)}, \end{cases}$$

where Λ is the channel sign matrix with $\Lambda_{ij} = \lambda_{ij}$ for $i \neq j$ and $\Lambda_{ii} = 0$ for i = 1, 2, 3.

In the following we present an example to illustrate the reason for loss in GDoF for singular $\Lambda + I$. The proof of Theorem 5 can be found in Appendix B. We will also show that this result can be generalized to arbitrary K provided that Λ satisfies certain conditions. Extension of this result to the quasi-symmetric Gaussian channel would be straight-forward from the coding scheme in Section V, and we skip it in sake of brevity.

Example 1. Consider a network with K = 3 users, and sign matrix given by

$$\Lambda = \begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \end{bmatrix}.$$

It is clear that the first and third rows of $\Lambda + \mathbf{I}$ are identical, and hence this matrix is singular. The channel model for this network for m = n can be written as

$$\begin{split} Y_1 &= X_1 - X_2 + X_3, \\ Y_2 &= X_1 + X_2 - X_3, \\ Y_3 &= X_1 - X_2 + X_3, \end{split}$$

in which $Y_1 = Y_3$. Consider an arbitrary reliable coding scheme with block length T for this network. Having the output of \mathbf{Rx}_1 over the whole block, one can find

$$Y_1^T \to W_1 \to X_1^T = g_1^T(W_1, Y_1^T).$$

Similarly

$$Y_1^T \to Y_3^T = Y_1^T \to W_3 \to X_3^T = g_3^T(W_3, Y_3^T).$$

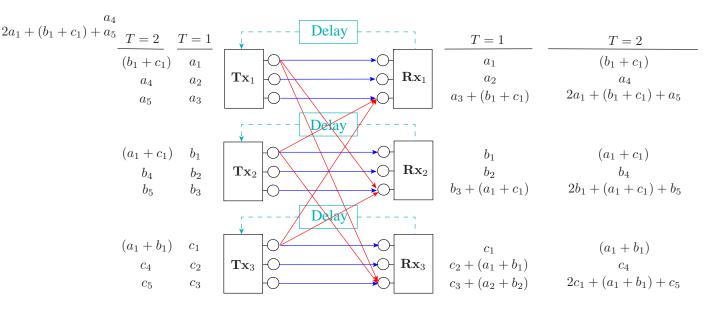


Fig. 3. Coding scheme for the linear deterministic model in the weak interference regime, for K = 3, n = 3, and m = 1.

Therefore,

and

$$X_2^T = -Y_1^T + X_1^T + X_3^T,$$
$$Y_2^T = X_1^T + X_2^T - X_3^T$$

can be found from Y_1^T , and finally W_2 can be decoded from Y_2^T . In other words, having Y_1^T , all three messages can be decoded, i.e.,

$$R_{1} + R_{2} + R_{3} \leq H(W_{1}, W_{2}, W_{3})$$

$$\leq I(W_{1} + W_{2} + W_{3}; Y_{1}^{T}) + T\epsilon_{T}$$

$$\leq H(Y_{1}^{T}) + T\epsilon_{T} \leq T(n + \epsilon_{T}),$$

which results in $R_{sym} \leq n/3$, which is the same rate claimed in Theorem 5. Achievability of this rate using time-sharing scheme is clear.

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 = 1.

For these specific parameters, we have $R_{\text{sym}} = 5/2$. As it is shown in Fig. 3, the proposed coding scheme is able to convey five 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, a_5 . Each transmitter sends three fresh symbols in its first channel use. Receivers get two interference-free symbols, and one more equation, 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 level. The two lower levels will be used to transmit the remaining fresh symbols.

Now, consider the received signals at \mathbf{Rx}_1 in two channel uses. It has received 6 linearly independent equations, involving 7 variables, which seems to be unsolvable at the first glance. However, we do not need to decode b_1 and c_1 individually. Instead, we can solve the system of linear of equations in a_1 , a_2 , a_3 , a_4 , a_5 , and $(b_1 + c_1)$, which can be solved for the intended variables. Hence, a per-user rate of 5/2 symbols/channel-use is achievable with feedback.

b) Strong Interference Regime (m > n): In this section we present an encoding scheme which can support a symmetric rate of $R_{sym} = \frac{m}{2}$. Again we focus on specific parameters, n = 1 and m = 3, which implies $R_{sym} = 3/2$.

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 $(a_1, a_2 \text{ and } a_3$ 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,

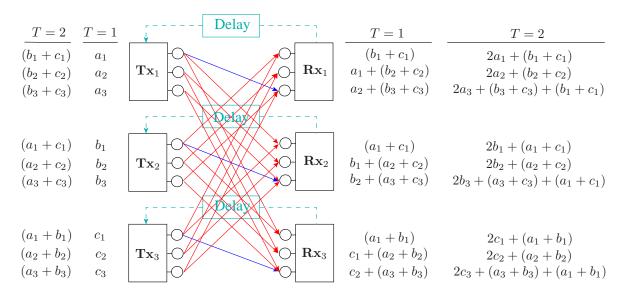


Fig. 4. Coding scheme for the linear deterministic model in the strong interference regime, for K = 3, n = 1, and m = 3.

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 $R_{sym} = n/K$ 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.

B. Outer Bound

In this section we derive an outer bound on the symmetric feedback capacity of the fully-connected interference channel. We may use shorthand notation $W_{[2:K]}$ to denote $(W_2, W_3, \ldots, W_K$. Similarly $Y_{[2:K]t}$ may be used to denote $(Y_{2t}, Y_{3t}, \ldots, Y_{Kt})$.

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:

$$H(W_{1}) + H(W_{2}) \stackrel{(a)}{=} H(W_{1}, W_{2}|W_{[3:K]})$$

$$\leq H(W_{1}, W_{2}, Y_{1}^{T}, Y_{2}^{T}|W_{[3:K]})$$

$$= H(Y_{2}^{T}|W_{[3:K]}) + H(W_{2}|W_{[3:K]}, Y_{2}^{T})$$

$$+ H(Y_{1}^{T}|W_{[3:K]}, Y_{2}^{T}) + H(W_{1}|W_{[3:K]}, Y_{1}^{T}, Y_{2}^{T})$$

$$\leq H(Y_{2}^{T}) + H(W_{2}|Y_{2}^{T})$$

$$+ H(Y_{1}^{T}|W_{[3:K]}, Y_{2}^{T}) + H(W_{1}|Y_{1}^{T})$$

$$\leq T[\max(m, n) + 2\epsilon_{T}] + H(Y_{1}^{T}|W_{[3:K]}, Y_{2}^{T}), \quad (9)$$

where (a) holds since messages are assumed to be independent, and (9) is due to Fano's inequality, in which $\epsilon_T \rightarrow 0$, as T grows. We can continue with bounding the remaining term in (9) as

$$\begin{aligned} H(Y_1^T | W_{[2:K]}, Y_2^T) \\ &\leq H(Y_1^T, Y_{[3:K]}^T, | W_{[2:K]} Y_2^T) \\ &= \sum_{t=1}^T H(Y_{1t}, Y_{[3:K]t} | W_{[2:K]}, Y_2^T, Y_1^{t-1}, Y_{[3:K]}^{t-1}) \\ &\stackrel{(b)}{=} \sum_{t=1}^T H(Y_{1t}, Y_{[3:K]t} | W_{[2:K]}, Y_2^T, Y_1^{t-1}, Y_{[3:K]}^{t-1}, X_{[2:K]t}) \\ &\stackrel{(c)}{\leq} \sum_{t=1}^T H(Y_{1t}, Y_{[3:K]t} | Y_2^T, X_{[2:K]t}) \\ &\stackrel{(d)}{=} \sum_{t=1}^T H(Y_{1t}, Y_{[3:K]t} | Y_2^T, X_{[2:K]t}, D^{q-m} X_{1t}) \\ &\stackrel{(e)}{=} \sum_{t=1}^T H(D^{q-n} X_{1t} | Y_2^T, X_{[2:K]t}, D^{q-m} X_{1t}) \\ &\stackrel{(e)}{\leq} \sum_{t=1}^T H(D^{q-n} X_{1t} | D^{q-m} X_{1t}) \\ &= T(n-m)^+, \end{aligned}$$

where (b) is due to the fact that $X_{jt} = f_{jt}(W_j, Y_j^{t-1})$; (c) holds because conditioning reduces entropy; (d) follows the fact that $D^{q-m}X_{1t} = Y_{2t} - D^{q-n}X_{2t} - D^{q-m}\sum_{j>2}X_{jt}$ is a deterministic function of $(Y_{2t}, X_{[2:K]t})$; and (e) holds because given $(D^{q-m}X_{1t}, X_{[2:K]t})$, the output $Y_{jt} = D^{q-n}X_{jt} + D^{q-m}\sum_{i\neq j}X_{it}$ is deterministically known for $j = 3, \ldots, K$; moreover, every term in Y_{1t} except $D^{q-n}X_{1t}$ is know given the same condition.

Replacing (10) in (9) we arrive at

$$R_{1} + R_{2} \leq \frac{1}{T} [H(W_{1}) + H(W_{2})]$$

$$\leq \max(m, n) + (n - m)^{+} + 2\epsilon_{T}$$

$$= \max(m, 2n - m) + 2\epsilon_{T}.$$
(11)

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

$$R_{\text{sym}} \le \max\left(\frac{m}{2}, n - \frac{m}{2}\right) + \epsilon_T.$$
 (12)

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

$$T\sum_{k=1}^{K} R_{k} = H(W_{1}, W_{2}, \dots, W_{K})$$

$$\leq I(W_{[1:K]}; Y_{[1:K]}^{T}) + KT\epsilon$$

$$\stackrel{(f)}{=} I(W_{[1:K]}; Y_{1}^{T}) + KT\epsilon$$

$$\leq H(Y_{1}^{T}) + KT\epsilon \leq Tn + KT\epsilon, \quad (13)$$

where (f) is due to the fact that $Y_1^T = Y_2^T = \cdots = Y_K^T$. Dividing (13) by KT and setting $R_1 = \cdots = R_K = R_{\text{sym}}$, we arrive at $R_{\text{sym}} \leq n/K$.

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 [14] 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 closedness 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*.

Lattice Codes: Lattice code is a class of codes that can achieve the capacity of the Gaussian channel [23], [24], with lower complexity compared to the conventional random codes. The structural behaviors of lattice codes is very important property which can also be exploited for interference alignment.

In the following we present a brief introduction for lattice codes which will be used later in our coding strategy.

A *T*-dimensional lattice Λ is subset of *T*-tuples with real elements, such that $\mathbf{x}, \mathbf{y} \in \Lambda$ implies $-\mathbf{x} \in \Lambda$ and $\mathbf{x} + \mathbf{y} \in \Lambda$. For an arbitrary $\mathbf{x} \in \mathbb{R}^T$, we define $[\mathbf{x} \mod \Lambda] = \mathbf{x} - Q(\mathbf{x})$, where

$$Q(\mathbf{x}) = \arg\min_{\mathbf{t}\in\Lambda} \| \mathbf{x} - \mathbf{t} \|$$

is the closet lattice point to x. The Voronoi cell of Λ denoted by \mathcal{V} is defined as

$$\mathcal{V} = \{ \mathbf{x} \in \mathbb{R}^T : Q(\mathbf{x}) = \mathbf{0} \}.$$

The Voronoi volume $V(\mathcal{V})$ and the second moment $\sigma^2(\Lambda)$ of the lattice are defined as

$$V(\mathcal{V}) = \int_{\mathcal{V}} d\mathbf{x}, \qquad \sigma^2(\Lambda) = \frac{\int_{\mathcal{V}} \|\mathbf{x}\|^2 d\mathbf{x}}{TV(\mathcal{V})}.$$

We further define the normalized second moment of Λ as

$$G(\Lambda) = \frac{\sigma^2(\Lambda)}{V(\mathcal{V})^{2/T}} = \frac{1}{T} \frac{\int_{\mathcal{V}} \|\mathbf{x}\|^2 \, d\mathbf{x}}{V(\mathcal{V})^{1+\frac{2}{T}}}.$$

A sequence of lattices $\{\Lambda_T\}$ is called *good quantization code* if

$$\lim_{T \to \infty} G(\Lambda_T) = \frac{1}{2\pi e}.$$

On the other hand a sequence of lattices is known to be *good* for AWGN channel coding if

$$\lim_{T \to \infty} \Pr[\mathbf{z}^T \in \mathcal{V}_T] = 1,$$

where $\mathbf{z}^T \sim \mathcal{N}(0, \sigma^2(\Lambda_T))$ is random zero-mean Gaussian noise with proper variance. It is shown in [25] that there exist sequences of lattices $\{\Lambda_T\}$ which are simultaneously good for quantization and AWGN channel coding.

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 separately presented.

A. Weak Interference Regime $2 \leq INR \leq SNR/2$

The coding scheme we use for this regime is based on the insight gained from studying the deterministic model. A careful review of the coding scheme illustrated in Appendix A-A reveals that the set of information symbols of each user can be split into three subsets: (1) $(S_k(1), \ldots, S_k(m))$ that are sent over the first channel use and cause interference for other receivers; (2) $(S_1(m+1), \ldots, S_k(n))$ which are corrupted by interference at $\mathbf{R}\mathbf{x}_k$, but do not cause interference at other receivers; and finally (3) $(S_1(n+1), \ldots, S_k(2n-m))$ which are sent on the second channel use on proper levels such that they do not cause interference at other receivers. The other m levels of each transmitter in the second channel use send the interfering signal received at $\mathbf{R}\mathbf{x}_k$ in the previous channel use. In the decoding process, each receiver first decodes the total interference from its channel output in the second channel use, and removes it to decode $(S_1(n+1), \ldots, S_k(2n-m))$. Then it also subtract the interference from its channel output in the first slot in order to decode $(S_1(1), \ldots, S_k(m))$ and $(S_1(m+1),\ldots,S_k(n)).$

Inspired by the this coding scheme and message splitting, 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} (which are counterparts of $(S_k(m+1), \ldots, S_k(n))$ and $(S_k(n+1), \ldots, S_k(2n-m))$), respectively) is performed using usual random Gaussian codebooks with block length T and average power 1, which results in codewords \mathbf{c}_{k1} and \mathbf{c}_{k2} . The power allocated to \mathbf{c}_{k1} and \mathbf{c}_{k2} is chosen such that they get received at other receivers at the noise level.

The third sub-message, w_{k0} (corresponding to $(S_k(1), \ldots, S_k(m))$) is the main interfering part from \mathbf{Tx}_k . Since we need the total interference to be decodable, we need to use a common lattice code which is shared between all transmitters.

We need a nested lattice code [18] which is generated using a good quantization lattice for shaping and a good channel coding lattice. We start with *T*-dimensional nested lattices $\Lambda_c \subseteq \Lambda_f$, where Λ_c is a good quantization lattice Λ_c with $\sigma^2(\Lambda_c) = 1$ and $G(\Lambda_c) \approx 1/2\pi e$, and Λ_f as a good channel coding lattice. We construct a codebook $\mathcal{C} = \Lambda_f \cap \mathcal{V}_c$, where \mathcal{V}_c is the Voronoi cell of the lattice Λ_c . The following properties are fairly standard in the context of lattice coding:

- a) Codebook C is a closed set with respect to summation under the "mod Λ_c" operation, i.e., if x₁, x₂ ∈ C are two codewords, then [x₁ + x₂] mod Λ_c ∈ C is also a codeword.
- b) Lattice code C can be used to reliably transmit up to rate⁴ $R = \frac{1}{2} \log(\text{SNR})$ over a Gaussian channel modelled by $Y = \sqrt{\text{SNR}X} + Z$ with $\mathbb{E}[Z^2] = 1$.

In order to encode w_{k0} , we use the common lattice code C defined above. Let s_{k0} be the lattice codeword to which w_{k0}

is mapped, and define $\mathbf{s}_0 = \mathbf{s}_{10} + \mathbf{s}_{20} + \cdots + \mathbf{s}_{K0}$.

Once the encoding process is performed, the signal transmitted by \mathbf{Tx}_k in the first block (of length T) is formed as

$$\mathbf{x}_{k1} = \sqrt{\frac{\mathsf{INR} - 1}{\mathsf{INR}}} \mathbf{c}_{k0} + \sqrt{\frac{1}{\mathsf{INR}}} \mathbf{c}_{k1}$$

where $\mathbf{c}_{k0} = [\mathbf{s}_{k0} - \mathbf{d}_k] \mod \Lambda_c$, and \mathbf{d}_k is a random dither uniformly distributed over \mathcal{V}_c , and shared between all the terminals in the network. Therefore, the signal received at \mathbf{Rx}_k can be written as

$$\begin{aligned} \mathbf{y}_{k1} &= \sqrt{\mathsf{SNR}} \mathbf{x}_{k1} + \sqrt{\mathsf{INR}} \sum_{i \neq k} \mathbf{x}_{i1} + \mathbf{z}_{k1} \\ &= \sqrt{\frac{\mathsf{SNR}}{\mathsf{INR}}} (\mathsf{INR} - 1) \mathbf{c}_{k0} + \sqrt{\frac{\mathsf{SNR}}{\mathsf{INR}}} \mathbf{c}_{k1} + \sqrt{\mathsf{INR}} - 1 \sum_{i \neq k} \mathbf{c}_{i0} \\ &+ \sum_{i \neq k} \mathbf{c}_{i1} + \mathbf{z}_{k1}. \end{aligned}$$

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

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

Recall that $\mathbf{s}_0 = [\sum \mathbf{s}_{i0} \mod \Lambda_c] = [\sum \mathbf{c}_{i0} + \sum \mathbf{d}_i \mod \Lambda_c] \in \mathcal{C}$. So it can be decoded from $\tilde{\mathbf{y}}_k$ by treating the rest as noise, provided that

$$R_0 \le \frac{1}{2} \log \left(\frac{\mathsf{INR} - 1}{K + 1} \right). \tag{14}$$

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

In the second block, having \mathbf{s}_0 decoded, $\mathbf{T}\mathbf{x}_k$ generates $\mathbf{c}_0 = [\mathbf{s}_0 - \mathbf{d}_0] \mod \Lambda_c$ and transmits

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

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

$$\mathbf{y}_{k2} = \sqrt{\mathsf{SNR}} \mathbf{x}_{k2} + \sqrt{\mathsf{INR}} \sum_{i \neq k} \mathbf{x}_{i2} + \mathbf{z}_{k2}$$
(15)
$$= \sqrt{\frac{\mathsf{SNR}}{\mathsf{INR}}} (\mathsf{INR} - 1) \mathbf{c}_0 + \sqrt{\frac{\mathsf{SNR}}{\mathsf{INR}}} \mathbf{c}_{k2} + \sqrt{\mathsf{INR}} - 1 \sum_{i \neq k} \mathbf{c}_0$$

$$+\sum_{i \neq k} \mathbf{s}_{i2} + \mathbf{z}_{k2}$$

$$= \left(\sqrt{\frac{\mathsf{SNR}}{\mathsf{INR}}} + K - 1\right) \sqrt{\mathsf{INR} - 1} \mathbf{c}_0 + \sqrt{\frac{\mathsf{SNR}}{\mathsf{INR}}} \mathbf{c}_{k2}$$

$$+ \sum_{i \neq k} \mathbf{c}_{i2} + \mathbf{z}_{k2}.$$
(16)

⁴A more sophisticated scheme can achieve rates $R = \frac{1}{2} \log (1 + \text{SNR})$. However, the simple scheme is sufficient for the purpose of approximate capacity characterization.

Receiver \mathbf{Rx}_k first decodes \mathbf{c}_0 treating everything else as Therefore, for this regime the symmetric rate of noise. This is possible as long as

$$R_0 \le \frac{1}{2} \log \left(\frac{(\mathsf{INR} - 1) \left(\sqrt{\mathsf{SNR}} + (K - 1) \sqrt{\mathsf{INR}} \right)^2}{\mathsf{SNR} + K\mathsf{INR}} \right).$$
(17)

After decoding and removing c_0 from the received signal, $\mathbf{R}\mathbf{x}_k$ can decode the Gaussian codeword c_{k2} , provided that

$$R_2 \le \frac{1}{2} \log \left(1 + \frac{\mathsf{SNR}}{K\mathsf{INR}} \right). \tag{18}$$

Next, the decoder uses \mathbf{c}_0 to remove the interference $\sum_{i \neq k} \mathbf{c}_{0i}$ from \mathbf{y}_{k1} in order to decode \mathbf{c}_{k0} and \mathbf{c}_{k1} . To this end, $\mathbf{R}\mathbf{x}_k$ first computes

$$\mathbf{y}_{k1} - \sqrt{\mathsf{INR} - 1}\mathbf{c}_0 = \left(\sqrt{\frac{\mathsf{SNR}}{\mathsf{INR}}} - 1\right)\sqrt{\mathsf{INR} - 1}\mathbf{c}_{k0} + \sqrt{\frac{\mathsf{SNR}}{\mathsf{INR}}}\mathbf{c}_{k1} + \sum_{i \neq k}\mathbf{c}_{i1} + \mathbf{z}_{k1},$$

from which codewords \mathbf{c}_{k0} and \mathbf{c}_{k1} can be sequentially decoded provided that

$$R_0 \le \frac{1}{2} \log \left(\frac{(\mathsf{INR} - 1) \left(\sqrt{\mathsf{SNR}} - \sqrt{\mathsf{INR}} \right)^2}{\mathsf{SNR} + K\mathsf{INR}} \right), \qquad (19)$$

$$R_1 \le \frac{1}{2} \log \left(1 + \frac{\mathsf{SNR}}{K\mathsf{INR}} \right). \tag{20}$$

It only remains to choose R_0 , R_1 , and R_2 that satisfy all constraints in (14)–(20). It is easy to verify that the choice of

$$R_0^{\star} = \frac{1}{2} \log \left(\frac{\mathsf{INR} - 1}{8(K+1)} \right)$$

$$R_1^{\star} = R_2^{\star} = \frac{1}{2} \log \left(1 + \frac{\mathsf{SNR}}{K\mathsf{INR}} \right)$$
(21)

satisfies all the constraints, and therefore

$$R_{\text{sym}} = \frac{1}{2} (R_0^* + R_1^* + R_2^*)$$

= $\frac{1}{4} \log \left(\frac{\text{INR} - 1}{8(K+1)} \right) + \frac{1}{2} \log \left(1 + \frac{\text{SNR}}{K\text{INR}} \right)$

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. It is easy to verify that for $2 \leq INR \leq \frac{1}{2}SNR$ we have

$$\frac{\mathsf{INR}-1}{8(K+1)}\left(1+\frac{\mathsf{SNR}}{K\mathsf{INR}}\right) \ge \frac{1}{16K(K+1)}\left(1+\mathsf{INR}+\mathsf{SNR}\right),\tag{22}$$

which implies

$$\frac{1}{4}\log\left(\frac{\mathsf{INR}-1}{8(K+1)}\right) + \frac{1}{2}\log\left(1+\frac{\mathsf{SNR}}{K\mathsf{INR}}\right)$$
$$\geq \frac{1}{4}\log\left(1+\mathsf{INR}+\mathsf{SNR}\right) + \frac{1}{4}\log\left(1+\frac{\mathsf{SNR}}{1+\mathsf{INR}}\right)$$
$$-\frac{1}{4}\log16K^2(K+1) \tag{23}$$

$$R_{\text{sym}} = \frac{1}{4} \log \left(1 + \text{INR} + \text{SNR} \right) + \frac{1}{4} \log \left(1 + \frac{\text{SNR}}{\text{INR}} \right) \\ - \frac{1}{4} \log 16K^2 (K+1)$$
(24)

is achievable.

Remark 2. It is worth mentioning that the coding schemes proposed for the weak interference regimes keep all messages except W_k almost secure from receiver $\mathbf{R}\mathbf{x}_k$, for all $k = 1, \ldots, K$. More precisely, one can show that for $K \geq 3$, the leakage rate of information is upper bounded by

$$\frac{1}{2T}I(W_k; y_j^{2T}) \le \frac{1}{2}\log\frac{K}{K-1}, \qquad k \ne j,$$
(25)

where 2T is the length of the entire course of communication. Here the upper bound on the leakage rate is a constant, independent of SNR, INR, and the actual rates of the messages. However, this secrecy is different from (and weaker than) the standard notion of secrecy, which imposes a vanishing total leakage rate in strong secrecy, or a vanishing per-symbol leakage rate in weak secrecy⁵.

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 messages 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 $\mathbf{R}\mathbf{x}_1$. Therefore, although $\mathbf{R}\mathbf{x}_1$ 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 [28] to guarantee informationsecrecy. 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. We will discuss this property of the encoding scheme in more detail in Appendix C.

B. Strong Interference Regime $INR > 2 \max(SNR, 1)$

The coding scheme for the strong interference regime is simpler than the last case. It is known that for strong interference regime in the usual interference channel (without feedback) it is optimum to decode the interference and remove it from the received signal before decoding the intended message [1]. Surprisingly, this is not the case when transmitters get feedback from their respective receivers (as far as approximate capacity is concerned). In this regime, the receivers do not need to decode the interference, and can cancel it using a zero-forcing scheme. This is implemented using Alamouti's scheme [29] in [14] for K = 2. The orthogonality of the design matrix in the 2×2 Alamouti's scheme causes the intended signal and the interference signal to be orthogonal, and so

⁵We refer the reader to [26] (and references therein) for details concerning information-theoretic secrecy. It is worth mentioning that both weak and strong secrecy are shown to be equivalent in [27], in the sense that substituting the weak secrecy criterion by the stronger version does not change the secrecy capacity.

zero-forcing the interference does not cause a loss in signal power. However, it is shown in [30] that $K \times K$ orthogonal designs exist only for K = 2, 4, 8 with real elements, and K = 2 with complex elements. When such matrices do not exist for arbitrary K, we may use non-orthogonal coding for the intended and interference signal. The key point is that if the transmitters can re-generate the interfering signal of one coding block at the receivers over another block, then the receiver can cancel two copies of interference, and decode its intended message. Since the transition matrix between the intended and interference signal on one side and the channel outputs on the other side is non-orthogonal, zeroforcing causes a power loss. However, this only affects the gap between the achievable rate and the upper bound, and does not cause a major problem when approximate capacity is concerned. We present this scheme in general detail in the rest of this section.

As in the previous case, the transmission is performed over two blocks. First note that since the interfering signals do not need to be decoded neither at the transmitters nor receivers, there is no need to use lattice codes to force the total interference to be a codeword. We associate a randomly generated Gaussian codebook of rate R_{sym} to each transmitter. Transmitter \mathbf{Tx}_k maps its message w_k to a codeword \mathbf{c}_k from its codebook, and sends $\mathbf{x}_{k1} = \mathbf{c}_k$ over the first block of transmission. At the end of the first block, each receiver sends back its received signal to its respective transmitter. Upon receiving \mathbf{y}_{k1} from the feedback link, \mathbf{Tx}_k removes its own signal, and resends the residual over the second block.

$$\mathbf{x}_{k2} = \gamma \left[\mathbf{y}_{k1} + (\sqrt{\mathsf{INR}} - \sqrt{\mathsf{SNR}}) \mathbf{x}_{k1} \right]$$
$$= \gamma (\sqrt{\mathsf{INR}} \sum_{i=1}^{K} \mathbf{c}_i + \mathbf{z}_{k1}),$$

where $\gamma = 1/\sqrt{K \text{INR} + 1}$ guarantees the transmit signal satisfies the power constraint.

At the end of the second transmission block, $\mathbf{R}\mathbf{x}_k$ has access to

$$\begin{aligned} \mathbf{y}_{k2} &= \sqrt{\mathsf{SNR}} \mathbf{x}_{k2} + \sqrt{\mathsf{INR}} \sum_{i \neq k} \mathbf{x}_{i2} + \mathbf{z}_{k2} \\ &= \gamma \left(\sqrt{\mathsf{SNR}} + (K-1)\sqrt{\mathsf{INR}} \right) \sqrt{\mathsf{INR}} \left(\mathbf{c}_k + \sum_{i \neq k} \mathbf{c}_i \right) \\ &+ \gamma \sqrt{\mathsf{SNR}} \mathbf{z}_{k1} + \gamma \sqrt{\mathsf{INR}} \sum_{i \neq k} \mathbf{z}_{i1} + \mathbf{z}_{k2}. \end{aligned}$$

Applying zero-forcing at $\mathbf{R}\mathbf{x}_k$ to remove $\sum_{i \neq k} \mathbf{c}_i$, we obtain an effective channel

$$\begin{split} \tilde{\mathbf{y}}_k &= \mathbf{y}_{k2} - \gamma \left(\sqrt{\mathsf{SNR}} + (K-1)\sqrt{\mathsf{INR}} \right) \mathbf{y}_{k1} \\ &= \gamma \left(\sqrt{\mathsf{SNR}} + (K-1)\sqrt{\mathsf{INR}} \right) (\sqrt{\mathsf{INR}} - \sqrt{\mathsf{SNR}}) \mathbf{c}_k \\ &- \gamma (K-1)\sqrt{\mathsf{INR}} \mathbf{z}_{k1} + \gamma \sqrt{\mathsf{INR}} \sum_{i \neq k} \mathbf{z}_{i1} + \mathbf{z}_{k2} \\ &= \gamma \left(\sqrt{\mathsf{SNR}} + (K-1)\sqrt{\mathsf{INR}} \right) (\sqrt{\mathsf{INR}} - \sqrt{\mathsf{SNR}}) \mathbf{c}_k + \tilde{\mathbf{z}}_{k2} \end{split}$$

which can be used for decoding c_k . The power of the total noise in this effective channel would be

$$\mathbb{E}[\tilde{\mathbf{z}}_{k2}^{2}] = \gamma^{2}(K-1)^{2} \mathsf{INR} \ \mathbb{E}[\mathbf{z}_{k1}^{2}] + \gamma^{2} \mathsf{INR} \sum_{i \neq k} \mathbb{E}[\mathbf{z}_{i1}^{2}] + \mathbb{E}[\mathbf{z}_{k2}^{2}]$$
$$= \frac{\mathsf{INR}}{K\mathsf{INR}+1} \left[(K-1)^{2} + (K-1) \right] + 1$$
$$= \frac{K^{2}\mathsf{INR}+1}{K\mathsf{INR}+1} < K.$$

On the other hand the power of the signal in the effective channel can be lower bounded by

$$\begin{split} \gamma^2 \Big(\sqrt{\mathsf{SNR}} + (K-1)\sqrt{\mathsf{INR}} \Big)^2 (\sqrt{\mathsf{INR}} - \sqrt{\mathsf{SNR}})^2 \\ &\geq \gamma^2 \left(\sqrt{\mathsf{SNR}} + \sqrt{\mathsf{INR}} \right)^2 (\sqrt{\mathsf{INR}} - \sqrt{\mathsf{SNR}})^2 \\ &= \frac{(\mathsf{INR} - \mathsf{SNR})^2}{K\mathsf{INR} + 1}. \end{split}$$

Therefore, since the course of communication is performed over two blocks, the symmetric rate

$$R_{\rm sym} = \frac{1}{4} \log \left(1 + \frac{(\mathsf{INR} - \mathsf{SNR})^2}{K(K\mathsf{INR} + 1)} \right)$$

can be simultaneously achieved for all pair of transmitter/receiver. We can simplify this expression to make it comparable to the rate claimed in Theorem 3. First note that

$$1 + \frac{(\mathsf{INR} - \mathsf{SNR})^2}{K(K\mathsf{INR} + 1)} \ge \frac{1}{8K^2} (1 + \mathsf{SNR} + \mathsf{INR})$$
(26)

for INR ≥ 2 SNR and $K \geq 2$. On the other hand, in this regime we have $1 + \frac{\text{SNR}}{1 + \text{INR}} < 2$, which implies

$$\begin{split} R_{\mathrm{sym}} \geq \frac{1}{4} \log \left(1 + \mathsf{SNR} + \mathsf{INR}\right) + \frac{1}{4} \log \left(1 + \frac{\mathsf{SNR}}{1 + \mathsf{INR}}\right) \\ &- \frac{1}{4} \log 16 K^2. \end{split}$$

C. Negligible Interference Regime INR < 2

In the discussion of Sections V-A and V-B we excluded the cases where INR is small. If this is the case, the standard *treat interference as noise* scheme is close to be optimum. Here we briefly discuss the achievable rate and its gap from the upper bound for completeness.

In this regime, each transmitter encodes its message using a Gaussian codebook, and sends it to the receiver. The course of communication is performed in a single block, and each receiver decodes its message at the end of the block by treating the interference as noise. This can support any positive rate not exceeding

$$R_{\rm sym} = \frac{1}{2} \log \left(1 + \frac{{\sf SNR}}{1 + (K-1){\sf INR}} \right)$$

This expression can be rephrased as

$$\begin{split} R_{\rm sym} &> \frac{1}{2}\log\frac{1}{K-1}\left(1+\frac{{\sf SNR}}{1+{\sf INR}}\right) \\ &= \frac{1}{4}\log\left(1+{\sf SNR}+{\sf INR}\right)-\frac{1}{4}\log(1+{\sf INR}) \\ &+\frac{1}{4}\log\left(1+\frac{{\sf SNR}}{{\sf INR}+1}\right)-\frac{1}{2}\log(K-1) \\ &\geq \tilde{C}_{\rm sym}-\frac{1}{4}\log 3(K-1)^2, \end{split}$$

which implies $\tilde{C}_{sym} - R_{sym} \le \frac{1}{4} \log 3(K-1)^2$.

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 $i = 3, 4, \dots, K$ and $t = 1, \dots, T$. Then, we can write

$$\begin{aligned} T(R_{1} + R_{2}) &\leq H(W_{1}) + H(W_{2}) \\ \stackrel{(a)}{=} H\left(W_{1}, W_{2} \middle| W_{[3:K]}\right) \\ &= H\left(W_{2} \middle| W_{3}, \dots, W_{K}\right) + H\left(W_{1} \middle| W_{[2:K]}\right) \\ &= I\left(W_{2}; y_{2}^{T} \middle| W_{[3:K]}\right) + H\left(W_{2} \middle| y_{2}^{T}, W_{[3:K]}\right) \\ &+ I\left(W_{1}; y_{1}^{T} y_{2}^{T} \middle| W_{[2:K]}\right) + H\left(W_{1} \middle| y_{1}^{T} y_{2}^{T}, W_{[2:K]}\right) \\ &\leq I\left(W_{2}; y_{2}^{T}, \tilde{z}_{[3:K]}^{T} \middle| W_{[3:K]}\right) \\ &+ I\left(W_{1}; y_{1}^{T} y_{2}^{T}, \tilde{z}_{[3:K]}^{T} \middle| W_{[2:K]}\right) + 2T\epsilon_{T} \\ &= h\left(y_{2}^{T}, \tilde{z}_{[3:K]}^{T} \middle| W_{[3:K]}\right) - h\left(y_{1}^{T} y_{2}^{T}, \tilde{z}_{[3:K]}^{T} \middle| W_{[1:K]}\right) \\ &+ 2T\epsilon_{T} \\ &= h\left(y_{2}^{T}, \tilde{z}_{[3:K]}^{T} \middle| W_{[3:K]}\right) + h\left(y_{1}^{T} \middle| y_{2}^{T}, \tilde{z}_{[3:K]}^{T} \middle| W_{[2:K]}\right) \\ &- h\left(y_{1}^{T} y_{2}^{T}, \tilde{z}_{[3:K]}^{T} \middle| W_{[1:K]}\right) + 2T\epsilon_{T}, \end{aligned}$$

where ϵ_T vanishes as T grows. Note that we used independence of the messages in (a). We can bound each term in (27)

individually. The first term can be bounded as

$$\begin{aligned} h\left(y_{2}^{T}, \tilde{z}_{[3:K]}^{T} \middle| W_{3}, \dots, W_{K}\right) &\leq h\left(y_{2}^{T}\right) + h\left(\tilde{z}_{3}^{T}\right) + \dots + h\left(\tilde{z}_{K}^{T}\right) \\ &\stackrel{(b)}{\leq} Th(y_{2}) + \frac{T(K-2)}{2} \log(4\pi e) \\ &\leq \frac{T}{2} \log\left(1 + \mathsf{SNR} + (K-1)\mathsf{INR} + 2\sqrt{\mathsf{SNR} \cdot \mathsf{INR}} \sum_{j \neq 2} \rho_{2j} \\ &\quad + 2\mathsf{INR} \sum_{\substack{i>j\\i,j \neq 2}} \rho_{ij}\right) + \frac{T(K-1)}{2} \log(4\pi e) \\ &\leq \frac{T}{2} \log\left(1 + \left(\sqrt{\mathsf{SNR}} + (K-1)\sqrt{\mathsf{INR}}\right)^{2}\right) \\ &\quad + \frac{T(K-1)}{2} \log(4\pi e), \end{aligned}$$
(28)

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

Bounding the second term is more involved. First note that

$$\begin{split} I\left(y_{1}^{T}; y_{[3:K]}^{T} \middle| y_{2}^{T}, \tilde{z}_{[3:K]}^{T}, W_{[2:K]}\right) \\ &= \sum_{t=1}^{T} I\left(y_{1}^{T}; y_{[3:K]t} \middle| y_{2}^{T}, \tilde{z}_{[3:K]}^{T}, W_{[2:K]}, y_{[3:K]}^{t-1}\right) \\ &\stackrel{(c)}{=} \sum_{t=1}^{T} I\left(y_{1}^{T}; y_{[3:K]t} \middle| y_{2}^{T}, \tilde{z}_{[3:K]}^{T}, W_{[2:K]}, y_{[3:K]}^{t-1}, x_{[2:K]t}\right) \\ &\stackrel{(d)}{=} 0 \end{split}$$
(29)

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

$$y_{jt} = \sqrt{\mathsf{SNR}} x_{jt} + \sqrt{\mathsf{INR}} \sum_{i \notin \{2, j\}} x_{it} + \sqrt{\mathsf{INR}} x_{2t} + z_{jt}$$
$$= \left[\sqrt{\mathsf{SNR}} x_{2t} + \sqrt{\mathsf{INR}} \sum_{i \notin \{2, j\}} x_{it} + \sqrt{\mathsf{INR}} x_{jt} + z_{2t} \right]$$
$$+ (\sqrt{\mathsf{SNR}} - \sqrt{\mathsf{INR}}) (x_{1t} - x_{2t}) + (z_{jt} - z_{2t})$$
$$= y_{2t} + (\sqrt{\mathsf{SNR}} - \sqrt{\mathsf{INR}}) (x_{jt} - x_{2t}) + \tilde{z}_{jt}, \qquad (30)$$

which implies that y_{jt} can be deterministically recovered from $(y_{2t}, x_{2t}, x_{jt}, \tilde{z}_{jt})$. Hence, each term in (29) is zero. From (29)

we can bound the second term in (27) as

Finally, we can bound the third term in (27) as follows:

$$h\left(y_{1}^{T}, y_{2}^{T}, \tilde{z}_{[3:K]}^{T} \middle| W_{[1:K]}\right)$$

$$= \sum_{t=1}^{T} h\left(y_{1t}, y_{2t}, \tilde{z}_{[3:K]t} \middle| y_{1}^{t-1}, y_{2}^{t-1}, \tilde{z}_{[3:K]}^{t-1}, W_{[1:K]}\right)$$

$$\geq \sum_{t=1}^{T} h\left(y_{1t}, y_{2t}, \tilde{z}_{[3:K]t} \middle| y_{1}^{t-1}, y_{2}^{t-1}, \tilde{z}_{[3:K]}^{t-1}, W_{[1:K]}, x_{[1:K]t}\right)$$

$$= \sum_{t=1}^{T} h\left(z_{1t}, z_{2t}, \tilde{z}_{[3:K]t} \middle| y_{1}^{t-1}, y_{2}^{t-1}, \tilde{z}_{[3:K]}^{t-1}, W_{[1:K]}, x_{[1:K]t}\right)$$

$$\stackrel{(e)}{=} \sum_{t=1}^{T} h\left(z_{1t}, z_{2t}, \tilde{z}_{[3:K]t}\right)$$

$$= \sum_{t=1}^{T} h\left(z_{1t}, z_{2t}, \tilde{z}_{[3:K]t}\right)$$

$$(32)$$

where (e) 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 (28), (31) and (32) in (27), and recalling the fact that we are interested in the maximum $R_1 = R_2 = R_{sym}$, we get

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

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

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

which implies

$$R_{\text{sym}} \leq \frac{1}{4} \log \left(1 + K^2 (\text{SNR} + \text{INR}) \right) + \frac{1}{4} \log \left(1 + \frac{\text{SNR}}{1 + \text{INR}} \right) \\ + \frac{K - 1}{4} \\ \leq \frac{1}{4} \log (1 + \text{SNR} + \text{INR}) + \frac{1}{4} \log \left(1 + \frac{\text{SNR}}{1 + \text{INR}} \right) \\ + \frac{K - 1}{4} + \frac{1}{2} \log K,$$
(33)

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,

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

The concept of generalized degrees of freedom for $\alpha = 1$ is more involved, and a finer look to the problem is necessary. For INR = SNR we claim that the degrees of freedom of the network is 1/K. Note that a simple time-sharing scheme, in which in each block all the transmitters except one are silent, guarantees a reliable rate of $R_{\text{sym}} = \frac{1}{2K} \log(1 + \text{SNR})$, which results in $d_{\text{FB}}(\text{INR} = \text{SNR}) \ge 1/K$.

On the other hand we may use a simple cut-set argument in order to show optimality of this DoF for INR = SNR. Recall that in the deterministic model, the received signal of all the receivers were identical 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

$$\begin{split} TKR_{\text{sym}} &= T\sum_{k=1}^{K} R_{k} = H\Big(W_{[1:K]}\Big) \\ &\leq I\Big(y_{[1:K]}^{T}; W_{[1:K]}\Big) + KT\epsilon_{T} \\ &= h\Big(y_{[1:K]}^{T}\Big) - h\Big(y_{[1:K]}^{T}\Big|W_{[1:K]}\Big) + KT\epsilon_{T} \\ &= h\Big(y_{1}^{T}, z_{2}^{T} - z_{1}^{T}, \dots, z_{K}^{T} - z_{1}^{T}\Big) \\ &- \sum_{t=1}^{T} h\Big(y_{[1:K]t}\Big|y_{[1:K]}^{t-1}, W_{[1:K]}\Big) + KT\epsilon_{T} \\ &\stackrel{(a)}{\leq} h(y_{1}^{T}) + \sum_{k=2}^{K} h\Big(z_{k}^{T} - z_{1}^{T}\Big) \\ &- \sum_{t=1}^{T} h\Big(y_{[1:K]t}\Big|y_{[1:K]}^{t-1}, W_{[1:K]}, x_{[1:K]t}\Big) + KT\epsilon_{T} \\ &= h(y_{1}^{T}) + \sum_{k=2}^{K} h\Big(z_{k}^{T} - z_{1}^{T}\Big) - \sum_{t=1}^{T} \sum_{k=1}^{K} h(z_{kt}) + KT\epsilon_{T} \\ &\leq \frac{T}{2} \log\Big(1 + (\sqrt{\mathsf{SNR}} + (K-1)\sqrt{\mathsf{INR}})^{2}\Big) \\ &+ \frac{(K-1)T}{2} \log 2 + KT\epsilon_{T} \\ &\leq \frac{T}{2} \log\Big(1 + K^{2}\mathsf{SNR}\Big) + \frac{(K-1)T}{2} + KT\epsilon_{T}, \end{split}$$

where (a) holds since $x_{kt} = f_{kt}(W_k, y_k^{t-1})$. Dividing by KT,

we get

$$R_{\rm sym} \le \frac{1}{K} \log(1 + K^2 {\rm SNR}) + \frac{K-1}{2K},$$
 (34)

1

which implies $d_{\rm FB}({\sf INR} = {\sf SNR}) \leq \frac{1}{K}$.

However, a more accurate relationship between INR and SNR has to be taken into account when SNR and INR are close to each other. The reason is that

$$\lim_{\mathsf{SNR}\to\infty}\frac{\log\mathsf{INR}}{\log\mathsf{SNR}} =$$

may include several regimes with different capacity behaviors. For example, if $INR = \beta SNR$ with constant $\beta \notin (\frac{1}{2}, 2)$, we still have $\alpha = 1$. Nevertheless, the result of Theorem 3 holds for this regime of parameters, and thus $d_{FB} = \frac{1}{2}$ can be achieved. An even more complicated scenario may happen when⁶ INR = SNR(1+o(SNR)) with $\lim_{SNR\to\infty} o(SNR) = 0$.

In other words, one has to be more careful when dealing with two simultaneous limiting behaviors, namely $\log(INR) \rightarrow \log(SNR)$ and $SNR \rightarrow \infty$, because depending on different rates of growth of SNR and convergence of INR to SNR, different numbers of degrees of freedom can be achieved. This discontinuous behavior is similar to the discontinuity of the DoF of the fully connected interference channel (without feedback) studied in [31], [32]. It is shown in [32] that the per-user DoF of the *K*-user FC-IC is strictly less than $\frac{1}{2}$ when INR = β SNR and β is a non-zero rational coefficient. However, DoF = 1/2 can be achieved for irrational β .

A slightly different (and perhaps more realistic) model to study DoF (GDoF at $\alpha = 1$) is to fix the channel gains, and allow the transmit power of all transmitters to increase simultaneously, i.e., $P \rightarrow \infty$ where $\mathbb{E}[x_k^2] \leq P$. Under this model, instead of having two independently⁷ growing variables (SNR and INR), we deal with a single variable P, and the relationship between the signal-to-noise ratio and interferenceto-noise ratio is controlled by the channel coefficients. A nice and generic result of Cadambe and Jafar [33] shows that the per-user DoF of K-user FC-IC with feedback and randomly chosen channel coefficients (not necessarily symmetric) under the latter model is 1/2, almost surely.

VIII. GAUSSIAN UPPER BOUND FOR GLOBAL FEEDBACK MODEL

In Sections V and VI we demonstrated the effect of *local* feedback on the symmetric capacity of the *K*-user fully connected interference channel. It is shown that providing each transmitter with the signal observed by its receiver in the past can be significantly beneficial. In particular, it can improve the GDoF of the network for certain regimes of interference. A natural question arises is whether availability of more information through the feedback can further improve the symmetric capacity of the network. In the rest of this section we study the effect of *global feedback* on the capacity

⁷Any rate of growth satisfying $\log INR / \log SNR \rightarrow 1$ is feasible in model in (1) and (3).

of the network, that is when each transmitter has access to the received signals of not only its respective receiver, but all other receivers. We will show that for the symmetric topology which is of interest in this paper, global feedback does *not* improve symmetric capacity beyond the local feedback (in approximate sense). This generalizes the result of [34] that in 2-user interference channel with local feedback providing additional feedback link does not improve the capacity.

In this model, the transmit signal of each user may depend on its message and all received signals in the past. Hence, in general we have

$$x_{kt} = g_{kt}(W_k, y_1^{t-1}, \dots, y_K^{t-1}).$$
(35)

Theorem 6. The symmetric capacity of the K user interference channel with global feedback with $\frac{\text{INR}}{\text{SNR}} \notin (0.5, 2)$ can be approximated by

$$\tilde{C}_{\text{sym}} = \frac{1}{4}\log(1 + \mathsf{SNR} + \mathsf{INR}) + \frac{1}{4}\log\left(1 + \frac{\mathsf{SNR}}{1 + \mathsf{INR}}\right).$$
(36)

Note that the coding scheme presented in Section V well suits this model, and can be applied to achieve similar rates. We only need to derive an upper bound on the symmetric capacity for the global feedback model.

Recall the proof of the upper bound of Theorem 3 in Section VI. It is clear that the initial bound in (27) is still valid, regardless of the feedback model. Moreover, bounding inequalities (28) and (32), used to bound the first and third terms in (27) respectively, would remain the same under the global feedback model. However, the argument we used to bound the second bound in (31) is not valid any more. The reason is that step (c) in (29) does not hold for global feedback model, because in this model the input signal x_{jt} depends not only on (W_j, y_j^{t-1}) , but also on $(y_1^{t-1}, y_2^{t-1}, \ldots, y_K^{t-1})$, and y_1^{t-1} is missing in the condition. Alternatively, we can use the following lemma.

Lemma 1. For any reliable coding scheme of block length T, the transmit signal of users k = 2, 3, ..., K at time t can be determined from

$$Q_t = \{y_1^{t-1}, y_2^T, \tilde{z}_3^T, \dots, \tilde{z}_K^T, W_2, W_3, \dots, W_K\},\$$

where $\tilde{z}_{kt} = z_{kt} - z_{2t}$ for k = 3, 4, ..., K. More precisely, for any family of coding functions $\{g_{kt}\}$ defined in (35), there exist corresponding coding functions $\{\tilde{g}_{kt}\}$ such that

$$x_{kt} = \tilde{g}_{kt}(y_1^{t-1}, y_2^T, \tilde{z}_3^T, \dots, \tilde{z}_K^T, W_2, W_3, \dots, W_K)$$

for k = 2, 3, ..., K and t = 1, ..., T.

Proof: We prove this claim by induction on t. For t = 1, the claim is obvious, since there is no feedback in the system and $x_{k1} = g_{k1}(W_k)$. Assume the claim is correct for $t = \ell - 1$. We will show that the claim valid for $t = \ell$, i.e., $x_{k\ell} = \tilde{g}_{k\ell}(Q_\ell) = \tilde{g}_{k\ell}(y_{1,\ell-1}, Q_{\ell-1})$. Similar to (30), we have

$$y_{k,\ell-1} = y_{2,\ell-1} + (\sqrt{\mathsf{SNR}} - \sqrt{\mathsf{INR}})(x_{k,\ell-1} - x_{2,\ell-1}) + \tilde{z}_{2,\ell-1}.$$

Note that $y_{2,\ell-1}$ and $\tilde{z}_{2,\ell-1}$ are given in $Q_{\ell-1}$. Moreover, $x_{k,\ell-1} = \tilde{g}_{k,\ell-1}(Q_{\ell-1})$ and $x_{2,\ell-1} = \tilde{g}_{2,\ell-1}(Q_{\ell-1})$. Hence,

⁶Note that this regime is not included in the statement of Theorem 3. In fact, the gap between the rate can be achieved by the proposed coding scheme and the upper bound is not bounded for this regime. The characterization of (approximate) capacity remains as an open question.

we can find $y_{k,\ell-1}$ from $Q_{\ell-1}$ for k = 2, ..., K. This means the output of all receivers at time $\ell - 1$ (except \mathbf{Rx}_1) is known from $Q_{\ell-1}$ (and hence from Q_{ℓ}). On the other hand the channel output for \mathbf{Rx}_1 at time $\ell - 1$ is explicitly given in Q_{ℓ} . Therefore, all the channel outputs y_{kt} for k = 1, ..., Kand $t = 1, ..., \ell - 1$ are given by Q_{ℓ} , which together with W_k can uniquely determine the transmit signals $x_{2\ell}, ..., x_{K\ell}$.

Now, from Lemma 1, we can bound the second term in (27) as follows.

$$h\left(y_{1}^{T}\middle|y_{2}^{T},\tilde{z}_{[3:K]}^{T},W_{[2:K]}\right) = \sum_{t=1}^{T} h\left(y_{1t}\middle|y_{1}^{t-1},y_{2}^{T},\tilde{z}_{[3:K]}^{T},W_{[2:K]}\right)$$
$$\stackrel{(a)}{=} \sum_{t=1}^{T} h\left(y_{1t}\middle|y_{1}^{t-1},y_{2}^{T},\tilde{z}_{[3:K]}^{T},W_{[2:K]},x_{[2:K]}^{T}\right)$$
$$\leq \frac{T}{2}\log\left(1+\frac{\mathsf{SNR}}{1+\mathsf{INR}}\right) + \frac{T}{2}\log(2\pi e) \tag{37}$$

where we used Lemma 1 in (a), and the last inequality follows the same argument used in (31).

The rest of the proof is similar that in Section VI, since all the other inequalities still hold under the global feedback. We skip the details in order to avoid repetition.

IX. 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 [14], 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 [16], 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. In the following, we use A(a : b) to denote a column vector $[A(a) \quad A(a+1) \quad \cdots \quad A(b)]'$, where $a \leq b$ are two positive integer numbers.

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, $\mathbf{S}_k \triangleq S_k(1:2n-m) = [S_k(1), \dots, S_k(2n-m)]'$, where $[\cdot]'$ denote matrix transpose. Each user sends p = n fresh symbols over its first channel use, i.e.,

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

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,

$$Y_{k1} = \left[S'_k(1:n-m), \\ S_k(n-m+1) + S_{\sim k}(1), \dots, S_k(n) + S_{\sim k}(m) \right]',$$

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:

$$X_{k2} = \left[S_{\sim k}(1), \dots, S_{\sim k}(m), S'_k(n+1:2n-m) \right]'.$$

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

$$Y_{k2} = X_{k2} + D^{n-m} \sum_{i \neq k} X_{i2}$$

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

$$= \begin{bmatrix} S_{\sim k}(1) \\ \vdots \\ S_{\sim k}(m) \\ S_{k}(n+1) \\ \vdots \\ S_{k}(2n-m) \end{bmatrix} + (K-1) \begin{bmatrix} 0 \\ \vdots \\ 0 \\ S_{k}(1) \\ \vdots \\ S_{k}(m) \end{bmatrix} + (K-2) \begin{bmatrix} 0 \\ \vdots \\ 0 \\ S_{\sim k}(1) \\ \vdots \\ S_{\sim k}(m) \\ \vdots \\ S_{\sim k}(m) \end{bmatrix}.$$

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 2nequations 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 solved⁸. Therefore,

⁸It is easy to verify that the coefficient matrix is full-rank.

 $\mathbf{R}\mathbf{x}_k$ can recover all its 2n - m symbols transmitted by $\mathbf{T}\mathbf{x}_k$, 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,

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

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

$$Y_{k1} = D^{m-n} \mathbf{S}_k + \mathbf{S}_{\sim k}$$

= $\left[S_{\sim k}(1), \dots, S_{\sim k}(m-n), S_k(1) + S_{\sim k}(m-n+1), \dots, S_k(n) + S_{\sim k}(m) \right]'.$

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,

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

Hence, we have

$$Y_{k2} = \begin{bmatrix} 0\\ \vdots\\ 0\\ S_{\sim k}(1)\\ \vdots\\ S_{\sim k}(n) \end{bmatrix} + \begin{bmatrix} \sum_{i \neq k} S_{\sim i}(1)\\ \vdots\\ \sum_{i \neq k} S_{\sim i}(m) \end{bmatrix}$$
$$= \begin{bmatrix} 0\\ \vdots\\ 0\\ S_{\sim k}(1)\\ \vdots\\ S_{\sim k}(n) \end{bmatrix} + (K-1) \begin{bmatrix} S_{k}(1)\\ \vdots\\ S_{k}(m) \end{bmatrix} + (K-2) \begin{bmatrix} S_{\sim k}(1)\\ \vdots\\ S_{\sim k}(m) \end{bmatrix}$$
$$= (K-1)\mathbf{S}_{k} + (D^{m-n} + (K-2)\mathbf{I})\mathbf{S}_{\sim k},$$

where **I** is the identity matrix of proper size $(m \times m)$ in the equation above). 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 $\mathbf{S}_{\sim k}$):

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

This system has a unique solution if and only if the coefficient matrix is full-rank. Note that $K \not\equiv 1 \pmod{q}$ is a necessary and sufficient condition for having a unique solution, which

can be easily satisfied for a proper choice⁹ of p. Fig. 4 pictorially demonstrates this coding scheme for 3-user case.

APPENDIX B QUASI-SYMMETRIC FULLY CONNECTED K-USER INTERFERENCE CHANNEL UNDER THE DETERMINISTIC MODEL

Note that the symmetry of the channels in the fullysymmetric model allows us to align the interfering signals in the second block and easily reconstruct the same interfering signal as in the first block at each receiver. This is not possible in the quasi-symmetric case, since it is not clear whether one can simultaneously align all interfering signals.

In the following we present a generic coding scheme together with a sufficient condition which guarantees feasibility of simultaneous alignment for the quasi-symmetric model. We will further show that this conditions holds for any choice of channel signs for K = 3. This scheme is based on opportunistically choosing the coefficient of the feedback signal in formation of the transmit signal in the second phase.

Lemma 2. Simultaneous alignment of interference in the quasi-symmetric fully connected K-user interference channel is feasible provided that there exist (non-zero) diagonal matrices A, B, U and V such that

$$\Lambda A + \Lambda B \Lambda = U + V \Lambda, \tag{38}$$

where $\Lambda = \{\lambda_{ij}\}$ is the network sign matrix with zero diagonal elements, and $\{\pm 1\}$ off-diagonal elements.

In the following we prove this lemma, and at the end of this section we show that (38) can be satisfied with diagonal matrices for any Λ of size 3×3 .

Proof of Lemma 2: Assume diagonal matrices A, B, U and V exist such that (38) holds. We present a coding scheme which guarantees simultaneous interference alignment at all the receivers.

Case I: Weak Interference Regime (n > m)

We borrow the notation from Appendix A, to denote the transmit signal of each user in the first block of transmission:

$$X_{k1} = S_k(1:n) = \left[S_k(1), S_k(2), \dots, S_k(n)\right]'.$$

Upon receiving feedback, transmitter \mathbf{Tx}_k can subtract its own contribution from the received signal at \mathbf{Rx}_k and compute

$$\mathcal{I}_{k} = \begin{bmatrix} I_{k}(1) \\ I_{k}(2) \\ \vdots \\ I_{k}(m) \end{bmatrix} \triangleq \begin{bmatrix} \sum_{i} \lambda_{ki} S_{i}(1) \\ \sum_{i} \lambda_{ki} S_{i}(2) \\ \vdots \\ \sum_{i} \lambda_{ki} S_{i}(m) \end{bmatrix}$$

⁹Note 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.

The transmit signal of user k in the second block, will be formed as

$$X_{k2} = \begin{bmatrix} A_k S_k(1) + B_k I_k(1) \\ \vdots \\ A_k S_k(m) + B_k I_k(m) \\ A_k S_k(m+1) + B_k S_k(n+1) \\ \vdots \\ A_k S_k(n) + B_k S_k(2n-m) \end{bmatrix},$$

where A_k and B_k are the *k*th diagonal elements of matrices A and B, respectively. Performing such coding scheme at each transmitter, the received signal at \mathbf{Rx}_k in the second block can be written as

$$Y_{k2} = X_{k2} + D^{n-m} \sum_{i} \lambda_{ki} X_{i2}$$

$$= X_{k2} + \sum_{i} \lambda_{ki} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ A_i S_i(1) + B_i \sum_j \lambda_{ij} S_j(1) \\ \vdots \\ A_i S_i(m) + B_i \sum_j \lambda_{ij} S_j(m) \end{bmatrix}$$

$$= X_{k2} + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \sum_{i} [\lambda_{ki} A_i + \sum_j \lambda_{kj} B_j \lambda_{ji}] S_i(1) \\ \vdots \\ \sum_{i} [\lambda_{ki} A_i + \sum_j \lambda_{kj} B_j \lambda_{ji}] S_i(m) \end{bmatrix}$$
(39)

Now note that $\lambda_{ki}A_i$ and $\sum_j \lambda_{kj}B_j\lambda_{ji}$ are (k, i)th elements of matrices ΛA and $\Lambda B\Lambda$, respectively, and from (38) we have

$$\lambda_{ki}A_i + \sum_j \lambda_{kj}B_j\lambda_{ji} = (\Lambda A + \Lambda B\Lambda)_{ki}$$
$$= (U + V\Lambda)_{ki} = U_{ki} + V_k\lambda_{ki}.$$

Therefore,

$$\sum_{i} \left[\lambda_{ki} A_{i} + \sum_{j} \lambda_{kj} B_{j} \lambda_{ji} \right] S_{i}(\ell) = \sum_{i} [U_{ki} + V_{k} \lambda_{ki}] S_{i}(\ell)$$
$$= U_{k} S_{k}(\ell) + V_{k} \sum_{i} \lambda_{ki} S_{i}(\ell)$$
$$= U_{k} S_{k}(\ell) + V_{k} I_{k}(\ell), \tag{40}$$

for $\ell = 1, 2, \ldots, m$. Replacing this into (39), we get

$$Y_{k2} = \begin{bmatrix} A_k S_k(1) + B_k I_k(1) \\ \vdots \\ A_k S_k(m) + B_k I_k(m) \\ A_k S_k(m+1) + B_k S_k(n+1) \\ \vdots \\ A_k S_k(n) + B_k S_k(2n-m) \end{bmatrix} + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ U_k S_k(1) + V_k I_k(1) \\ \vdots \\ U_k S_k(m) + V_k I_k(m) \end{bmatrix}$$

Therefore, all the interfering signals received at $\mathbf{R}\mathbf{x}_k$ in the second block are scaled version of those received in the first block, and so it has to decode 2n - m information symbols $S_k(1), \ldots, S_k(2n - m)$ as well as m interference symbols

 $I(1), \ldots, I(m)$ based on its 2n equations. This system is given by

$$\begin{bmatrix} Y_{k1} \\ Y_{k2} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & D^{n-m} \\ A_k \mathbf{I} + U_k D^{n-m} & B_k \mathbf{I} + V_k D^{n-m} \end{bmatrix} \times \begin{bmatrix} S_k(1:n) \\ \hline I_k(1:m) \\ \hline S_k(n+1:2n-m) \end{bmatrix}$$

One can easily verify that the determinant of the coefficient matrix of this system of equations is

$$\det(\mathbf{I}) \det \left(B_k \mathbf{I} + V_k D^{n-m} - (A_k \mathbf{I} + U_k D^{n-m}) \mathbf{I}^{-1} D^{n-m} \right)$$

=
$$\det(B_k \mathbf{I} + (V_k - A_k) D^{n-m} - U_k D^{2(n-m)})$$

=
$$B_k^K,$$

and hence it is full-rank and so decoding is feasible provided that $B_k \neq 0$ for k = 1, ..., K.

Case II: Strong Interference Regime (n < m)

Similar to the last case, each transmitter first sends m information symbol in the first block. Upon receiving the feedback, transmitter \mathbf{Tx}_k subtract its own contribution to find the interfering symbols at \mathbf{Rx}_k

$$\mathcal{I}_{k} = \begin{bmatrix} I_{k}(1) \\ I_{k}(2) \\ \vdots \\ I_{k}(m) \end{bmatrix} \triangleq \begin{bmatrix} \sum_{i} \lambda_{ki} S_{i}(1) \\ \sum_{i} \lambda_{ki} S_{i}(2) \\ \vdots \\ \sum_{i} \lambda_{ki} S_{i}(m) \end{bmatrix}$$

It forms a linear combination of its own symbols and the interfering symbols, and transmits the result over the second block:

$$X_{k2} = \begin{bmatrix} A_k S_k(1) + B_k I_k(1) \\ \vdots \\ A_k S_k(m) + B_k I_k(m) \end{bmatrix}$$

Using (40), the received signal at \mathbf{Rx}_k in the second block can be written as (41) given at the top of next page. Note that this is a function of only $S_k(1), \ldots, S_k(m)$ and $I_k(1), \ldots, I_k(m)$ variables; so the receiver can can solve the system of available 2m equations for these variables. More precisely, it has to solve

$$\begin{bmatrix} Y_{k1} \\ Y_{k2} \end{bmatrix} = \begin{bmatrix} D^{m-n} & \mathbf{I} \\ \hline U_k \mathbf{I} + A_k D^{m-n} & V_k \mathbf{I} + B_k D^{m-n} \end{bmatrix} \begin{bmatrix} S_k(1:m) \\ \hline I_k(1:m) \end{bmatrix}$$

It is easy to verify that the coefficient matrix is non-singular, since its determinant can be written as

$$\det(\mathbf{I}) \det\left((U_k \mathbf{I} + A_k D^{m-n}) - (V_k \mathbf{I} + B_k D^{m-n}) \mathbf{I}^{-1} D^{m-n} \right)$$
$$= \det\left(U_k \mathbf{I} + (A_k - V_k) D^{m-n} - B_k D^{2(m-n)} \right)$$
$$= U_k^K$$

which is non-zero, provided that $U_k \neq 0$ for k = 1, 2, ..., K.

Case III: Moderate Interference Regime (n = m)

The coding scheme and argument for this case is the same as the previously discussed cases. At the end of the second channel use, \mathbf{Rx}_k has to solve

$$Y_{k2} = D^{m-n}X_{k2} + \sum_{i} \lambda_{ki}X_{i2} = D^{m-n}X_{k2} + \sum_{i} \lambda_{ki} \begin{bmatrix} A_iS_i(1) + B_i\sum_{j}\lambda_{ij}S_j(1) \\ \vdots \\ A_iS_i(m) + B_i\sum_{j}\lambda_{ij}S_j(m) \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ \vdots \\ 0 \\ A_kS_k(1) + B_kI_k(1) \\ \vdots \\ A_kS_k(n) + B_kI_k(n) \end{bmatrix} + \begin{bmatrix} U_kS_k(1) + V_kI_k(1) \\ \vdots \\ U_kS_k(m-n) + V_kI_k(m-n) \\ U_kS_k(m-n+1) + V_kI_k(m-n+1) \\ \vdots \\ U_kS_k(m) + V_kI_k(m) \end{bmatrix}$$
(41)

$$\begin{bmatrix} Y_{k1} \\ Y_{k2} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{I} \\ \hline (U_k + A_k)\mathbf{I} & (V_k + B_k)\mathbf{I} \end{bmatrix} \begin{bmatrix} S_k(1:n) \\ \hline I_k(1:n) \end{bmatrix}$$

for $S_k(1:n)$. It is clear that this system can be uniquely solved provided that $B_k - A_k + V_k + U_k \neq 0$ for k = 1, ..., K. Now we are ready to prove Theorem 5 based on Lemma 2.

A. Proof of Theorem 5

The proof of the converse part is similar to that of Theorem 4 for $m \neq n$, and hence we skip it. For m = n we may distinguish the following two cases:

Case I: $\Lambda + I$ *is full-rank:* In this case we can use the upper bound (12) presented in Section IV-B, which yields

$$R_{\text{sym}} \le \max\left(\frac{n}{2}, n - \frac{n}{2}\right) = \frac{n}{2}$$

Case II: $\Lambda + \mathbf{I}$ is singular: It is easy to check that if rank $(\Lambda + \mathbf{I}) > 0$. If rank $(\Lambda + \mathbf{I}) = 1$, then $\lambda_{ij} = 1$ for $i \neq j$, and therefore the argument in (13) holds, which shows $R_{\text{sym}} \leq n/3$.

Now, assume $\operatorname{rank}(\Lambda + \mathbf{I}) = 2$. It is easy to verify that a 3×3 matrix with elements in $\{\pm 1\}$ has $\operatorname{rank} = 2$ if and only if it has two (up to negative sign) identical rows. Without loss of generality, we may assume the first and second rows are identical, which yields Y_2t can be deterministically recovered from Y_{1t} .

So, we can write

$$\begin{aligned} H(W_1, W_2, W_3 | Y_1^T) &= H(W_1, W_2, W_3 | Y_1^T, Y_2^T) \\ &= H(W_3 | W_1, W_2, Y_1^T, Y_2^T) + H(W_1, W_2 | Y_1^T, Y_2^T) \\ &\leq H(W_3 | W_1, W_2, Y_1^T, Y_2^T, X_1^T, X_2^T) + T\epsilon_T \\ &= H(W_3 | W_1, W_2, Y_1^T, Y_2^T, X_1^T, X_2^T, X_3^T, Y_3^T) + T\epsilon_T \\ &\leq 2T\epsilon_T, \end{aligned}$$

and therefore,

$$T(R_1 + R_2 + R_3) = H(W_1, W_2, W_3)$$

$$\leq I(W_1, W_2, W_3; Y_1^T) + H(W_1, W_2, W_3 | Y_1^T)$$

$$\leq H(Y_1^T) + 2T\epsilon_T \leq nT + 2T\epsilon_T$$

which yields in $R_{\text{sym}} \leq n/3$.

In order to prove the achievability part of Theorem 5 it remains to be shown that for any sign matrix of size 3×3 there exist a solution for (38). **Lemma 3.** For any given sign matrix Λ of size 3×3 , there exist diagonal matrices A, B, U, and V such that

$$\Lambda A + \Lambda B \Lambda = U + V \Lambda.$$

and one of the following holds:

- a) $B_k \neq 0$, $\forall k$ (for the weak interference regime); or
- b) $U_k \neq 0$, $\forall k$ (for the strong interference regime); or
- c) $B_k A_k + V_k U_k \neq 0$, $\forall k$ (for the moderate interference regime).

Proof of Lemma 3: Note that for K = 3 we are free to choose 12 variables (the diagonal elements of matrices A, B, U and V) such that they satisfy 9 linear equations. It is clear that this system of equations has multiple solutions. Hence, depending in the interference regime of interest, we can choose 3 variables in order to make the overall transition matrix of two channel uses full rank, i.e., we set $B_k \neq 0$ in the weak interference regime for k = 1, 2, 3, and $U_k \leq 0$ for the strong interference regime.

Regarding the moderate interference regime (m = n), the constraint to have a solvable system of equation is $B_k - A_k + V_k + U_k \neq 0$ for k = 1, 2, 3. It can be shown that when $\Lambda + \mathbf{I}$ is a full-rank matrix, these additional constraints are feasible with the solution for matrices A, B, U, and V, and therefore $R_{\text{sym}} = n/2$ can be achieved using cooperative interference alignment. However, if $\Lambda + \mathbf{I}$ is a singular matrix, $R_{\text{sym}} = n/3$ can be simply achieved by time-sharing scheme. This completes the proof.

APPENDIX C

WEAK SECRECY PROVIDED BY USING LATTICE CODES

In this section we prove the claim of Remark 2. We can break the output signal of \mathbf{Rx}_{i} into two blocks, and write

$$I(y_{j}^{2T}; W_{k}) = I(y_{j1}^{T}, y_{j2}^{T}; W_{k}) \leq I(y_{j1}^{T}, y_{j2}^{T}; \mathbf{s}_{k0}, \mathbf{c}_{k1}, \mathbf{c}_{k2})$$

$$\leq I(y_{j1}^{T}, y_{j2}^{T}, \mathbf{s}_{j0}, \mathbf{c}_{j1}, \mathbf{c}_{j2}; \mathbf{s}_{k0}, \mathbf{c}_{k1}, \mathbf{c}_{k2})$$

$$= I\left(\sum_{i \neq j} \mathbf{s}_{i0}, \sum_{i \neq j} \mathbf{c}_{i1} + \mathbf{z}_{j1}, \sum_{i \neq j} \mathbf{c}_{i2} + \mathbf{z}_{j2}; \mathbf{s}_{k0}, \mathbf{c}_{k1}, \mathbf{c}_{k2}\right)$$

$$= I\left(\sum_{i \neq j} \mathbf{s}_{i0}; \mathbf{s}_{k0}\right) + I\left(\sum_{i \neq j} \mathbf{c}_{i1} + \mathbf{z}_{j1}; \mathbf{c}_{k1}\right)$$

$$+ I\left(\sum_{i \neq j} \mathbf{c}_{i2} + \mathbf{z}_{j2}; \mathbf{c}_{k2}\right)$$
(42)

where (42) holds because the three pairs $(\sum_{i \neq j} \mathbf{s}_{i0}, \mathbf{s}_{k0})$, $(\sum_{i \neq j} \mathbf{c}_{i1} + \mathbf{z}_{j1}; \mathbf{c}_{k1})$ and $(\sum_{i \neq j} \mathbf{c}_{i2} + \mathbf{z}_{j2}; \mathbf{c}_{k2})$ are mutually independent. The first term in (42) is zero for $K \geq 3$ due to the crypto lemma [35]. The second and third terms can be upper bounded using the mutual information expression for Gaussian variables. Hence,

$$I(y_j^{2T}; W_k) \le 2\frac{T}{2}\log\left(1 + \frac{1}{K-1}\right) = T\log\frac{K}{K-1}$$

which is constant with respect to SNR.

ACKNOWLEDGEMENT

We are grateful to the associate editor and the reviewers for their suggestions which led to the enhancement of the scope and the readability of this work.

References

- T. S. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Information Theory*, vol. 27, no. 1, 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, no. 4, 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] C. Shannon, "The zero error capacity of a noisy channel," *IRE Trans. Information Theory*, vol. 2, no. 3, pp. 8–19, 1956.
- [9] A. El Gamal and Y. H. Kim, *Network information theory*. Cambridge University Press, 2011.
- [10] G. Kramer, "Feedback strategies for white Gaussian interference networks," *IEEE Trans. Information Theory*, vol. 48, no. 6, pp. 1423–1438, June 2002.
- [11] 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.
- [12] 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.
- [13] V. R. Cadambe and S. A. Jafar, "Feedback improves the generalized degrees of freedom of the strong interference channel," 2008, available from http://escholarship.org/uc/item/02w92010.
- [14] 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.
- [15] 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.
- [16] 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.
- [17] O. Ordentlich, U. Erez, and N. B., "The approximate sum capacity of the symmetric Gaussian K-user interference channel," Arxiv preprint arXiv:1206.0197, 2012.
- [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, Oct. 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, 2009, pp. 88–93.
- [21] S. Mohajer, R. Tandon, and H. V. Poor, "Generalized degrees of freedom of the symmetric K-user interference channel with feedback," in *IEEE International Symposium on Information Theory (ISIT)*, Cambridge, USA, July 2012.
- [22] S. A. Jafar, "Interference alignment: A new look at signal dimensions in a communication network," *Foundations and Trends in Communications* and Information Theory, vol. 7, no. 1, pp. 1–134, 2010.
- [23] R. Urbanke and B. Rimoldi, "Lattice codes can achieve capacity on the AWGN channel," *IEEE Trans. Information Theory*, vol. 44, no. 1, pp. 273–278, 1998.
- [24] U. Erez and R. Zamir, "Achieving ¹/₂ log(1 + SNR) on the AWGN channel with lattice encoding and decoding," *IEEE Trans. Information Theory*, vol. 50, no. 10, pp. 2293–2314, Jan. 2004.
- [25] U. Erez, S. Litsyn, and R. Zamir, "Lattices which are good for (almost) everything," *IEEE Trans. Information Theory*, vol. 51, no. 10, pp. 3401– 3416, Nov. 2005.
- [26] Y. Liang, H. Poor, and S. Shamai (Shitz), "Information theoretic security," *Foundations and Trends in Communications and Information Theory*, vol. 5, no. 4-5, pp. 355–580, 2008.
- [27] U. Maurer and S. Wolf, "Information-theoretic key agreement: From weak to strong secrecy for free," in Advances in CryptologyEURO-CRYPT 2000. Springer, pp. 351–368.
- [28] 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.
- [29] S. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, 1998.
- [30] V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Information Theory*, vol. 45, no. 5, pp. 1456–1467, May 1999.
- [31] A. Motahari, S. O. Gharan, and Khandani, "On the degrees of freedom of the 3-user Gaussian interference channel: The symmetric case," in *IEEE International Symposium on Information Theory (ISIT)*, 2009, pp. 1914–1918.
- [32] R. Etkin and E. Ordentlich, "The degrees-of-freedom of the K-user Gaussian interference channel is discontinuous at rational channel coefficients," *IEEE Trans. Information Theory*, vol. 55, no. 11, pp. 4932– 4946, Nov. 2009.
- [33] V. R. Cadambe and S. A. Jafar, "Degrees of freedom of wireless networks with relays, feedback, cooperation and full duplex operation," *IEEE Trans. Information Theory*, vol. 55, no. 5, pp. 2334–2344, Nov. 2009.
- [34] A. Sahai, V. Aggarwal, M. Yuksel, and A. Sabharwal, "Effective relaying in two-user interference channel with different models of channel output feedback," *Arxiv preprint arXiv:1104.4805*, 2011.
- [35] G. Forney, "On the role of MMSE estimation in approaching the information-theoretic limits of linear Gaussian channels: Shannon meets Wiener," in *Proc. the 41th Annual Allerton Conference on Communication, Control, and Computing*, 2003, pp. 430–439.

Soheil Mohajer received the B.Sc. degree in electrical engineering from the Sharif University of Technology, Tehran, Iran, in 2004, and the M.Sc. and Ph.D. degrees in communication systems both from Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland, in 2005 and 2010, respectively. He then joined Princeton University, New Jersey, as a postdoctoral research associate. Dr. Mohajer has been a post-doctoral researcher at the University of California at Berkeley, since October 2011.

His research interests include network information theory, data compression, wireless communication, and bioinformatics. **Ravi Tandon** (S03, M09) received the B.Tech degree in electrical engineering from the Indian Institute of Technology (IIT), Kanpur in 2004 and the Ph.D. degree in electrical and computer engineering from the University of Maryland, College Park in 2010. From 2010 until 2012, he was a post-doctoral research associate with Princeton University. In 2012, he joined Virginia Polytechnic Institute and State University (Virginia Tech) at Blacksburg, where he is currently a Research Assistant Professor in the Department of Electrical and Computer Engineering. His research interests are in network information theory, communication theory for wireless networks and information theoretic security.

Dr. Tandon is a recipient of the Best Paper Award at the Communication Theory symposium at the 2011 IEEE Global Telecommunications Conference.

H. Vincent Poor (S72, M77, SM82, F87) received the Ph.D. degree in electrical engineering and computer science from Princeton University in 1977. From 1977 until 1990, he was on the faculty of the University of Illinois at Urbana-Champaign. Since 1990 he has been on the faculty at Princeton, where he is the Dean of Engineering and Applied Science, and the Michael Henry Strater University Professor of Electrical Engineering. Dr. Poor's research interests are in the areas of stochastic analysis, statistical signal processing and information theory, and their applications in wireless networks and related fields including social networks and smart grid. Among his publications in these areas are *Smart Grid Communications and Network-ing* (Cambridge University Press, 2012) and *Principles of Cognitive Radio* (Cambridge University Press, 2013).

Dr. Poor is a member of the National Academy of Engineering and the National Academy of Sciences, a Fellow of the American Academy of Arts and Sciences, and an International Fellow of the Royal Academy of Engineering (U. K.). He is also a Fellow of the Institute of Mathematical Statistics, the Optical Society of America, and other organizations. In 1990, he served as President of the IEEE Information Theory Society, in 2004-07 as the Editor-in-Chief of these *Transactions*, and in 2009 as General Co-chair of the IEEE International Symposium on Information Theory, held in Seoul, South Korea. He received a Guggenheim Fellowship in 2002 and the IEEE Education Medal in 2005. Recent recognition of his work includes the 2010 IET Ambrose Fleming Medal for Achievement in Communications, the 2011 IEEE Eric E. Sumner Award, the 2011 IEEE Information Theory Paper Award, and honorary doctorates from Aalborg University, the Hong Kong University of Science and Technology, and the University of Edinburgh.