On the Symmetric Feedback Capacity of the K-user Cyclic Z-Interference Channel

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

arXiv:1109.1507v2 [cs.IT] 21 Dec 2012

Abstract—The K-user cyclic Z-interference channel models a situation in which the kth transmitter causes interference only to the $(k-1)$th receiver in a cyclic manner, e.g., the first transmitter causes interference only to the Kth receiver. The impact of noiseless feedback on the capacity of this channel is studied by focusing on the Gaussian cyclic Z-interference channel. To this end, the symmetric feedback capacity of the linear shift deterministic cyclic Z-interference channel (LD-CZIC) is completely characterized for all interference regimes. Using insights from the linear deterministic channel model, the symmetric feedback capacity of the Gaussian cyclic Z-interference channel is characterized up to within a constant number of bits. As a byproduct of the constant gap result, the symmetric generalized degrees of freedom with feedback for the Gaussian cyclic Z-interference channel are also characterized. These results highlight that the symmetric feedback capacities for both linear and Gaussian channel models are in general functions of K, the number of users. Furthermore, the capacity gain obtained due to feedback decreases as K increases.

I. INTRODUCTION

Managing the effects of interference is a key issue in currently deployed wireless networks. Among several ways to mitigate or perhaps constructively using interference is to make use of cooperation amongst interfering users. In this paper, we focus on one such important issue by studying the impact of noiseless receiver-to-transmitter feedback on the capacity of the K-user cyclic Z-interference channel (CZIC). In this model, K transmitters intend to transmit independent messages to K respective receivers and the kth transmitter causes interference to the $(k-1)$th receiver in a cyclic manner. The motivation for studying the cyclic Z-interference channel

comes from the modified Wyner model [1], which describes the soft handoff scenario of a cellular network. In the original Wyner model [2], each receiver can suffer interference from its adjacent transmitters. In the modified Wyner model, one can assume that the terminals are situated along a circular array (see Figure 1). If in addition, we assume that the mobile communicates with the intended base-station on its left (or right), while suffering interference due to the BS to its right (or left), then the resulting channel model is the K-user CZIC, which is considered in this paper. The K-user Gaussian CZIC (G-CZIC) *without feedback* was recently investigated in [3], where it was shown that the generalized degrees-of-freedom of the symmetric K-user G-CZIC are the same as for the 2-user Gaussian interference channel. By an interesting generalization of the results of Etkin, Tse and Wang [4], the approximate symmetric capacity was characterized for the weak interference regime and the exact capacity region was characterized for the strong interference regime. A simpler variation of the Gaussian K-user CZIC has been studied in [5], where the results of [3] are strengthened for the 3-user case. It is shown in [5] that a generalization of the Han-Kobayashi [6] scheme can achieve sum-capacity for some interference regimes.

In this paper we focus on the K-user CZIC *with feedback*, i.e., we assume the presence of noiseless and causal feedback from the kth receiver to the kth transmitter. For $K = 2$, this model reduces to the conventional 2-user interference channel with feedback. For $K > 2$, this model is a special case of the general K-user interference channel with feedback (see Figure 2). The 2-user interference channel with various forms of feedback has been investigated recently. Feedback coding schemes for K-user Gaussian interference networks have been developed by Kramer in [7]. Outer bounds for the 2-user interference channel with generalized feedback have been derived in [8], [9] and [10] (also see references therein). The 2-user Gaussian interference channel with noiseless (channel output) feedback was considered in [11] and the feedback capacity region was characterized to within two bits. One of the main findings in [11] is that feedback provides multiplicative gain at high signal-to-noise ratio (SNR) and the gain can be arbitrarily large for certain channel parameters. The key insights that led to this result were obtained by characterizing the feedback capacity region of the linear deterministic (LD) 2-user interference channel. The linear deterministic model despite its simplicity can provide valuable insights for the Gaussian channel model.

With this correspondence at hand, we first focus on the linear deterministic K-user CZIC with feedback. We characterize

Fig. 1. Modified Wyner Model.

the symmetric feedback capacity, $\mathcal{C}^{\mathrm{FB}}_{\mathrm{sym,LD}}$, which is defined as the maximum R such that the rate K-tuple $(R, R, \ldots, R)$ is achievable with feedback. We use insights from the linear deterministic model to characterize the symmetric feedback capacity of the K-user Gaussian CZIC within a constant number of bits (independent of the channel gains) for all interference regimes. As a consequence of our constant gap results, we also establish the generalized degrees of freedom of the Gaussian CZIC with feedback. For the scope of this paper, we restrict our attention to the case of symmetric channel parameters. For instance, for the Gaussian CZIC with feedback, we assume that the direct channel gain from the kth transmitter to the kth receiver is the same for all k, and that the interference channel gain from the kth transmitter to the $(k - 1)\mathrm{mod}\,(K)$th receiver is the same for all k.

The symmetric K-user Gaussian CZIC is described by the pair $(\mathsf{SNR}, \mathsf{INR})$, where SNR denotes the direct channel gain and INR denotes the interference channel gain. The per-user generalized degrees of freedom (**GDoF**) of the Gaussian CZIC *without feedback* is defined as

$$\mathbf{GDoF}(\alpha, K) = \frac{1}{K} \lim_{\mathsf{SNR} \to \infty} \frac{\mathcal{C}_{\mathrm{sum,G}}(K)}{\frac{1}{2} \log(1 + \mathsf{SNR})}, \quad (1)$$

where $\mathcal{C}_{\mathrm{sum,G}}(K)$ is the sum-capacity *without feedback*, and α is the interference parameter, defined as $\alpha \triangleq \frac{\log(\mathsf{INR})}{\log(\mathsf{SNR})}$. Analogously to (1), we define the per-user generalized degrees of freedom of the Gaussian CZIC *with feedback* as

$$\mathbf{GDoF}^{\mathrm{FB}}(\alpha, K) = \frac{1}{K} \lim_{\mathsf{SNR} \to \infty} \frac{\mathcal{C}^{\mathrm{FB}}_{\mathrm{sum,G}}(K)}{\frac{1}{2} \log(1 + \mathsf{SNR})}, \quad (2)$$

where $\mathcal{C}^{\mathrm{FB}}_{\mathrm{sum,G}}(K)$ is the sum-capacity *with feedback*.

In the breakthrough work [4], [12], several novel results were obtained for the 2-user Gaussian interference channel. Among the results, is the characterization of the **GDoF**. In particular, it was shown that $\mathbf{GDoF}(\alpha, 2) = \min(\max(1 - \alpha, \alpha), 1 - \alpha/2)$. On the other hand, the feedback **GDoF** for $K = 2$ was recently characterized in [11] and is given as $\mathbf{GDoF}^{\mathrm{FB}}(\alpha, 2) = \max(1 - \alpha/2, \alpha/2)$. From the results of [3], it is clear that the **GDoF** without feedback for the K-user Gaussian CZIC is the same for all $K \geq 2$, i.e., it is *independent* of K, the number of users. It is natural to ask whether this equivalence continues to hold in presence of feedback for $K > 2$.
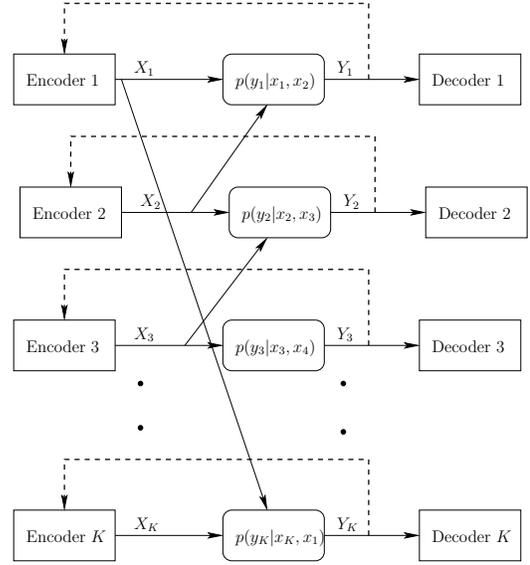
Fig. 2. K-user cyclic Z-interference channel with feedback.

We answer this question in the negative by showing that the feedback **GDoF** of the K-user Gaussian CZIC is in general a function of K. Of particular interest is the very strong interference regime, corresponding to $\alpha \geq 2$. In this regime the feedback **GDoF** for the 2-user case is given as $\mathbf{GDoF}^{\mathrm{FB}}(\alpha, 2) = \alpha/2$. This implies that the feedback gain can be unbounded as α increases. For this regime, we show that the feedback **GDoF** of the K-user Gaussian CZIC is given as $\mathbf{GDoF}^{\mathrm{FB}}(\alpha, K) = 1 + \frac{(\alpha - 2)}{K}$. This result shows that for a fixed α, as the number of users increases, the feedback gain decreases and completely vanishes in the limit $K \to \infty$. The outer bounds derived in this paper to establish capacity/constant bit gap results can be regarded as genie aided bounds derived for the 2-user case considered in [11]. However, as K, the number of users increases, selecting appropriate genies becomes prohibitively complex. In particular, for the K-user CZIC, we have a total of $K!$ sum-rate upper bounds. Depending on the interference parameter α, we carefully select the best upper bound among the $K!$ upper bounds.

This paper is organized as follows. In Section II, we describe the K-user cyclic Z-interference channel with feedback. In Section III we describe our main results for both linear deterministic and Gaussian K-user CZICs. We provide intuition as to why the feedback gain decreases as the number of users increases. Proofs for the K-user linear deterministic CZIC are presented in Sections IV and V. Constant gap results for the feedback capacity of the K-user Gaussian CZIC are established in Section VI. We conclude the paper in Section VII. Parts of this paper have appeared in [13].

II. K-USER CYCLIC Z-IC WITH FEEDBACK

The K-user cyclic Z-interference channel is described by K conditional probabilities $\{p(y_1|x_1, x_2), \ldots, p(y_K|x_K, x_1)\}$. A $(T, M_1, \ldots, M_K)$ feedback code for the CZIC consists of sequences of K encoding functions

$$f_{k,t} : \{1, \ldots, M_k\} \times \mathcal{Y}_k^{t-1} \to \mathcal{X}_{k,t}, \quad k = 1, \ldots, K, \quad (3)$$

for $t = 1, \ldots, T$, and K decoding functions

$$g_k : \mathcal{Y}_k^T \to \{1, \ldots, M_k\}, \quad k = 1, \ldots, K. \tag{4}$$

The probability of decoding error at decoder k is denoted by P_k and is defined as $P_k = \mathbb{P}(g_k(Y_k^T) \neq W_k)$, where W_k is the message of transmitter k.

A rate K-tuple $(R_1, \ldots, R_K)$ is achievable for the K-user CZIC if there exists a $(T, M_1, \ldots, M_K)$ feedback code such that $\log(M_k)/T \leq R_k - \epsilon_{k,T}$ and $P_k \leq \epsilon_{k,T}$, where $\epsilon_{k,T} \to 0$ as $T \to \infty$ for all k. The feedback capacity region $\mathcal{C}^{\mathrm{FB}}(K)$ is the set of all achievable K-tuples.

In this paper, we focus on the symmetric feedback capacity of the K-user CZIC, denoted by $\mathcal{C}_{\mathrm{sym}}^{\mathrm{FB}}(K)$, which is defined as the maximum R such that $(R, \ldots, R) \in \mathcal{C}^{\mathrm{FB}}(K)$.

A. Linear deterministic CZIC with Feedback

The symmetric linear deterministic CZIC is described by a pair of integers (n, m), where n denotes the number of signal (direct) levels and m denotes the number of interference levels observed at the receivers.

The channel input of transmitter k, denoted by X_k, for $k = 1, \ldots, K$, is assumed to be of length $\max(n, m)$.

When $n \geq m$, we denote

$$
\begin{aligned}
U_k &: \text{ top-most } (n - m) \text{ bits of } X_k \\
V_k &: \text{ top-most } m \text{ bits of } X_k \\
L_k &: \text{ lower-most } m \text{ bits of } X_k.
\end{aligned}
\tag{5}
$$

With this notation, we can write the channel outputs for the K-user LD-CZIC as follows:

$$Y_k = (U_k, L_k \oplus V_{k+1}), \tag{6}$$

for $k = 1, \ldots, K$.

When $n < m$, we denote

$$
\begin{aligned}
U_k &: \text{ top-most } (m - n) \text{ bits of } X_k \\
V_k &: \text{ top-most } n \text{ bits of } X_k \\
L_k &: \text{ lower-most } n \text{ bits of } X_k.
\end{aligned}
\tag{7}
$$

With this notation, we can write the channel outputs for the K-user LD-CZIC as follows:

$$Y_k = (U_{k+1}, L_{k+1} \oplus V_k), \tag{8}$$

for $k = 1, \ldots, K$, where we define

$$V_{K+1} \triangleq V_1, \quad U_{K+1} \triangleq U_1, \quad L_{K+1} \triangleq L_1 \tag{9}$$

for consistency.

For instance, when $n \geq m$, the 3-user LD-CZIC is described by the following input-output relationships:

$$
\begin{aligned}
Y_1 &= (U_1, L_1 \oplus V_2) \\
Y_2 &= (U_2, L_2 \oplus V_3) \\
Y_3 &= (U_3, L_3 \oplus V_1).
\end{aligned}
$$

B. Gaussian K-user CZIC with Feedback

To describe the K-user Gaussian CZIC, we denote[1] the signal transmitted by user k as X_k. We impose an average unit power constraint at each user; that is $\mathbb{E}[X_k^2] \leq 1$. The signal observed at receiver k is obtained by

$$Y_k = \sqrt{\mathsf{SNR}} X_k + \sqrt{\mathsf{INR}} X_{k+1} + Z_k, \qquad k = 1, 2, \ldots, K, \tag{10}$$

where we define $X_{K+1} \triangleq X_1$ for consistency, and the noise Z_k at receiver k is a zero mean Gaussian random variable with unit variance. Moreover, the noises across the receivers, i.e., Z_k and $Z_{k'}$ are independent for $k \neq k'$.

III. MAIN RESULTS

The results for the linear deterministic model are presented in terms of the interference parameter α, which is defined in this model as the ratio of the number of interference levels to the number of signal levels, i.e.,

$$\alpha \triangleq \frac{m}{n}. \tag{11}$$

We define the normalized[2] symmetric feedback capacity *per-user* of the LD-CZIC as follows:

$$\mathcal{C}_{\mathrm{sym,LD}}^{\mathrm{FB}}(\alpha, K) \triangleq \frac{1}{K} \frac{\mathcal{C}_{\mathrm{sum,LD}}^{\mathrm{FB}}(K)}{n}, \tag{12}$$

where $\mathcal{C}_{\mathrm{sum,LD}}^{\mathrm{FB}}(K)$ is the feedback sum-capacity of the K-user LD-CZIC.

We present our first result in the following theorem:

Theorem 1. *The normalized symmetric feedback capacity, $\mathcal{C}_{\mathrm{sym,LD}}^{\mathrm{FB}}(\alpha, K)$ of the K-user LD-CZIC is given by*

$$
\mathcal{C}_{\mathrm{sym,LD}}^{\mathrm{FB}}(\alpha, K) =
\begin{cases}
(1 - \alpha) + \frac{\alpha}{K}, & 0 \leq \alpha \leq 1/2 \\
\alpha + \frac{(2 - 3\alpha)}{K}, & 1/2 \leq \alpha \leq 2/3 \\
1 - \frac{\alpha}{2}, & 2/3 \leq \alpha \leq 1 \\
\frac{\alpha}{2}, & 1 \leq \alpha \leq 2 \\
1 + \frac{(\alpha - 2)}{K}, & \alpha \geq 2.
\end{cases}
\tag{13}
$$

Theorem 1 is proved in two parts: feedback coding schemes are presented in Section IV and corresponding upper bounds for the normalized symmetric feedback capacity are obtained in Section V.

The constant bit gap results for the Gaussian model are presented in terms of two parameters $(C_{\mathsf{SNR}}, C_{\mathsf{INR}})$, defined as

$$C_{\mathsf{SNR}} \triangleq \frac{1}{2} \log(1 + \mathsf{SNR}) \tag{14}$$

$$C_{\mathsf{INR}} \triangleq \frac{1}{2} \log(1 + \mathsf{INR}). \tag{15}$$

We also define the interference parameter for the Gaussian model as

$$\alpha = \frac{\log(\mathsf{INR})}{\log(\mathsf{SNR})}. \tag{16}$$

---

[1] With slight abuse of notation, we use similar notation for both LD-CZIC and G-CZIC channel models. However, the corresponding notation should be clear from the context.

[2] The normalization is with respect to the number of direct levels, n.

We further define the symmetric feedback capacity *per-user* of the K-user Gaussian CZIC as follows:

$$\mathcal{C}_{\mathrm{sym,G}}^{\mathrm{FB}}(K) \triangleq \frac{\mathcal{C}_{\mathrm{sum,G}}^{\mathrm{FB}}(K)}{K}, \tag{17}$$

where $\mathcal{C}_{\mathrm{sum,G}}^{\mathrm{FB}}(K)$ is the feedback sum-capacity of the K-user Gaussian CZIC. We next define the feedback degrees of freedom *per-user* for the K-user Gaussian CZIC as follows:

$$\mathbf{GDoF}^{\mathrm{FB}}(\alpha, K) = \lim_{\mathsf{SNR} \to \infty} \frac{\mathcal{C}_{\mathrm{sym,G}}^{\mathrm{FB}}(K)}{\frac{1}{2} \log(1 + \mathsf{SNR})}. \tag{18}$$

We present our next result in the following theorem:

Theorem 2. *The symmetric feedback capacity per user,* $\mathcal{C}_{\mathrm{sym,G}}^{\mathrm{FB}}(K)$ *of the K-user Gaussian CZIC satisfies*

$$\mathcal{C}_{\mathrm{sym,G}}^{\mathrm{FB}}(K) \simeq \begin{cases} C_{\mathsf{SNR}} - C_{\mathsf{INR}} + \frac{C_{\mathsf{INR}}}{K}, & 0 \le \alpha \le 1/2 \\ C_{\mathsf{INR}} + \frac{(2C_{\mathsf{SNR}} - 3C_{\mathsf{INR}})}{K}, & 1/2 \le \alpha \le 2/3 \\ C_{\mathsf{SNR}} - \frac{C_{\mathsf{INR}}}{2}, & 2/3 \le \alpha \le 1 \\ \frac{C_{\mathsf{INR}}}{2}, & 1 \le \alpha \le 2 \\ C_{\mathsf{SNR}} + \frac{(C_{\mathsf{SNR}} - 2C_{\mathsf{INR}})}{K}, & \alpha \ge 2, \end{cases} \tag{19}$$

where the notation $A \simeq B$ implies that $(A - B) \le 3$, i.e., the worst case gap (for all interference regimes) between the upper and lower bounds is at most 3 bits/user.

Theorem 2 is proved in Section VI, where we use key insights from the linear deterministic model to construct feedback coding schemes and corresponding upper bounds on the feedback sum capacity. Further analysis of these bounds shows that they differ by a constant number of bits, which is independent of $(\mathsf{SNR}, \mathsf{INR})$. We note here that the worst case gap of 3 can be reduced depending on the interference regime. For instance, in our proof of Theorem 2, we show that the gap for the case when $\alpha > 2$ is at most 2 bits.

As a consequence of Theorem 2, we have the following corollary:

Corollary 1. *The feedback degrees of freedom per-user of the K-user Gaussian CZIC is given by*

$$\mathbf{GDoF}^{\mathrm{FB}}(\alpha, K) = \begin{cases} (1 - \alpha) + \frac{\alpha}{K}, & 0 \le \alpha \le 1/2; \\ \alpha + \frac{(2 - 3\alpha)}{K}, & 1/2 \le \alpha \le 2/3; \\ 1 - \frac{\alpha}{2}, & 2/3 \le \alpha \le 1; \\ \frac{\alpha}{2}, & 1 \le \alpha \le 2; \\ 1 + \frac{(\alpha - 2)}{K}, & \alpha \ge 2. \end{cases} \tag{20}$$

We recall the no-feedback degrees of freedom per-user of the K-user Gaussian CZIC [3]:

$$\mathbf{GDoF}(\alpha, K) = \begin{cases} (1 - \alpha), & 0 \le \alpha \le 1/2 \\ \alpha, & 1/2 \le \alpha \le 2/3 \\ 1 - \frac{\alpha}{2}, & 2/3 \le \alpha \le 1 \\ \frac{\alpha}{2}, & 1 \le \alpha \le 2 \\ 1, & 2 \le \alpha. \end{cases} \tag{21}$$

Note that $\mathbf{GDoF}(\alpha, K)$ is *independent* of K, i.e., $\mathbf{GDoF}(\alpha, K) = \mathbf{GDoF}(\alpha, 2)$, for all K. This implies that from the $\mathbf{GDoF}$ point of view, the behavior of the
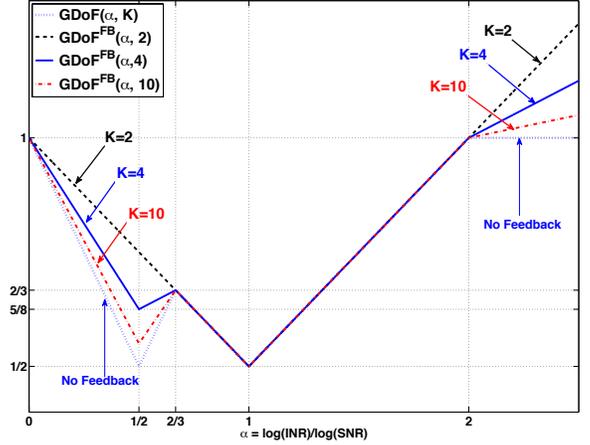
Fig. 3. **GDoF** of the K-user Gaussian CZIC with and without feedback.

K-user system is similar to the $K = 2$ user system in the *absence* of feedback.

On the other hand, we note that the feedback **GDoF** for $K = 2$ is given by [11]

$$\mathbf{GDoF}^{\mathrm{FB}}(\alpha, 2) = \begin{cases} 1 - \frac{\alpha}{2}, & \alpha \le 1 \\ \frac{\alpha}{2}, & \alpha \ge 1. \end{cases} \tag{22}$$

In the light of above observations, it is natural to ask whether the behavior of the K-user Gaussian CZIC mimics the behavior of the $K = 2$ system in the presence of feedback. Corollary 1 answers this question in the negative by showing that the feedback **GDoF** per-user for $K > 2$ is in general a function of K. Moreover, the feedback **GDoF** of $K = 2$ always serves as an upper bound for the feedback **GDoF** for $K > 2$ users.

In Figure 3, the feedback **GDoF**s are shown for the K-user Gaussian CZIC, when $K = 2, 4$ and 10.

Remark 1. *Corollary 1 also shows that* $\mathbf{GDoF}^{\mathrm{FB}}(\alpha, K)$ *can be strictly less than* $\mathbf{GDoF}^{\mathrm{FB}}(\alpha, 2)$ *(see Figure 3). Secondly, it also shows that* $\mathbf{GDoF}^{\mathrm{FB}}(\alpha, K)$ *is monotonically decreasing in K. Hence, as the number of users in the system increases, the **GDoF** gain obtained via feedback decreases. Furthermore, in the limit $K \to \infty$, the feedback gain vanishes, i.e., we have*

$$\lim_{K \to \infty} \mathbf{GDoF}^{\mathrm{FB}}(\alpha, K) = \mathbf{GDoF}(\alpha, K). \tag{23}$$

The results presented so far show that the capacity gain provided by feedback decreases as K increases. We should remark here that this behavior of capacity is not a universal phenomenon and is not necessarily dependent on the cyclic network topology. We show in the next section that this phenomenon is an artifact of the *local* feedback assumption. Under the local feedback assumption, receiver k feeds back its channel output *only* to transmitter k. To avoid any confusion, whenever we refer to feedback, it should be clear that we are referring to *local* feedback.

On the contrary, under the stronger model of global feedback, i.e., a model in which all receivers feed their channel outputs back to all the transmitters, the feedback gain is *independent* of K. We present the sum capacity of the LD-CZIC with global feedback in the following theorem.
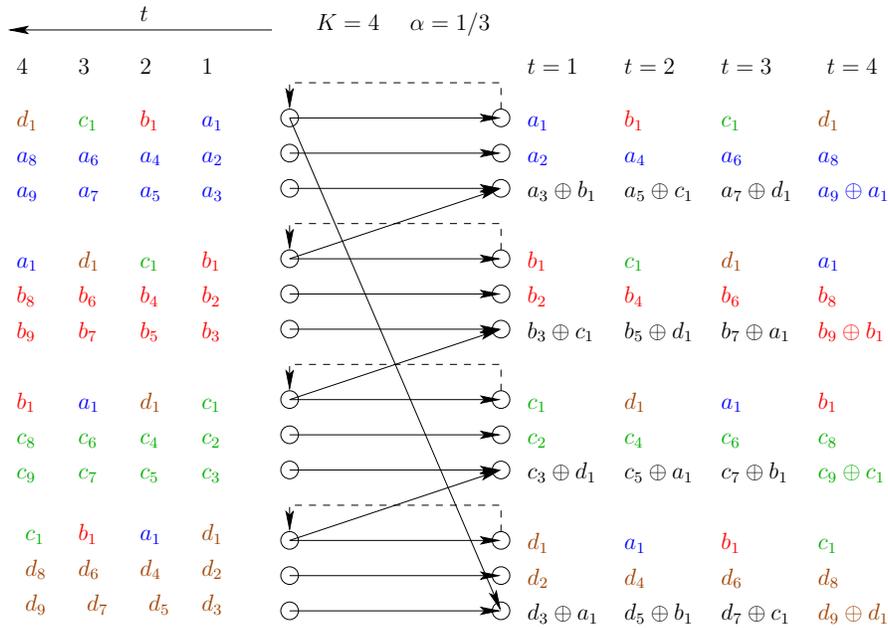
Fig. 4.   Feedback coding scheme for $\alpha = 1/3$, $K = 4$.

Theorem 3. *The normalized symmetric feedback capacity of the K-user LD-CZIC with global feedback is given by*

$$\mathcal{C}^{\text{FB}}_{\text{sym,LD}}(\alpha, K) = \max\left(1 - \frac{\alpha}{2}, \frac{\alpha}{2}\right). \qquad (24)$$

We present the coding scheme with global feedback in Section IV-E and mention the converse after Theorem 4. The setting of symmetric K-user fully connected LD-IC (and extensions to the Gaussian model) have also been considered in the literature and the interested reader is referred to [14].

IV. FEEDBACK CODING SCHEMES FOR LD-CZIC

A. Very-weak interference: $0 \leq \alpha \leq 1/2$

In this regime, we show that $K(n - m) + m$ bits per user can be reliably sent in K channel uses.

As an example, we start with the case in which $K = 4$, $m = 1$ and $n = 3$, so that $\alpha = 1/3$. To achieve 9 bits per user in 4 channel uses, the following coding scheme is used (see Figure 4):

- In the first channel use, each encoder transmits fresh bits on all levels (for example, encoder 1 sends a_1, a_2, a_3, encoder 2 sends b_1, b_2, b_3 etc.).
- Upon receiving feedback, each encoder can decode the upper most bit of the next encoder (encoder 1 decodes b_1, encoder 2 decodes c_1, etc.).
- In all subsequent channel uses, each encoder transmits the previously decoded bit on the top most level and fresh information bits in the remaining lower two levels (at $t = 2$ encoder 1 transmits b_1 on the top level and a_4, a_5 on the two lower levels).
- From Figure 4 it is clear that each user can reliably transmit 9 bits to its decoder in 4 channel uses. Hence, this scheme yields a normalized symmetric rate of $(9/4) \times (1/3) = 3/4$.

This scheme can be readily generalized for *arbitrary* numbers of users K and for any $\alpha \in [0, 1/2]$ as follows: each

transmit signal can split its signal into three mutually exclusive sets of levels as $X_k(t) = (X_{k,1}(t), X_{k,2}(t), X_{k,3}(t))$, where the number of bits in $X_{k,1}(t)$, $X_{k,2}(t)$, and $X_{k,3}(t)$ are m, $(n-2m)$, and m respectively. At $t = 1$, each encoder transmits n fresh information bits over all its transmit levels. Using feedback, at the end of slot time $t - 1$ ($1 < t \leq K$), the kth encoder decodes the upper most m bits transmitted by the $(k + 1)$th encoder, that is $X_{k+1,1}(t - 1)$. For all the subsequent channel uses the kth encoder transmits the previously decoded m bits on its top most m levels, and transmits fresh information in the lower $(n - m)$ levels. More precisely, $X_{k,1}(t) = X_{k+1,1}(t - 1)$, and $X_{k,2}(t)$ and $X_{k,3}(t)$ consist of $(n-m)$ fresh symbols intended for the kth receiver. At the end time slot t, the kth receiver can decode $X_{k,1}(t)$ and $X_{k,2}(t)$ which are received cleanly. The remaining part of the received signal would be $X_{k,3}(t) \oplus X_{k+1,1}(t)$, $X_{k,3}(t)$, and so $X_{k,3}(t)$ cannot be decoded right away because it is corrupted by interference. However, note that this interference will be retransmitted by the kth encoder in the next time slot. Hence, once $X_{k,1}(t + 1) = X_{k+1,1}(t)$ is decoded in the next block, receiver k can remove it from the interfered signal, and determine $X_{k,3}(t)$. This scheme achieves $K(n - m) + m$ bits per user in K channel uses and the achievable rate is $(n - m) + m/K$. Hence for $\alpha \in [0, 1/2]$, we have

$$\mathcal{C}^{\text{FB}}_{\text{sym,LD}}(\alpha, K) \geq (1 - \alpha) + \frac{\alpha}{K}. \qquad (25)$$

B. Weak interference: $1/2 \leq \alpha \leq 2/3$

For this regime, we present a feedback coding scheme that achieves $Km + (2n - 3m)$ bits per user in K channel uses. We break the channel input of encoder k into four mutually exclusive sets of levels as follows: $X_k(t) = (X_{k,1}(t), X_{k,2}(t), X_{k,3}(t), X_{k,4}(t))$, where the number of bits in $X_{k,r}(t)$ are $(2m - n), (2n - 3m), (2m - n)$ and $(n - m)$, for $r = 1, 2, 3$ and 4, respectively. Note that condition

$$K = 3, \quad \alpha = \tfrac{4}{7} \in \left(\tfrac{1}{2}, \tfrac{2}{3}\right)$$

$n = 7, m = 4$

Left grid (columns $t = 3, 2, 1$):

		3	2	1
$X_{11}(t)$	$(2m-n)$	a_{11}	a_7	a_1
$X_{12}(t)$	$(2n-3m)$	c_2	b_2	a_2
		c_3	b_3	a_3
$X_{13}(t)$	$(2m-n)$	—	—	—
$X_{14}(t)$	$(n-m)$	a_{12}	a_8	a_4
		a_{13}	a_9	a_5
		a_{14}	a_{10}	a_6
		b_{11}	b_7	b_1
		a_2	c_2	b_2
		a_3	c_3	b_3
		—	—	—
		b_{12}	b_8	b_4
		b_{13}	b_9	b_5
		b_{14}	b_{10}	b_6
		c_{11}	c_7	c_1
		b_2	a_2	c_2
		b_3	a_3	c_3
		—	—	—
		c_{12}	c_8	c_4
		c_{13}	c_9	c_5
		c_{14}	c_{10}	c_6

Right grid (columns $t = 1, 2, 3$):

1	2	3
a_1	a_7	a_{11}
a_2	b_2	c_2
a_3	b_3	c_3
b_1	b_7	b_{11}
$a_4 \oplus b_2$	$a_8 \oplus c_2$	$a_{12} \oplus a_2$
$a_5 \oplus b_3$	$a_9 \oplus c_3$	$a_{13} \oplus a_3$
a_6	a_{10}	a_{14}
b_1	b_7	b_{11}
b_2	c_2	a_2
b_3	c_3	a_3
c_1	c_7	c_{11}
$b_4 \oplus c_2$	$b_8 \oplus a_2$	$b_{12} \oplus b_2$
$b_5 \oplus c_3$	$b_9 \oplus a_3$	$b_{13} \oplus b_3$
b_6	b_{10}	b_{14}
c_1	c_7	c_{11}
c_2	a_2	b_2
c_3	a_3	b_3
a_1	a_7	a_{11}
$c_4 \oplus a_2$	$c_8 \oplus b_2$	$c_{12} \oplus c_2$
$c_5 \oplus a_3$	$c_9 \oplus b_3$	$c_{13} \oplus c_3$
c_6	c_{10}	c_{14}

Fig. 5. Feedback coding scheme for $\alpha = 4/7$, $K = 3$.

$1/2 \leq \alpha \leq 2/3$ ensures that the number of levels in each partition is non-negative.

At $t = 1$, each encoder transmits fresh information bits on $X_{k,1}(1)$, $X_{k,2}(1)$ and $X_{k,4}(1)$ levels. For all $t \in \{1, \ldots, K\}$, all encoders remains silent in the $X_{k,3}(t)$ level. At t, due to feedback, encoder k can decode the bits transmitted by the encoder $(k+1)$ over its top m levels, including $X_{(k+2),2}$, since $(2m-n) + (2n-3m) \leq m$. . For all $1 < t \leq K$, encoder k transmits

$$X_k(t) = (X_{k,1}(t), X_{(k+1),2}(t-1), \phi, X_{k,4}(t)),$$

where $X_{k,1}(t)$ and $X_{k,4}(t)$ consist of fresh information bits. It is clear that Km bits are achievable from the levels 1 and 4. A gain of $(2n-3m)$ bits is provided by feedback in K uses of the channel. It can be easily verified that this coding scheme yields $Km + (2n-3m)$ bits per user in K channel uses. Hence, we have

$$\mathcal{C}_{\text{sym,LD}}^{\text{FB}}(\alpha, K) \geq \alpha + \frac{(2-3\alpha)}{K}. \qquad (26)$$

This scheme is illustrated through an example in Figure 5 for the case in which $K = 3$ and $n = 7, m = 4$, so that $\alpha = 4/7$. The partition of the input at each user into four mutually exclusive levels can be readily understood through this example. It is clear that for these parameters, the scheme achieves $Km + (2n-3m) = 14$ bits per user in $K = 3$ channel uses.

C. Moderate-strong interference: $2/3 \leq \alpha \leq 2$

In this regime, Theorem 1 shows that feedback does not increase the normalized symmetric capacity and hence the no-feedback coding scheme in [3] suffices.

D. Very-strong interference: $\alpha \geq 2$

In this regime, we will show that $(K-2)n + m$ bits per user are achievable in K channel uses.

As an example, we start with the case in which $K = 4$, $m = 3$ and $n = 1$, so that $\alpha = 3$. To achieve 5 bits per user in 4 channel uses, the following coding scheme is used (see Figure 6):

- In all channel uses, each encoder remains silent in the lower-most bit. In the first channel use, each encoder transmits 2 fresh bits (for instance, encoder 1 sends a_1, a_2 and encoder 2 sends b_1, b_2). Using feedback, each encoder can decode the second bit transmitted by the encoder interfering with its decoder (encoder 1 decodes b_2, encoder 2 decodes c_2, etc.).
- In all subsequent channel uses, each encoder transmits a fresh information bit in the top-most level and the previously decoded bit in the second level (for instance, at $t = 2$, encoder 2 sends the fresh bit b_3 in the top-most level and the decoded bit c_2 in the second level).
- From Figure 6, it is clear that in 4 channel uses, using the top-most level, each decoder receives 4 bits. One more bit is received from the interfering user in the final channel use (for instance, the bit a_2 is received at decoder 1 in a delayed manner). This scheme yields a rate of $5/4$ per user.

This scheme can be readily generalized for *arbitrary* numbers of users K and for any $\alpha \geq 2$ as follows: for any $1 \leq t \leq K$, all encoders transmit no information in the lower-most n levels. At $t = 1$, the kth encoder transmits $(m-n)$ fresh information bits in the top $(m-n)$ levels. Using feedback, it decodes the $(m-n)$ bits transmitted by
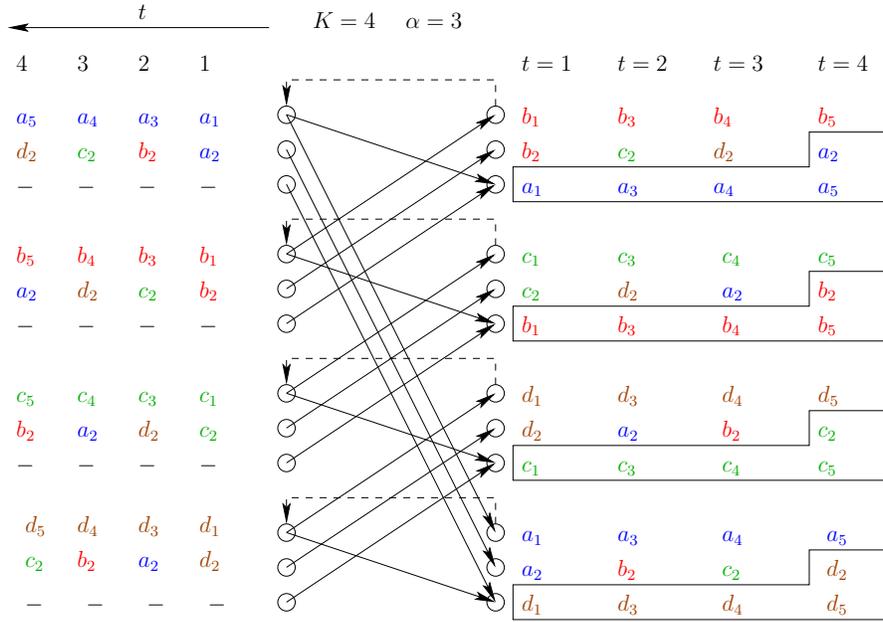
Fig. 6. Feedback coding scheme for $\alpha = 3$, $K = 4$.

the $(k + 1)$th encoder. For any $1 < t \leq K$, the kth encoder transmits fresh information on the top-most n levels and in the remaining $(m - 2n)$ levels, it transmits the lower $(m - 2n)$ bits decoded at $(t - 1)$. This scheme achieves $Kn + (m - 2n)$ bits per user in K channel uses. Hence for $\alpha \geq 2$, we have

$$\mathcal{C}^{\text{FB}}_{\text{sym,LD}}(\alpha, K) \geq 1 + \frac{(\alpha - 2)}{K}. \tag{27}$$

Theorem 1 shows that the feedback coding scheme presented above is optimal. Therefore, it is clear that the gain obtained via feedback should decrease as the number of users increases. To substantiate this claim, we note that when $\alpha \geq 2$ (corresponding to $m \geq 2n$), a normalized per-user rate of 1 can always be achieved without feedback by remaining silent on the lower most $(m - n)$ levels and sending fresh information in the top-most n levels. However, with feedback, each user can send additional information in the middle $(m - 2n)$ levels in the first channel use. This additional information can eventually reach the intended decoder via the delayed feedback path in K channel uses. For instance, in Figure 6, the bit a_2 is eventually received at decoder 1 in the last channel use. Therefore feedback can boost the reliable transmission from Kn bits to $Kn + (m - 2n)$ bits in K uses of the channel. Thus, the rate gain obtained via feedback is $(m - 2n)/K$ which decreases as K increases.

E. Coding for LD-CZIC with Global Feedback

We presented the effects of local feedback on the DoF in the previous section, which vanish as K grows for the cyclic network. However, we do not claim the same behavior for a general topology and feedback model. In this section we study the cyclic network under the deterministic model with global feedback; that is each transmitter receives the output signal of *all* the receivers with a unit delay. We will show that the V-curve for the 2-user channel is still a valid characterization for the sum capacity of this network.

We present the encoding scheme with global feedback for $K = 3$. To this end, we will show that the following rate-triple is achievable with global feedback:

$$(R_1, R_2, R_3) = (R^*, R^*, R^*), \tag{28}$$

where

$$R^* = \frac{\max(2n - m, m)}{2}. \tag{29}$$

In the following subsections, we describe the scheme for specific values of (n, m). The generalization to arbitrary K and arbitrary (n, m) is straightforward.

1) Weak Interference ($n \geq m$) : We illustrate the basic idea behind the scheme when $n = 3$ and $m = 1$ so that $\alpha = 1/3$. Figure 7 shows a scheme that achieves a rate of $5/2$ per user via global feedback. At $t = 1$, all transmitters send fresh information on all $n = 3$ levels. The least significant $m = 1$ bits suffer interference from the adjacent transmitters. Upon receiving global feedback, all $Kn = 3n = 9$ bits can be obtained at all the transmitters. This is the additional benefit of global feedback. Subsequently at $t = 2$, each transmitter sends the bits a_3, b_3 and c_3 respectively on the top $m = 1$ levels. In the remaining lower $(n - m)$ levels, transmitter j sends fresh information bits. However, in the lower-most $m = 1$ level, transmitter j cancels the interference that receiver j will suffer from transmitter $j + 1$ at $t = 2$ (since it is known via global feedback). In particular transmitter 1 sends $a_5 \oplus b_1$ in the lowest level at $t = 2$ etc. It is clear that 5 bits are reliably transmitted to each receiver in 2 channel uses, thus achieving the rate of $(2m - m)/2 = 5/2$ bits/channel-use per user.

2) Strong Interference ($n < m$) : For the strong interference case, we focus on $n = 1$ and $m = 3$ so that $\alpha = 3$. Figure 7 shows a scheme that achieves a rate of $3/2$ per user via global feedback. At $t = 1$, all transmitters send fresh information on all $m = 3$ levels. transmitters. Upon receiving global feedback, all $Km = 3m = 9$ bits can be
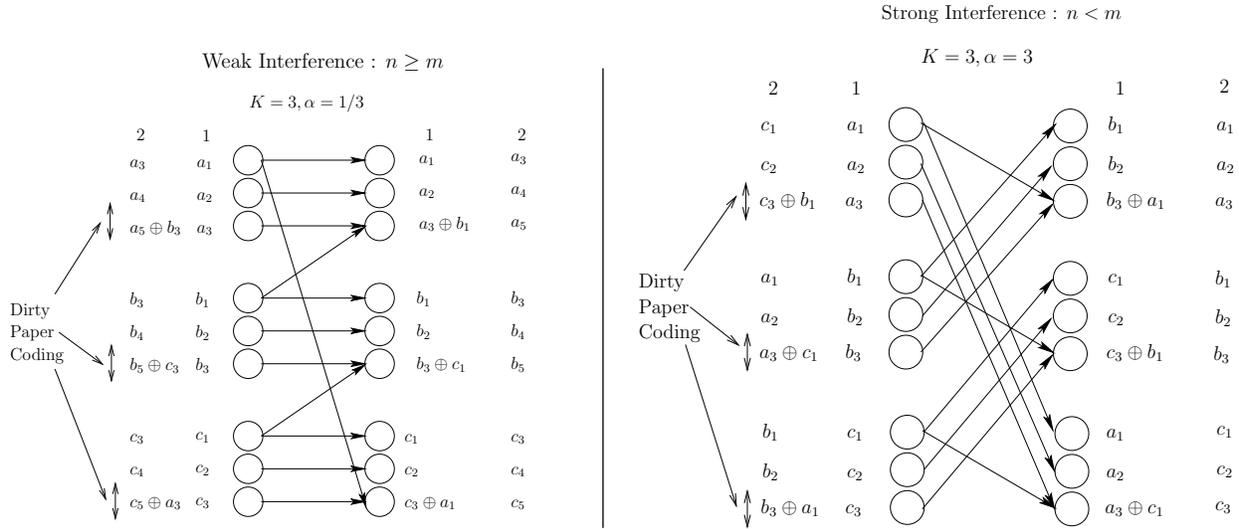
Fig. 7. Coding for LD-CZIC with global feedback.

obtained at all the transmitters. This is the additional benefit of global feedback. Subsequently at $t = 2$, transmitter j sends the bits required by receiver $(j - 1) \mod K$. In addition, in the lower-most $n = 1$ level, it cancels the interference that receiver $(j - 1) \mod K$ will face at $t = 2$ from its own transmitter. It is clear that $m = 3$ bits are reliably transmitted to each receiver in 2 channel uses, thus achieving the rate of $m/2 = 3/2$ bits/channel-use per user.

From these schemes, it becomes clear that it is the local feedback constraint that causes the gain due to feedback to decrease as K increases. As we have shown, under the global feedback assumption, the idea of canceling known interference (commonly referred to as dirty paper coding) can be employed to obtain the V-curve.

V. UPPER BOUNDS ON THE FEEDBACK SUM-CAPACITY

In this section, we present two types of upper bounds on the sum-capacity of the K-user LD-CZIC. The type-I upper bound allows us to show that the normalized symmetric feedback capacity for the K-user LD-CZIC is always upper bounded by the symmetric feedback capacity of the 2-user system. The type-II upper bound is in fact a set of $K!$ genie-aided upper bounds, in which each upper bound corresponds to a permutation of K users. These type-II upper bounds are in fact valid for the general K-user interference channel with noiseless channel output feedback, i.e., they are not specifically derived for the cyclic interference channel.

We present the type-I upper bound in the following theorem:

Theorem 4. *The normalized symmetric feedback capacity of the K-user LD-CZIC satisfies*

$$\mathcal{C}^{\text{FB}}_{\text{sym,LD}}(\alpha, K) \leq \max\left(1 - \frac{\alpha}{2}, \frac{\alpha}{2}\right). \tag{30}$$

The proof of Theorem 4 is given in the appendix. The main idea behind this upper bound is to show that

$$R_j + R_{(j+1)} \leq \max(2n - m, m), \tag{31}$$

for $j = 1, \ldots, K$. By adding all such K upper bounds and normalizing by $2nK$, we obtain the desired bound stated in
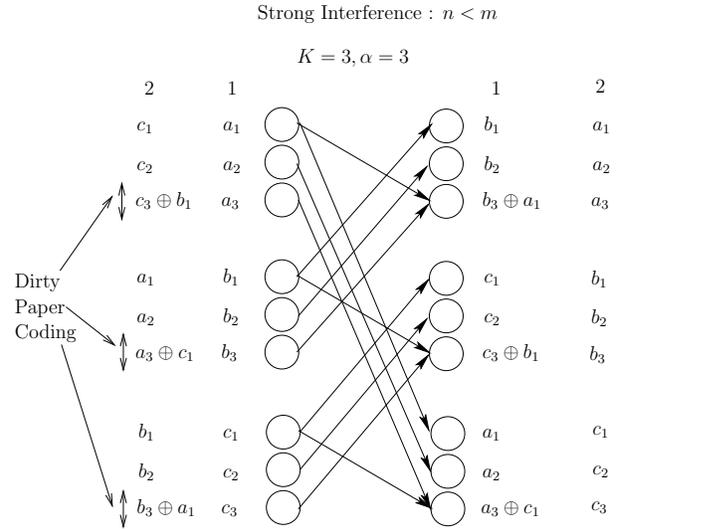
Theorem 4. Theorem 4 along with (21) leads to the conclusion that feedback does not increase the symmetric capacity of the K-user LD-CZIC in the regime $\alpha \in [2/3, 2]$. We also remark here that this upper bound also holds under the global feedback assumption. The intuition behind this can be seen as follows: the pairwise upper bounds can be regarded as genie aided bounds in which the messages of the remaining $(K - 2)$ users are supplied to both receivers and both transmitters, which is tantamount to global feedback.

We next present the type-II upper bound:

Theorem 5. *Fix a permutation order $\pi = \{\pi_1, \ldots, \pi_K\}$. Then the feedback sum-capacity of the general K-user interference channel is upper bounded as follows:*

$$\mathcal{C}^{\text{FB}}_{\text{sum}}(K)$$

$$\leq \max_{p(x_1, \ldots, x_K)} \sum_{k=1}^{K} \bigg[H(Y_{\pi_k} | X_{\pi_1}, Y_{\pi_1}, \ldots, X_{\pi_{k-1}}, Y_{\pi_{k-1}})$$

$$- H(Y_1, \ldots, Y_K | X_1, \ldots, X_K) \bigg].$$

To illustrate by an example, consider the case in which $K = 3$, for which Theorem 5 yields 6 upper bounds on the feedback sum capacity:

$$\max_{p(x_1, x_2, x_3)} \bigg[H(Y_1) + H(Y_2 | X_1, Y_1) + H(Y_3 | X_1, X_2, Y_1, Y_2)$$

$$- H(Y_1, Y_2, Y_3 | X_1, X_2, X_3) \bigg]$$

$$\max_{p(x_1, x_2, x_3)} \bigg[H(Y_1) + H(Y_3 | X_1, Y_1) + H(Y_2 | X_1, X_3, Y_1, Y_3)$$

$$- H(Y_1, Y_2, Y_3 | X_1, X_2, X_3) \bigg]$$

$$\max_{p(x_1, x_2, x_3)} \bigg[H(Y_2) + H(Y_1 | X_2, Y_2) + H(Y_3 | X_1, X_2, Y_1, Y_2)$$

$$- H(Y_1, Y_2, Y_3 | X_1, X_2, X_3) \bigg]$$

$$\max_{p(x_1, x_2, x_3)} \bigg[H(Y_2) + H(Y_3 | X_2, Y_2) + H(Y_1 | X_2, X_3, Y_2, Y_3)$$

$$- H(Y_1, Y_2, Y_3 | X_1, X_2, X_3) \bigg]$$

$$\max_{p(x_1,x_2,x_3)} \Big[H(Y_3) + H(Y_1|X_3,Y_3) + H(Y_2|X_1,X_3,Y_1,Y_3)$$
$$- H(Y_1,Y_2,Y_3|X_1,X_2,X_3) \Big]$$
$$\max_{p(x_1,x_2,x_3)} \Big[H(Y_3) + H(Y_2|X_3,Y_3) + H(Y_1|X_2,X_3,Y_2,Y_3)$$
$$- H(Y_1,Y_2,Y_3|X_1,X_2,X_3) \Big].$$

For an arbitrary K, Theorem 5 gives a total of $K!$ upper bounds. Optimization of these bounds for the general K user case and asymmetric channel gains is prohibitively complex. For the scope of this paper, we are interested in the case of CZIC with symmetric channel parameters. Depending on the range of the interference parameter α, we carefully select one of the type-II bounds and evaluate it to obtain the desired converse result as stated in Theorem 1.

A. Very Weak and Weak interference regimes: $0 \leq \alpha \leq 2/3$

In this regime, we select the type-II upper bound corresponding to the identical permutation order:

$$\pi = (1, 2, \ldots, K). \tag{32}$$

Theorem 5 yields the following bound on the sum-capacity:

$$\mathcal{C}^{\mathrm{FB}}_{\mathrm{sum,LD}}(K)$$
$$\leq \max_{p(x_1,\ldots,x_K)} \Bigg[\sum_{k=1}^{K} H(Y_k|X_1,Y_1,\ldots,X_{k-1},Y_{k-1})$$
$$- H(Y_1,\ldots,Y_K|X_1,\ldots,X_K) \Bigg] \tag{33}$$
$$= \max_{p(x_1,\ldots,x_K)} \sum_{k=1}^{K} H(Y_k|X_1,Y_1,\ldots,X_{k-1},Y_{k-1}) \tag{34}$$
$$= \max_{p(x_1,\ldots,x_K)} \Big[H(Y_1) + H(Y_2|X_1,Y_1) + \ldots$$
$$+ H(Y_K|X_1,Y_1,\ldots,X_{K-1},Y_{K-1}) \Big] \tag{35}$$
$$\leq n + \max_{p(x_1,\ldots,x_K)} \sum_{k=2}^{K-1} H(Y_k|X_{k-1},Y_{k-1})$$
$$+ \max_{p(x_1,\ldots,x_K)} H(Y_K|X_1,Y_1,X_{K-1},Y_{K-1}), \tag{36}$$

where (34) follows from the fact that the channel outputs $(Y_1,\ldots,Y_K)$ are deterministic functions of the channel inputs $(X_1,\ldots,X_K)$, and (36) follows from the fact that $H(Y_1) \leq \max(m,n) = n$.

To further upper bound (36), we first recall the notation used for $n \geq m$ in (5):

$$U_k : \text{ top-most } (n-m) \text{ bits of } X_k$$
$$V_k : \text{ top-most } m \text{ bits of } X_k$$
$$L_k : \text{ lower-most } m \text{ bits of } X_k.$$

For any $2 \leq k \leq (K-1)$, we have the following sequence of inequalities:

$$H(Y_k|X_{k-1},Y_{k-1}) = H(Y_k|X_{k-1},Y_{k-1},V_k) \tag{37}$$
$$= H(U_k, L_k \oplus V_{(k+1)}|X_{k-1},Y_{k-1},V_k) \tag{38}$$

$$\leq H(U_k|V_k) + H(L_k \oplus V_{k+1}) \tag{39}$$
$$\leq \max(0, n-2m) + m, \tag{40}$$

where (37) is due to the fact that V_k can be determined from (X_{k-1},Y_{k-1}).

Finally we upper bound the last term in (36) as follows:

$$H(Y_K|X_1,Y_1,X_{K-1},Y_{K-1})$$
$$= H(Y_K|V_1,X_1,Y_1,X_{K-1},Y_{K-1},V_K) \tag{41}$$
$$= H(U_K, L_K \oplus V_1|V_1,V_K,X_1,Y_1,X_{K-1},Y_{K-1}) \tag{42}$$
$$\leq H(U_K, L_K|V_K) \tag{43}$$
$$= H(X_K|V_K) \tag{44}$$
$$\leq (n-m). \tag{45}$$

Using (40) and (45), we can further upper bound (36) to obtain

$$\mathcal{C}^{\mathrm{FB}}_{\mathrm{sum,LD}}(K)$$
$$\leq n + (K-2)\Big[\max(0, n-2m) + m \Big] + (n-m). \tag{46}$$

Therefore, the normalized symmetric feedback capacity is upper bounded as follows:

$$\mathcal{C}^{\mathrm{FB}}_{\mathrm{sym,LD}}(\alpha, K) \leq \max(\alpha, 1-\alpha) + \frac{\min(\alpha, 2-3\alpha)}{K}, \tag{47}$$

which can also be written as

$$\mathcal{C}^{\mathrm{FB}}_{\mathrm{sym,LD}}(\alpha, K) \leq \begin{cases} (1-\alpha) + \frac{\alpha}{K}, & 0 \leq \alpha \leq 1/2 \\ \alpha + \frac{(2-3\alpha)}{K}, & 1/2 \leq \alpha \leq 2/3. \end{cases} \tag{48}$$

Note that the upper bound alone shows that in the limit $K \to \infty$ the upper bound converges to the no-feedback symmetric capacity. This implies that in the limit of large K, the feedback gain vanishes.

B. Very strong interference: $\alpha \geq 2$

In this regime, we select the type-II upper bound corresponding to the following permutation order:

$$\pi = (1, K, K-1, K-2, \ldots, 3, 2). \tag{49}$$

Theorem 5 yields the following upper bound on the sum-capacity:

$$C^{\mathrm{FB}}_{\mathrm{sum,LD}}(K)$$
$$\leq \max_{p(x_1,\ldots,x_K)} \Big[H(Y_1) + H(Y_K|X_1,Y_1) + \ldots$$
$$\ldots + H(Y_2|X_1,X_3,\ldots,X_K,Y_1,Y_3,\ldots,Y_K) \Big] \tag{50}$$
$$\leq \max_{p(x_1,\ldots,x_K)} \Big[H(Y_1) + H(Y_K|X_1,Y_1)$$
$$+ \sum_{k=3}^{K-1} H(Y_k|X_{k+1},Y_{k+1}) + H(Y_2|X_1,Y_1,X_3,Y_3) \Big]. \tag{51}$$

To further upper bound (51), we recall the notation used for $n < m$ in (7):

$$U_k : \text{ top-most } (m-n) \text{ bits of } X_k$$
$$V_k : \text{ top-most } n \text{ bits of } X_k$$
$$L_k : \text{ lower-most } n \text{ bits of } X_k.$$

We now upper bound the terms in (51) as follows. We first have the trivial upper bound $H(Y_1) \leq \max(m,n) = m$. We then bound the second term in (51) as follows:

$$H(Y_K|X_1,Y_1) = H(U_1, L_1 \oplus V_K|X_1,Y_1) \quad (52)$$
$$= H(L_1 \oplus V_K|X_1,Y_1) \quad (53)$$
$$\leq n. \quad (54)$$

Next, for any $3 \leq k \leq (K-1)$, we have

$$H(Y_k|X_{k+1},Y_{k+1}) = H(U_{k+1}, L_{k+1} \oplus V_k|X_{k+1},Y_{k+1}) \quad (55)$$
$$= H(L_{k+1} \oplus V_k|X_{k+1},Y_{k+1}) \quad (56)$$
$$\leq n, \quad (57)$$

which implies that

$$\sum_{k=3}^{K-1} H(Y_k|X_{k+1},Y_{k+1}) \leq (K-3)n. \quad (58)$$

Finally, we have

$$H(Y_2|X_1,Y_1,X_3,Y_3) = H(U_3, L_3 \oplus V_2|X_1,Y_1,X_3,Y_3) \quad (59)$$
$$= H(V_2|X_1,Y_1,X_3,Y_3) \quad (60)$$
$$= H(V_2|U_2,X_1,Y_2,X_3,Y_3) \quad (61)$$
$$= 0, \quad (62)$$

where (62) follows from the fact that $\alpha \geq 2$ corresponds to the case in which $m - n \geq n$ and therefore V_2 is completely determined by U_2.

Using (54), (58) and (62), we have the following upper bound from (51):

$$\mathcal{C}_{\text{sum,LD}}^{\text{FB}}(K)$$
$$\leq H(Y_1) + H(Y_K|X_1,Y_1) + \sum_{k=3}^{K-1} H(Y_k|X_{k+1},Y_{k+1})$$
$$+ H(Y_2|X_1,Y_1,X_3,Y_3) \quad (63)$$
$$\leq m + (K-2)n. \quad (64)$$

Normalizing this upper bound by nK, we obtain

$$\mathcal{C}_{\text{sym,LD}}^{\text{FB}}(\alpha, K) = \frac{\mathcal{C}_{\text{sum,LD}}^{\text{FB}}(K)}{nK} \quad (65)$$
$$\leq \frac{m + (K-2)n}{nK} \quad (66)$$
$$= 1 + \frac{(\alpha - 2)}{K}, \quad (67)$$

which is the desired upper bound on the normalized symmetric feedback capacity.

VI. GAUSSIAN K-USER CZIC WITH FEEDBACK

In this section, we consider the K-user Gaussian CZIC with feedback. The signal transmitted by user k is denoted by X_k. We impose an average unit power constraint at each user; that is $\mathbb{E}[X_k^2] \leq 1$. The signal observed at receiver k is obtained by

$$Y_k = \sqrt{\text{SNR}}X_k + \sqrt{\text{INR}}X_{k+1} + Z_k, \qquad k = 1,2,\ldots,K, \quad (68)$$

where we define $X_{K+1} = X_1$ for consistency.

In the following, we study five different regimes depending on the parameter α (again, defined for this model as $\alpha = \log(\text{INR})/\log(\text{SNR})$, and propose upper bounds and feedback coding schemes for each one. We analyze the performance of the proposed schemes, and derive symmetric achievable rates for them. In the rest of this section, we use bold symbols to denote blocks of length T, e.g.,

$$\mathbf{x}_k[j] = (X_k((j-1)T+1), (X_k((j-1)T+2), \ldots, (X_k(jT)).$$

The encoding schemes that we propose for each regime involve message splitting and power/rate allocation to the resulting sub-messages, motivated by the analysis of the linear deterministic model. The message splitting at the encoders is similar to that we have seen for the LD model. The powers allocated to the sub-messages at the kth transmitter are chosen so that they are received at the kth and $(k-1)$th receiver at a proper power. For sake of clarity, we explain the relationship between the coding scheme for the LD model and the Gaussian model for one of the regimes (the very weak interference regime) in full detail. However, we avoid repeating the same argument for other regimes as they follow similarly.

Moreover, for each regime of parameters the power/rate allocations proposed in the following subsections are only meaningful under a certain underlying assumption on the values of SNR, INR, and their proportion. More precisely, in our analysis we have excluded some marginal ranges of SNR and INR for which the rates are constant. It is worth mentioning that the K-user cyclic Z-Interference channel studied in this work can be approximated by simple and easily analyzable models for the excluded range of parameters. We present analysis of the excluded ranges for the very weak interference regime in Appendix VIII-D for completeness, and avoid repeating a similar argument for other regimes for the sake of brevity.

We state our upper bounds in terms of following expressions:

$$A \triangleq \frac{1}{2} \log\left(1 + \text{SNR} + \text{INR} + 2\sqrt{\text{SNR} \cdot \text{INR}}\right) \quad (69)$$
$$B \triangleq \frac{1}{2} \log\left(1 + \text{SNR} + 2\text{INR} + \text{INR}^2 + 2\sqrt{\text{SNR} \cdot \text{INR}}\right) \quad (70)$$
$$C \triangleq \frac{1}{2} \log\left(1 + \text{SNR} + \text{INR}\right) \quad (71)$$
$$D \triangleq \frac{1}{2} \log(1 + \text{SNR}) \quad (72)$$
$$E \triangleq \frac{1}{2} \log(1 + \text{INR}). \quad (73)$$

A. Very Weak Interference: $0 \leq \alpha \leq 1/2$

1) Coding Scheme: Recall the coding strategy proposed for the LD model under the very weak interference regime in Section IV-A, where a total of $K(n-m)+m$ information symbols were transmitted to each receiver over K channel uses. There each transmit signal is split into three set of mutually exclusive levels, namely $X_k(t) = (X_{k,1}(t), X_{k,2}(t), X_{k,3}(t))$, and user k sends its information symbols in $X_{k,1}(1)$ (including m bits), $\{X_{k,2}(1),\ldots,X_{k,2}(K)\}$ (including $(n-2m)K$ bits),

and $\{X_{k,3}(1), \ldots, X_{k,3}(K)\}$ (including mK bits). Moreover, $X_{k,1}(t)$ includes those levels that are heard at the $(k-1)$th receiver as interference, and the received power of $(X_{k,2}, X_{k,3}(t))$ at the $(k-1)$th receiver is at the noise level.

Motivated by the summary above, the encoding scheme we propose here also takes K blocks, each of length T. We split the message of user k into a total of $(2K+1)$ messages, namely

$$\left(M_k^{(h)}, M_k^{(1,m)}, M_k^{(2,m)} \ldots, M_k^{(K,m)}, M_k^{(1,l)}, M_k^{(2,l)}, M_k^{(K,l)}\right),$$

that it wishes to send to its respective receiver over K transmission blocks. $M_k^{(h)}$, $M_k^{(t,m)}$ and $M_k^{(t,l)}$ in this message splitting are respectively the counterparts of $X_{k,1}(1)$, $X_{k,2}(t)$ and $X_{k,3}(t)$.

Moreover, we set the size of these message sets with rates given by

$$\frac{\log |\mathcal{M}_k^{(h)}|}{n} = R_1 = \frac{1}{2} \log^+ \left(\frac{\mathsf{INR}+1}{3}\right),$$
$$k = 1, \ldots, K,$$

$$\frac{\log |\mathcal{M}_k^{(j,m)}|}{n} = R_2 = \frac{1}{2} \log^+ \left(\frac{\frac{\mathsf{SNR}}{\mathsf{INR}}+1}{2\mathsf{INR}+1}\right), \quad (74)$$
$$k = 1, \ldots, K, \ j = 1, \ldots, K,$$

$$\frac{\log |\mathcal{M}_k^{(j,l)}|}{n} = R_3 = \frac{1}{2} \log^+ \left(\frac{\mathsf{INR}+1}{2}\right),$$
$$k = 1, \ldots, K, \ j = 1, \ldots, K,$$

where $\log^+(x) = \max(\log(x), 0)$. The messages are encoded using individual Gaussian codebooks with unit average power, to obtain

$$\left(\mathbf{s}_k^{(h)}, \mathbf{s}_k^{(1,m)}, \ldots, \mathbf{s}_k^{(K,m)}, \mathbf{s}_k^{(1,l)}, \mathbf{s}_k^{(K,l)}\right).$$

The signals transmitted by user k in block k are formed as combinations of such Gaussian codewords through a proper message splitting, i.e.,

$$\mathbf{x}_k[j] = \beta_h \mathbf{x}_{k,h}[j] + \beta_m \mathbf{x}_{k,m}[j] + \beta_l \mathbf{x}_{k,l}[j],$$

where the power factors β_h, β_m and β_l can be chosen such that $\beta_h^2 + \beta_m^2 + \beta_l^2 \leq 1$. Here, we have $\mathbf{x}_{k,m}[j] = \mathbf{s}_k^{(j,m)}$ and $\mathbf{x}_{k,l} = \mathbf{s}_k^{(j,l)}$. Moreover, in the first block, we set $\mathbf{x}_{k,h}[1] = \mathbf{s}_k^{(h)}$. For subsequent blocks, the high power part of the signal consists of the high power codeword of the neighboring transmitter sent over the last block, that is $\mathbf{x}_{k,h}[j] = \mathbf{x}_{k+1,h}[j-1]$. Of course, this is only possible if transmitter k can decode $\mathbf{x}_{k+1,h}[j-1]$ from $\mathbf{y}_k[j-1]$ received over the feedback link.

It remains to set the fraction of power allocated to each part of the transmit signal. In the rest of this section we assume that $\mathsf{INR} \geq 2$, and $\mathsf{SNR} \geq 2\mathsf{INR}^2$, which are conditions that ensure that all the rates in (74) are positive. We will separately analyze the excluded range of SNR and INR in Appendix VIII-D. Having the fact that $(\mathbf{x}_{k,m}[j], \mathbf{x}_{k,l}[j])$ are the counterparts of $(X_{k,2}(j), X_{k,3}(j))$ in the LD model, we allocate them a fraction of power so that they are received at the $(k-1)$th receiver at the noise level. This allows us to safely treat them as noise when decoding at the $(k-1)$th receiver. Note from (68) that the power of $(\mathbf{x}_{k,m}[j], \mathbf{x}_{k,l}[j])$

would be magnified by INR at the $(k-1)$th receiver, and we wish the result to be at the noise power. Hence we allocate a total of $1/\mathsf{INR}$ of the available power at the kth transmitter to these codewords, and the remaining $(\mathsf{INR}-1)/\mathsf{INR}$ to the high power part, i.e.,

$$\beta_h = \sqrt{\frac{\mathsf{INR}-1}{\mathsf{INR}}}.$$

On the other hand, $\mathbf{x}_{k,l}[j]$ is the counterpart of $X_{k,3}(j)$ which is corrupted by interference when received at its intended receiver. Remember from Section IV-A that we can only decode $X_{k,3}(j)$ once the interfering signal is decoded and removed in the next channel use. The power of the interference at the kth receiver is $1/\mathsf{INR}$. Hence, we choose the power of $\mathbf{x}_{k,l}[j]$ so that after magnification by SNR over the channel to the kth receiver, it gets the same power as the interference, i.e., we choose

$$\beta_l = \sqrt{\frac{\mathsf{INR}}{\mathsf{SNR}}},$$

and the remaining power will be allocated to $\mathbf{x}_{k,m}[j]$, that is

$$\beta_m = \sqrt{\frac{\mathsf{SNR}-\mathsf{INR}^2}{\mathsf{SNR} \cdot \mathsf{INR}}}.$$

Therefore, the transmit signal from the kth transmitter over the jth block can be written as

$$\mathbf{x}_k[j] = \sqrt{\frac{\mathsf{INR}-1}{\mathsf{INR}}} \mathbf{x}_{k,h}[j] + \sqrt{\frac{\mathsf{SNR}-\mathsf{INR}^2}{\mathsf{SNR} \cdot \mathsf{INR}}} \mathbf{x}_{k,m}[j]$$
$$+ \sqrt{\frac{\mathsf{INR}}{\mathsf{SNR}}} \mathbf{x}_{k,l}[j].$$

a) Decoding the feedback signal at encoder: Upon receiving $\mathbf{y}_k[j-1]$, transmitter k removes its own signal to obtain

$$\mathbf{y}_k[j-1] - \sqrt{\mathsf{SNR}} \mathbf{x}_k[j-1]$$
$$= \sqrt{\mathsf{INR}} \mathbf{x}_{k+1}[j-1] + \mathbf{z}_k[j-1]$$
$$= \sqrt{\mathsf{INR}-1} \mathbf{x}_{k+1,h}[j-1] + \sqrt{\frac{\mathsf{SNR}-\mathsf{INR}^2}{\mathsf{SNR}}} \mathbf{x}_{k+1,m}[j-1]$$
$$+ \sqrt{\frac{\mathsf{INR}^2}{\mathsf{SNR}}} \mathbf{x}_{k+1,l}[j-1] + \mathbf{z}_k[j-1]. \quad (75)$$

It then decodes $\mathbf{x}_{k+1,h}[j-1]$, up to rate

$$R_1 \leq \frac{1}{2} \log \left(\frac{\mathsf{INR}+1}{2}\right). \quad (76)$$

b) Decoding process at the receiver: The receiver node k has to decode $\mathbf{x}_{k,h}[j]$ and $\mathbf{x}_{k,m}[j]$. This is done using a sequential decode-and-remove scheme, which can support rates that satisfy

$$R_1 \leq \frac{1}{2} \log \left(\frac{\mathsf{SNR}+\mathsf{INR}+1}{\frac{\mathsf{SNR}}{\mathsf{INR}}+\mathsf{INR}+1}\right) \quad (77)$$

$$R_2 \leq \frac{1}{2} \log \left(\frac{\frac{\mathsf{SNR}}{\mathsf{INR}}+\mathsf{INR}+1}{2\mathsf{INR}+1}\right). \quad (78)$$

The receiver also stores the remaining part of its received signal,

$$
\begin{aligned}
&\tilde{\mathbf{y}}_k[j] \\
&= \sqrt{\mathsf{INR}}\mathbf{x}_{k,l}[j] + \sqrt{\mathsf{INR}-1}\mathbf{x}_{k+1,h}[j] \\
&\quad + \sqrt{\frac{\mathsf{SNR}-\mathsf{INR}^2}{\mathsf{SNR}}}\mathbf{x}_{k+1,m}[j] + \sqrt{\frac{\mathsf{INR}^2}{\mathsf{SNR}}}\mathbf{x}_{k+1,l}[j] + \mathbf{z}_k[j],
\end{aligned}
\tag{79}
$$

for further processing. In the next block, upon decoding $\mathbf{x}_{k,h}[j+1]$, it can use it to remove a part of the interference in $\tilde{\mathbf{y}}_k[j]$. Recall that $\mathbf{x}_{k,h}[j+1] = \mathbf{x}_{k+1,h}[j]$. Therefore, by removing $\mathbf{x}_{k+1,h}[j]$ from $\tilde{\mathbf{y}}_k[j]$, it can decode $\mathbf{x}_{k,l}[j]$, provided that its rate satisfies

$$
R_3 \le \frac{1}{2}\log\left(\frac{\mathsf{INR}+2}{2}\right).
\tag{80}
$$

The total achievable rate would be

$$
\begin{aligned}
R_{\mathrm{sym}} &= \frac{R_1 + KR_2 + KR_3}{K} \tag{81} \\
&\ge \frac{1}{2}\log\left(\frac{\mathsf{SNR}}{\mathsf{INR}}+1\right) + \frac{1}{2K}\log\left(1+\mathsf{INR}\right) \\
&\quad - \frac{1}{2}\log\left(\frac{2(2\mathsf{INR}+1)}{\mathsf{INR}+1}\right) - \frac{\log 3}{2K} \tag{82} \\
&\ge (D-E) + \frac{E}{K} - 1 - \frac{\log 3}{2K}. \tag{83}
\end{aligned}
$$

2) Upper Bound: In this regime, we use the following upper bound from Theorem 5 on the feedback sum-capacity:

$$
\begin{aligned}
&\mathcal{C}^{\mathrm{FB}}_{\mathrm{sum,G}}(K) \\
&\le \max_{p(x_1,\ldots,x_K)} \Big[h(Y_1) + h(Y_2|X_1,Y_1) + \ldots \\
&\qquad\qquad + h(Y_K|X_1,Y_1,\ldots,X_{K-1},Y_{K-1}) \\
&\qquad\qquad - h(Y_1,\ldots,Y_K|X_1,\ldots,X_K)\Big]. \tag{84}
\end{aligned}
$$

We first note the following:

$$
\begin{aligned}
h(Y_1) &\le \frac{1}{2}\log\left(1+\mathsf{SNR}+\mathsf{INR}+2\sqrt{\mathsf{SNR}\cdot\mathsf{INR}}\right) + c \tag{85} \\
&= A + c, \tag{86}
\end{aligned}
$$

where $c = 1/2\log(2\pi\mathrm{e})$.

For $2 \le k \le (K-1)$, we have

$$
\begin{aligned}
&h(Y_k|X_1,Y_1,\ldots,X_{k-1},Y_{k-1}) \\
&\le h(Y_k|X_{k-1},Y_{k-1}) \tag{87} \\
&= h(\sqrt{\mathsf{SNR}}X_k + \sqrt{\mathsf{INR}}X_{k+1} + Z_k|\sqrt{\mathsf{INR}}X_k + Z_{k-1}, X_{k-1}) \tag{88} \\
&\le h(\sqrt{\mathsf{SNR}}X_k + \sqrt{\mathsf{INR}}X_{k+1} + Z_k|\sqrt{\mathsf{INR}}X_k + Z_{k-1}) \tag{89} \\
&\le \frac{1}{2}\log\left(\frac{1+\mathsf{SNR}+2\mathsf{INR}+\mathsf{INR}^2+2\sqrt{\mathsf{SNR}\cdot\mathsf{INR}}}{1+\mathsf{INR}}\right) + c \tag{90} \\
&= B - E + c. \tag{91}
\end{aligned}
$$

Similarly, we have

$$
\begin{aligned}
&h(Y_K|X_1,Y_1,\ldots,X_{K-1},Y_{K-1}) \\
&\le h(Y_K|X_1,Y_1,X_{K-1},Y_{K-1}) \tag{92} \\
&\le h(\sqrt{\mathsf{SNR}}X_K + Z_K|\sqrt{\mathsf{INR}}X_K + Z_{K-1}) \tag{93} \\
&\le \frac{1}{2}\log\left(\frac{1+\mathsf{SNR}+\mathsf{INR}}{1+\mathsf{INR}}\right) + c \tag{94} \\
&= C - E + c. \tag{95}
\end{aligned}
$$

Finally, we have

$$
\begin{aligned}
h(Y_1,\ldots,Y_K|X_1,\ldots,X_K) &= h(Z_1,\ldots,Z_K) \tag{96} \\
&= Kc. \tag{97}
\end{aligned}
$$

Hence, from (84), we have

$$
\begin{aligned}
\mathcal{C}^{\mathrm{FB}}_{\mathrm{sum,G}}(K) &\le A + (K-2)(B-E) + C - E \tag{98} \\
&= K(B-E) + (A+C+E-2B), \tag{99}
\end{aligned}
$$

which implies that

$$
\begin{aligned}
\mathcal{C}^{\mathrm{FB}}_{\mathrm{sym,G}}(K) &\le (B-E) + \frac{A+C+E-2B}{K} \tag{100} \\
&= (B-E) + \frac{E}{K} + \frac{(A+C-2B)}{K} \tag{101} \\
&\le (B-E) + \frac{E}{K}, \tag{102}
\end{aligned}
$$

where we have used the fact that

$$
\begin{aligned}
A &\le B \tag{103} \\
C &\le B, \tag{104}
\end{aligned}
$$

which implies that $(A+C-2B) \le 0$.

We also note that

$$
\begin{aligned}
2B &= \log(1+\mathsf{SNR}+2\mathsf{INR}+\mathsf{INR}^2+2\sqrt{\mathsf{SNR}\cdot\mathsf{INR}}) \tag{105} \\
&\le \log(1+6\mathsf{SNR}) \tag{106} \\
&\le \log(6) + \log(1+\mathsf{SNR}) \tag{107} \\
&= \log(6) + 2D. \tag{108}
\end{aligned}
$$

Collecting all the bounds, we have

$$
\begin{aligned}
\mathcal{C}^{\mathrm{FB}}_{\mathrm{sym,G}}(K) &\le (B-E) + \frac{E}{K} \tag{109} \\
&\le (D-E) + \frac{E}{K} + \frac{\log(6)}{2}. \tag{110}
\end{aligned}
$$

Hence, the symmetric feedback capacity satisfies

$$
\begin{aligned}
&\left[(D-E) + \frac{E}{K}\right] - 1 - \frac{\log 3}{2K} \\
&\qquad \le \mathcal{C}^{\mathrm{FB}}_{\mathrm{sym,G}}(K) \\
&\qquad \le \left[(D-E) + \frac{E}{K}\right] + \frac{\log(6)}{2} \tag{111}
\end{aligned}
$$

so that the gap is given as

$$
\begin{aligned}
\Delta &= \frac{\log(6)}{2} + 1 + \frac{\log(3)}{2K} \tag{112} \\
&< \frac{3}{2} + 1 + \frac{2}{2K} \tag{113} \\
&\le \frac{11}{4}, \tag{114}
\end{aligned}
$$

and the degrees of freedom are given as

$$\mathbf{GDoF}^{\mathrm{FB}}(\alpha, K) = (1 - \alpha) + \frac{\alpha}{K}, \quad 0 \leq \alpha \leq 1/2. \quad (115)$$

B. Weak Interference: $1/2 \leq \alpha \leq 2/3$

1) Coding Scheme: In this regime we have $\mathsf{SNR}^{1/2} \leq \mathsf{INR} \leq \mathsf{SNR}^{2/3}$. The encoding scheme for this regime takes advantage of the feedback link. We create a cycle of length K consisting of all the interfering and feedback links. A part of the message of each user is conveyed through this cycle.

The encoding scheme is performed over K blocks. Assume each user has $2K + 1$ messages, namely

$$\left(M_k^{(1,h)}, M_k^{(2,h)}, \ldots, M_k^{(K,h)}, M_k^{(m)}, M_k^{(1,l)}, M_k^{(2,l)}, M_k^{(K,l)} \right),$$

The rates of the messages are given by[3]

$$\frac{\log |\mathcal{M}_k^{(j,h)}|}{n} = R_1 = \frac{1}{2} \log^+ \left(\frac{\mathsf{INR} + 1}{2\frac{\mathsf{SNR}}{\mathsf{INR}} + 1} \right),$$
$$k = 1, \ldots, K, \ j = 1, \ldots, K,$$
$$\frac{\log |\mathcal{M}_k^{(m)}|}{n} = R_2 = \frac{1}{2} \log^+ \left(\frac{1 + \mathsf{SNR}^2}{1 + 2\mathsf{INR}^3} \right), \quad (116)$$
$$k = 1, \ldots, K,$$
$$\frac{\log |\mathcal{M}_k^{(j,l)}|}{n} = R_3 = \frac{1}{2} \log^+ \left(\frac{\frac{\mathsf{SNR}}{\mathsf{INR}} + 2}{2} \right),$$
$$k = 1, \ldots, K, \ j = 1, \ldots, K.$$

The kth transmitter encodes its message using a individual Gaussian codebook with unit average power, to obtain

$$\left(\mathbf{s}_k^{(1,h)}, \ldots, \mathbf{s}_k^{(K,h)}, \mathbf{s}_k^{(m)}, \mathbf{s}_k^{(1,l)}, \mathbf{s}_k^{(K,l)} \right).$$

The transmitting signal in block k is formed as

$$\mathbf{x}_k[j] = \sqrt{\frac{\mathsf{INR}^2 - \mathsf{SNR}}{\mathsf{INR}^2}} \mathbf{x}_{k,h}[j] + \sqrt{\frac{\mathsf{SNR} - \mathsf{INR}}{\mathsf{INR}^2}} \mathbf{x}_{k,m}[j]$$
$$+ \sqrt{\frac{1}{\mathsf{INR}}} \mathbf{x}_{k,l}[j], \quad (117)$$

where $\mathbf{x}_{k,h}[j] = \mathbf{s}_k^{(j,h)}$ and $\mathbf{x}_{k,l}[j] = \mathbf{s}_k^{(j,l)}$. The moderate power codeword transmitted during the first block is the codeword corresponding to $M_k^{(m)}$. However, in the next block, $\mathbf{x}_{k,m}$ would be the moderate power codeword of the neighbor sent during the past block. More precisely,

$$\mathbf{x}_{k,m}[j] = \mathbf{x}_{k+1,m}[j-1] = \mathbf{x}_{k+2,m}[j-2] \cdots = \mathbf{x}_{k+j-1,m}[1]$$
$$= \mathbf{s}_{k+j-1}^{(m)}.$$

Note that $\mathbf{x}_{k+1,m}[j-1]$ has to be decoded from the signal sent to the transmitter k over the feedback link at the end of block $(j-1)$.

[3]Here we assume $\mathsf{SNR} \geq \mathsf{INR} \geq 1$ to guarantee positive rates. If this does not hold, similarly to the analysis in Appendix VIII-D, the bounded gap result can be established.

a) Decoding the feedback signal at encoder: The signal received at receiver k in block j is forwarded to its respective transmitter at the end of the block. The transmitter will use it for forming its transmitting signal in the next block. The transmitter k removes the part of the signal sent by it to obtain

$$\mathbf{y}_k[j] - \sqrt{\mathsf{SNR}} \mathbf{x}_k[j]$$
$$= \sqrt{\mathsf{INR}} \mathbf{x}_{k+1}[j] + \mathbf{z}_k[j]$$
$$= \sqrt{\frac{\mathsf{INR}^2 - \mathsf{SNR}}{\mathsf{INR}}} \mathbf{x}_{k+1,h}[j] + \sqrt{\frac{\mathsf{SNR} - \mathsf{INR}}{\mathsf{INR}}} \mathbf{x}_{k+1,m}[j]$$
$$+ \mathbf{x}_{k+1,l}[j] + \mathbf{z}_k[j]. \quad (118)$$

The transmitter needs the moderate power codeword for the next transmission. In order to decode $\mathbf{x}_{k+1,m}[j]$, it first decodes and removes the high power codeword, and then decodes the moderate power one. This can be done provided

$$R_1 \leq \frac{1}{2} \log \left(\frac{\mathsf{INR} + 1}{\frac{\mathsf{SNR}}{\mathsf{INR}} + 1} \right) \quad (119)$$

$$R_2 \leq \frac{1}{2} \log \left(\frac{\frac{\mathsf{SNR}}{\mathsf{INR}} + 1}{2} \right). \quad (120)$$

It is easy to check that the rates in (116) satisfy both constraints.

b) Decoding process at the receiver: The decoding procedure at decoder k is as follows. Upon receiving

$$\mathbf{y}_k[j] = \sqrt{\frac{\mathsf{SNR}}{\mathsf{INR}^2}(\mathsf{INR}^2 - \mathsf{SNR})} \mathbf{x}_{k,h}[j]$$
$$+ \sqrt{\frac{\mathsf{SNR}}{\mathsf{INR}^2}(\mathsf{SNR} - \mathsf{INR})} \mathbf{x}_{k,m}[j]$$
$$+ \sqrt{\frac{\mathsf{INR}^2 - \mathsf{SNR}}{\mathsf{INR}}} \mathbf{x}_{k+1,h}[j]$$
$$+ \sqrt{\frac{\mathsf{SNR}}{\mathsf{INR}}} \mathbf{x}_{k,l}[j] + \sqrt{\frac{\mathsf{SNR} - \mathsf{INR}}{\mathsf{INR}}} \mathbf{x}_{k+1,m}[j]$$
$$+ \mathbf{x}_{k+1,l}[j] + \mathbf{z}_k[k], \quad (121)$$

it decodes $\mathbf{x}_{k,h}[j]$, $\mathbf{x}_{k,m}[j]$, and $\mathbf{x}_{k+1,h}$ sequentially; that is in each step it treats everything else as noise, decodes the codeword, and removes the corresponding part from the received signal. This can be done as long as

$$R_1 \leq \frac{1}{2} \log \left(\frac{\mathsf{SNR} + \mathsf{INR} + 1}{\frac{\mathsf{SNR}^2}{\mathsf{INR}^2} + \mathsf{INR} + 1} \right) \quad (122)$$

$$R_2 \leq \frac{1}{2} \log \left(\frac{\frac{\mathsf{SNR}^2}{\mathsf{INR}^2} + \mathsf{INR} + 1}{\frac{\mathsf{SNR}}{\mathsf{INR}} + \mathsf{INR} + 1} \right) \quad (123)$$

$$R_1 \leq \frac{1}{2} \log \left(\frac{\frac{\mathsf{SNR}}{\mathsf{INR}} + \mathsf{INR} + 1}{2\frac{\mathsf{SNR}}{\mathsf{INR}} + 1} \right), \quad (124)$$

which are all satisfied with the rates in (116). The remaining part of the signal would be

$$\tilde{\mathbf{y}}_k[j] = \sqrt{\frac{\mathsf{SNR}}{\mathsf{INR}}} \mathbf{x}_{k,l}[j] + \sqrt{\frac{\mathsf{SNR} - \mathsf{INR}}{\mathsf{INR}}} \mathbf{x}_{k+1,m}[j]$$
$$+ \mathbf{x}_{k+1lm}[j] + \mathbf{z}_k[k], \quad (125)$$

which will be stored for further processing and for decoding $\mathbf{x}_{k,l}[j]$ later. In the next block, once $\mathbf{x}_{k,m}[j+1]$ is decoded, the decoder again recalls $\tilde{\mathbf{y}}_k[j]$ and subtracts from it the part corresponding to $\mathbf{x}_{k,m}[j+1] = \mathbf{x}_{k+1,m}[j]$. Therefore, it obtains

$$\tilde{\mathbf{y}}_k[j] - \sqrt{\frac{\mathsf{SNR} - \mathsf{INR}}{\mathsf{INR}}} \mathbf{x}_{k+1,m}[j]$$
$$= \sqrt{\frac{\mathsf{SNR}}{\mathsf{INR}}} \mathbf{x}_{k,l}[j] + \mathbf{x}_{k+1,l}[j] + \mathbf{z}_k[k], \tag{126}$$

from which $\mathbf{x}_{k,l}[j]$ can be decoded as long as

$$R_3 \leq \frac{1}{2} \log \left(\frac{\frac{\mathsf{SNR}}{\mathsf{INR}} + 2}{2} \right), \tag{127}$$

which clearly holds with R_3 in (116). Therefore, in each block, the receiver can decode the low power codeword of the previous block after removing the moderate power interfering signal. However, this process does not have to continue forever, since in the K-th block, the moderate power interfering signal at receiver k would be

$$\mathbf{x}_{k+1,m}[K] = \mathbf{x}_{k+2,m}[K-1] = \cdots = \mathbf{x}_{k+K,m}[1] = \mathbf{x}_{k,m}[1],$$

which was already decoded in the first block.

In summary, the total rate can be achieved per user per block would be

$$R_{\mathrm{sym}}$$
$$= \frac{KR_1 + R_2 + KR_3}{K} \tag{128}$$
$$= \frac{1}{2} \log (\mathsf{INR} + 1) + \frac{1}{2} \log \left(\frac{\frac{\mathsf{SNR}}{\mathsf{INR}} + 2}{2\frac{\mathsf{SNR}}{\mathsf{INR}} + 1} \right)$$
$$\quad + \frac{1}{K} \left[\frac{1}{2} \log(1 + \mathsf{SNR}^2) - \frac{1}{2} \log(1 + \mathsf{INR}^3) \right] \tag{129}$$
$$\geq \frac{1}{2} \log (\mathsf{INR} + 1) + \frac{1}{2} \log \frac{1}{2}$$
$$\quad + \frac{1}{K} \left[\frac{1}{2} \log(1 + \mathsf{SNR})^2 - \frac{1}{2} - \frac{1}{2} \log(1 + \mathsf{INR})^3 \right] \tag{130}$$
$$= E + \frac{(2D - 3E)}{K} - \frac{1}{2} - \frac{1}{2K}. \tag{131}$$

2) Upper Bound: In this regime, we use the same upper bound as in the case of very weak interference regime. In particular, from (99), we have

$$\mathcal{C}_{\mathrm{sym,G}}^{\mathrm{FB}}(K) \leq (B - E) + \frac{(A + C + E - 2B)}{K} \tag{132}$$
$$\leq 1 + E + \frac{(A + C + E - 4E)}{K} \tag{133}$$
$$\leq 1 + E + \frac{(1 + D + 1/2 + D + E - 4E)}{K} \tag{134}$$
$$= E + \frac{(2D - 3E)}{K} + 1 + \frac{3}{2K}. \tag{135}$$

Here, in (134), we have used the fact that for the weak interference regime, we have $2E \leq B \leq 1 + 2E$, $A \leq 1 + D$,

and $C \leq 1/2 + D$. Therefore, we have

$$\left[E + \frac{(2D - 3E)}{K} \right] - \frac{1}{2} - \frac{1}{2K}$$
$$\leq \mathcal{C}_{\mathrm{sym,G}}^{\mathrm{FB}}(K)$$
$$\leq \left[E + \frac{(2D - 3E)}{K} \right] + 1 + \frac{3}{2K}, \tag{136}$$

which implies that the gap is bounded as follows:

$$\Delta \leq \frac{3}{2} + \frac{2}{K} \tag{137}$$

which is at most 3 bits per user-pair and we have

$$\mathbf{GDoF}^{\mathrm{FB}}(\alpha, K) = \alpha + \frac{(2 - 3\alpha)}{K}, \ 1/2 \leq \alpha \leq 2/3. \tag{138}$$

C. Moderate Interference: $2/3 \leq \alpha \leq 1$

1) Coding Scheme: In this regime we use the private and common message for the encoding scheme. Assume each transmitter has two messages, namely the high power (common) message $M_k^{(h)}$, and the low power (private) message $M_k^{(l)}$. The following transmission scheme aims to convey the common message $M_k^{(h)}$ to both receivers k and $k+1$. However, the private message $M_k^{(l)}$ can be decoded only by the respective receiver.

We assume that the high power messages of all users have the same rate. Similarly, the rate of the low power messages for all users are the same; that is

$$R_1 = \frac{\log |\mathcal{M}_k^{(h)}|}{n}, \qquad k = 1, \ldots, K,$$
$$R_2 = \frac{\log |\mathcal{M}_k^{(l)}|}{n}, \qquad k = 1, \ldots, K. \tag{139}$$

The encoder first maps its messages to Gaussian codewords with unit average power, $\mathbf{x}_{k,h}$ and $\mathbf{x}_{k,l}$, and sends

$$\mathbf{x}_k = \sqrt{\frac{\mathsf{INR} - 1}{\mathsf{INR}}} \mathbf{x}_{k,h} + \sqrt{\frac{1}{\mathsf{INR}}} \mathbf{x}_{k,l}.$$

The receiver node k, upon receiving $\mathbf{y}_k$, with

$$\mathbf{y}_k[j]$$
$$= \sqrt{\mathsf{SNR}} \mathbf{x}_k[j] + \sqrt{\mathsf{INR}} \mathbf{x}_{k+1}[j] + \mathbf{z}_k[j] \tag{140}$$
$$= \sqrt{\frac{\mathsf{SNR}}{\mathsf{INR}}(\mathsf{INR} - 1)} \mathbf{x}_{k,h}[j] + \sqrt{\mathsf{INR} - 1} \mathbf{x}_{k+1,h}[j]$$
$$\quad + \sqrt{\frac{\mathsf{SNR}}{\mathsf{INR}}} \mathbf{x}_{k,l}[j] + \mathbf{x}_{k+1,l}[j] + \mathbf{z}_k[j], \tag{141}$$

first jointly decodes the codewords $\mathbf{x}_{k,h}^{(h)}$ and $\mathbf{x}_{k+1,h}^{(h)}$ treating all the rest as noise. Here we deal with a multiple access channel, whose achievable rate is characterized by

$$R_1 \leq \frac{1}{2} \log \left(\frac{\mathsf{SNR} + 2}{\frac{\mathsf{SNR}}{\mathsf{INR}} + 2} \right), \tag{142}$$
$$R_1 \leq \frac{1}{2} \log \left(\frac{\mathsf{INR} + \frac{\mathsf{SNR}}{\mathsf{INR}} + 1}{\frac{\mathsf{SNR}}{\mathsf{INR}} + 2} \right), \tag{143}$$
$$2R_1 \leq \frac{1}{2} \log \left(\frac{\mathsf{SNR} + \mathsf{INR} + 1}{\frac{\mathsf{SNR}}{\mathsf{INR}} + 2} \right). \tag{144}$$

In particular, it is easy to show that[4]

$$R_1 = \frac{1}{4}\log^+\left(\frac{\mathsf{SNR}+\mathsf{INR}+1}{\frac{\mathsf{SNR}}{\mathsf{INR}}+2}\right) \quad (145)$$

satisfies the above constraints. After decoding the high power codewords, and removing them from the received signal, receiver k decodes its own low power message by treating the other private codeword as noise. This private message can be reliably decoded provided that

$$R_2 \leq \frac{1}{2}\log\left(\frac{\frac{\mathsf{SNR}}{\mathsf{INR}}+2}{2}\right), \quad (146)$$

which yields an achievable total rate of

$$R_{\mathrm{sym}} = R_1 + R_2 \quad (147)$$

$$= \frac{1}{4}\log\left(\mathsf{SNR}+\mathsf{INR}+1\right) + \frac{1}{4}\log\left(2+\frac{\mathsf{SNR}}{\mathsf{INR}}\right) - \frac{1}{2} \quad (148)$$

$$\geq \frac{1}{2}\log(1+\mathsf{SNR}) - \frac{1}{4}\log(1+\mathsf{INR}) - \frac{1}{2} \quad (149)$$

$$= D - \frac{E}{2} - \frac{1}{2}. \quad (150)$$

2) Upper Bound: In this regime, we will develop a different upper bound that is analogous to the type-I upper bound obtained for the linear deterministic channel model. We have the following bound on the feedback sum capacity:

$$\mathcal{C}_{\mathrm{sum,G}}^{\mathrm{FB}}(K) \leq \frac{K}{2}(A+C-E). \quad (151)$$

The proof of (151) is given in the appendix. Hence, (151) implies that the symmetric feedback capacity is upper bounded as

$$\mathcal{C}_{\mathrm{sym,G}}^{\mathrm{FB}}(K) \leq \frac{A+C-E}{2} \quad (152)$$

$$= \frac{A+C}{2} - \frac{E}{2}. \quad (153)$$

Note that in this regime, we have

$$2A = \log(1+\mathsf{SNR}+\mathsf{INR}+2\sqrt{\mathsf{SNR}\cdot\mathsf{INR}}) \quad (154)$$

$$\leq \log(1+4\mathsf{SNR}) \quad (155)$$

$$\leq \log(4) + \log(1+\mathsf{SNR}) \quad (156)$$

$$= 2 + 2D, \quad (157)$$

and similarly,

$$2C = \log(1+\mathsf{SNR}+\mathsf{INR}) \quad (158)$$

$$\leq \log(1+2\mathsf{SNR}) \quad (159)$$

$$\leq \log(2) + \log(1+\mathsf{SNR}) \quad (160)$$

$$= 1 + 2D, \quad (161)$$

which implies that

$$\frac{A+C}{2} \leq D + \frac{3}{4}. \quad (162)$$

---

[4]We assume $\mathsf{INR} \geq 1$. If this condition does not hold, then a bounded gap result can be shown in a similar fashion to Appendix VIII-D.

Hence, we have

$$\left[D - \frac{E}{2}\right] - \frac{1}{2} \leq \mathcal{C}_{\mathrm{sym,G}}^{\mathrm{FB}}(K) \leq \left[D - \frac{E}{2}\right] + \frac{3}{4}, \quad (163)$$

so that the gap between the upper and lower bounds is at most $5/4$ bits and we have

$$\mathbf{GDoF}^{\mathrm{FB}}(\alpha, K) = 1 - \frac{\alpha}{2}, \quad 2/3 \leq \alpha \leq 1. \quad (164)$$

D. Strong Interference: $1 \leq \alpha \leq 2$

1) Coding Scheme: In this regime, we have $\mathsf{SNR} \leq \mathsf{INR} \leq \mathsf{SNR}^2$. The encoding scheme for this interference regime is simple, and the desired degrees of freedom can be achieved in one block. Denote the message of user k by $M_k \in \mathcal{M}_k$, where all the message sets have the same size which results in a symmetric rate of R. Each user takes a random Gaussian codebook with rate R and unit average power. Then it randomly maps its message to $\mathbf{x}_k$ and sends it over the channel. The k-th receiver observes $\mathbf{y}_k$ through a multiple access channel from the k-th and $(k+1)$-th transmitters, in which it has to decode both messages. The achievable rate of the MAC is well-known as [15]

$$R \leq \frac{1}{2}\log\left(\mathsf{SNR}+1\right), \quad (165)$$

$$R \leq \frac{1}{2}\log\left(\mathsf{INR}+1\right), \quad (166)$$

$$2R \leq \frac{1}{2}\log\left(\mathsf{INR}+\mathsf{SNR}+1\right). \quad (167)$$

Hence, it is clear that by choosing

$$R_{\mathrm{sym}} = \frac{1}{4}\log\left(1+\mathsf{INR}+\mathsf{SNR}\right) \quad (168)$$

$$= \frac{C}{2}, \quad (169)$$

all constraints are satisfied and a symmetric rate of R_{sym} is therefore achievable.

2) Upper bound: For this regime, we use the same upper bound developed in the previous section:

$$\mathcal{C}_{\mathrm{sym,G}}^{\mathrm{FB}}(K) \leq \frac{A+C-E}{2}. \quad (170)$$

Therefore, the symmetric feedback capacity satisfies

$$\frac{C}{2} \leq \mathcal{C}_{\mathrm{sym,G}}^{\mathrm{FB}}(K) \leq \frac{C}{2} + \frac{A-E}{2}, \quad (171)$$

and the gap between the bounds is

$$\Delta = \frac{A-E}{2} \quad (172)$$

$$= \frac{\log(1+\mathsf{SNR}+\mathsf{INR}+2\sqrt{\mathsf{SNR}\cdot\mathsf{INR}}) - \log(1+\mathsf{INR})}{4} \quad (173)$$

$$\leq \frac{\log(1+4\mathsf{INR}) - \log(1+\mathsf{INR})}{4} \quad (174)$$

$$\leq \frac{\log(4) + \log(1+\mathsf{INR}) - \log(1+\mathsf{INR})}{4} \quad (175)$$

$$= \frac{1}{2}. \quad (176)$$

Moreover, from (171), it is straightforward to show that

$$\mathbf{GDoF}^{\mathrm{FB}}(\alpha, K) = \frac{\alpha}{2}, \quad 1 \leq \alpha \leq 2. \quad (177)$$

E. Very-Strong Interference: $\alpha \geq 2$

1) Coding Scheme: The encoding scheme we propose here takes K blocks, each of length T. We assume each user k has a total of $K+1$ messages, namely $\left(M_k^{(l)}, M_k^{(1,h)}, \ldots, M_k^{(K,h)} \right)$, that it wishes to send to its respective receiver over K transmission blocks. We assume that $M_k^{(l)} \in \mathcal{M}_k^{(l)}$ and $M_k^{(j,h)} \in \mathcal{M}_k^{(j,h)}$, where $\mathcal{M}$'s are the message sets. Moreover, we set the size of these message sets so that[5]

$$
\begin{aligned}
\frac{\log |\mathcal{M}_k^{(l)}|}{n} &= R_1 = \frac{1}{2}\log^+\left(\frac{\frac{\text{INR}}{\text{SNR}^2}+1}{2} \right), \\
&\quad k = 1, \ldots, K \\
\frac{\log |\mathcal{M}_k^{(j,h)}|}{n} &= R_2 = \frac{1}{2}\log^+\left(\frac{\text{SNR}+1}{3} \right), \\
&\quad k = 1, \ldots, K, \; j = 1, \ldots, K.
\end{aligned}
\tag{178}
$$

That is, all the first messages of all the users have the same rate. Furthermore, the rates of all the remaining messages are also identical. Each message is encoded by a capacity achieving Gaussian codebook with unit variance. Hence user k has $K+1$ Gaussian codewords, $\mathbf{s}_k^{(l)}, \mathbf{s}_k^{(1,h)}, \ldots, \mathbf{s}_k^{(K,h)}$, each of length n.

The signal sent by transmitter k in block j is composed of two parts, the high power part $\mathbf{x}_{k,h}[j]$ and low power $\mathbf{x}_{k,l}[j]$:

$$
\mathbf{x}_k[j] = \sqrt{\frac{\text{SNR}-1}{\text{SNR}}}\mathbf{x}_{k,h}[j] + \sqrt{\frac{1}{\text{SNR}}}\mathbf{x}_{k,l}[j].
\tag{179}
$$

In all blocks, the high power part is the codeword corresponding to a fresh message. In the first block, since the nodes have not yet received any feedback, their low power codewords also describe fresh messages. However, for all blocks $j \geq 2$, the low level codeword used to form the transmitting signal is the low level codeword of their neighbor sent on the previous block. We will show that this message can be decoded from the signal received over the feedback link at the end of the last block. More precisely,

$$
\mathbf{x}_{k,h}[j] = \mathbf{s}_k^{(j,h)}, \quad k = 1, 2, \ldots, K, \; j = 1, 2, \ldots, K,
\tag{180}
$$

$$
\mathbf{x}_{k,l}[1] = \mathbf{s}_k^{(l)}, \quad k = 1, 2, \ldots, K,
\tag{181}
$$

$$
\mathbf{x}_{k,l}[j] = \mathbf{x}_{k+1,l}[j-1], \quad k = 1, 2, \ldots, K, \; j = 2, \ldots, K.
\tag{182}
$$

Therefore we have the following recursive relationship between the low power codewords:

$$
\begin{aligned}
\mathbf{x}_{k,l}[j] &= \mathbf{x}_{k+1,l}[j-1] = \mathbf{x}_{k+2,l}[j-2] = \cdots = \mathbf{x}_{k+j-1,l}[1] \\
&= \mathbf{s}_{k+j-1}^{(l)},
\end{aligned}
\tag{183}
$$

where all the user and block indicators are modulo K, e.g., $\mathbf{x}_{k+j-1}[1] = \mathbf{x}_{(k+j-1) \mod K}[1]$.

[5]Here we assume $\text{SNR} \geq 2$ to make sure that R_2 is positive. Note that R_1 is positive since $\text{INR} \geq \text{SNR}^2$. A similar analysis as in Appendix VIII-D can be done if $\text{SNR} < 2$.

a) Decoding the feedback signal at encoder: As stated above, in order to form the transmitting signal in block j, transmitter k uses the low power codeword sent by user $k+1$ in block $j-1$. We first show that this codeword can be decoded based on the signal it receives over the feedback link at the end of block $j-1$.

Once $\mathbf{y}_k[j-1]$ is received, transmitter k first removes its own signal, $\mathbf{x}_k[j-1]$, to obtain

$$
\begin{aligned}
&\mathbf{y}_k[j-1] - \sqrt{\text{SNR}}\mathbf{x}_k[j-1] \\
&= \sqrt{\text{INR}}\mathbf{x}_{k+1}[j-1] + \mathbf{z}_k[j-1]
\end{aligned}
\tag{184}
$$

$$
\begin{aligned}
&= \sqrt{\frac{\text{INR}}{\text{SNR}}(\text{SNR}-1)}\mathbf{x}_{k+1,h}[j-1] + \sqrt{\frac{\text{INR}}{\text{SNR}}}\mathbf{x}_{k+1,l}[j-1] \\
&\quad + \mathbf{z}_k[j-1],
\end{aligned}
\tag{185}
$$

from which it has to decode both $\mathbf{x}_{k+1,h}[j-1]$ and $\mathbf{x}_{k+1,l}[j-1]$. It first decodes $\mathbf{x}_{k+1,h}[j-1]$, treating everything else as noise. Then, it removes $\mathbf{x}_{k+1,h}[j-1]$ from the signal and decodes $\mathbf{x}_{k+1,l}[j-1]$ in a similar manner. This is possible as long as

$$
R_2 \leq \frac{1}{2}\log\left(\frac{\text{INR}+1}{\frac{\text{INR}}{\text{SNR}}+1} \right)
\tag{186}
$$

$$
R_1 \leq \frac{1}{2}\log\left(\frac{\text{INR}}{\text{SNR}}+1 \right),
\tag{187}
$$

which are clearly satisfied by the rates chosen in (178). Therefore, the transmitter k has access to $\mathbf{x}_{k+1,l}[j-1]$, which will be used as its low power codeword for block j.

b) Decoding process at the receiver: The signal sent by user k over the j-th block is given by

$$
\mathbf{x}_k[j] = \sqrt{\frac{\text{SNR}-1}{\text{SNR}}}\mathbf{x}_{k,h}[j] + \sqrt{\frac{1}{\text{SNR}}}\mathbf{x}_{k,l}[j],
\tag{188}
$$

which results in

$$
\mathbf{y}_k[j] = \sqrt{\text{SNR}}\mathbf{x}_k[j] + \sqrt{\text{INR}}\mathbf{x}_{k+1}[j] + \mathbf{z}_k[j]
\tag{189}
$$

$$
\begin{aligned}
&= \sqrt{\frac{\text{INR}}{\text{SNR}}(\text{SNR}-1)}\mathbf{x}_{k+1,h}[j] + \sqrt{\frac{\text{INR}}{\text{SNR}}}\mathbf{x}_{k+1,l}[j] \\
&\quad + \sqrt{\text{SNR}-1}\mathbf{x}_{k,h}[j] + \mathbf{x}_{k,l}[j] + \mathbf{z}_k[j].
\end{aligned}
\tag{190}
$$

At the end of the j-th block, user k sequentially decodes the codewords $\mathbf{x}_{k+1,h}[j]$, $\mathbf{x}_{k+1,l}[j]$, and $\mathbf{x}_{k,h}[j]$. At each step, it decodes the corresponding codewords, treating all the remaining parts as noise. Once one codeword is decoded, it removes it from its received signal, and proceeds with the next codeword. This can be done provided that

$$
R_2 \leq \frac{1}{2}\log\left(\frac{\text{INR}+\text{SNR}+1}{\frac{\text{INR}}{\text{SNR}}+\text{SNR}+1} \right),
\tag{191}
$$

$$
R_1 \leq \frac{1}{2}\log\left(\frac{\frac{\text{INR}}{\text{SNR}}+\text{SNR}+1}{\text{SNR}+1} \right),
\tag{192}
$$

$$
R_2 \leq \frac{1}{2}\log\left(\frac{\text{SNR}+1}{2} \right).
\tag{193}
$$

It is easy to check that all constraints are satisfied by the choice of R_1 and R_2 in (178).

At the end of each block, each receiver can decode its respective high power codeword, as well as some high power and low power codewords from other users which it is not intended to decode. However, from (183), the low power codeword decoded at receiver k at the very last block would be

$$\mathbf{x}_{k+1,l}[K] = \mathbf{x}_{k+2,l}[K-1] = \cdots = \mathbf{x}_{k+K,l}[1] \stackrel{(*)}{=} \mathbf{x}_{k,l}[1] = \mathbf{s}_k^{(l)},$$

where $(*)$ holds since $k + K = k \mod K$. Therefore all the intended messages for receiver k, can be decoded using this scheme in K blocks. The total rate of communication would be

$$
\begin{aligned}
R_{\text{sym}} = \frac{R_1 + K R_2}{K} &= \frac{1}{2} \log \left(\mathsf{SNR} + 1 \right) \\
&+ \frac{1}{2K} \log \left(\frac{\mathsf{INR}}{\mathsf{SNR}^2} + 1 \right) - \frac{K \log 3 + 1}{2K} \quad (194) \\
&= D + \frac{1}{2K} \log \left(\frac{\mathsf{INR}}{\mathsf{SNR}^2} + 1 \right) - \frac{K \log 3 + 1}{2K} \quad (195) \\
&\geq D + \frac{1}{2K} \log \left(\frac{1 + \mathsf{INR}}{(1 + \mathsf{SNR})^2} \right) - \frac{K \log 3 + 1}{2K} \quad (196) \\
&= D + \frac{(E - 2D)}{K} - \frac{(K \log 3 + 1)}{2K}. \quad (197)
\end{aligned}
$$

2) Upper Bound: For this regime we use the following upper bound from Theorem 5, similar to the LD case:

$$
\begin{aligned}
\mathcal{C}_{\text{sum,G}}^{\text{FB}}(K) \\
\leq \max_{p(x_1,\ldots,x_K)} &\Big[h(Y_K) + h(Y_{K-1}|X_K, Y_K) + \ldots \\
&+ h(Y_1|X_2, Y_2, \ldots, X_K, Y_K) \\
&- h(Y_1, \ldots, Y_K | X_1, \ldots, X_K) \Big]. \quad (198)
\end{aligned}
$$

We upper bound the first term in (198) as

$$h(Y_K) \leq \frac{1}{2} \log \left(1 + \mathsf{SNR} + \mathsf{INR} + 2\sqrt{\mathsf{SNR} \cdot \mathsf{INR}} \right) + c \quad (199)$$

$$= A + c, \quad (200)$$

where $c = (1/2) \log(2\pi e)$.

For any $2 \leq k \leq (K-1)$, we bound

$$h(Y_k|X_{k+1}, Y_{k+1}, \ldots, X_K, Y_K) \leq h(Y_k|X_{k+1}) \quad (201)$$

$$= h(\sqrt{\mathsf{SNR}}X_k + Z_k | X_{k+1}) \quad (202)$$

$$\leq h(\sqrt{\mathsf{SNR}}X_k + Z_k) \quad (203)$$

$$\leq \frac{1}{2} \log (1 + \mathsf{SNR}) + c \quad (204)$$

$$= D + c. \quad (205)$$

Finally, we bound the penultimate term in (198) as follows:

$$h(Y_1|X_2, Y_2, \ldots, X_K, Y_K)$$

$$\leq h(Y_1|X_2, Y_2, X_K, Y_K) \quad (206)$$

$$= h(\sqrt{\mathsf{SNR}}X_1 + Z_1 | X_2, Y_2, X_K, Y_K) \quad (207)$$

$$\leq h(\sqrt{\mathsf{SNR}}X_1 + Z_1 | \sqrt{\mathsf{INR}}X_1 + Z_K) \quad (208)$$

$$\leq \frac{1}{2} \log \left(\frac{1 + \mathsf{SNR} + \mathsf{INR}}{1 + \mathsf{INR}} \right) + c \quad (209)$$

$$= C - E + c, \quad (210)$$

where in (208), we used the fact that $Y_K = \sqrt{\mathsf{SNR}}X_K + \sqrt{\mathsf{INR}}X_1 + Z_K$ and the fact that conditioning reduces differential entropy. Finally, we note that

$$h(Y_1, \ldots, Y_K | X_1, \ldots, X_K) = h(Z_1, \ldots, Z_K | X_1, \ldots, X_K) \quad (211)$$

$$= h(Z_1, \ldots, Z_K) \quad (212)$$

$$= \sum_{k=1}^{K} h(Z_k) \quad (213)$$

$$= Kc. \quad (214)$$

Hence, the feedback sum capacity is upper bounded as follows:

$$\mathcal{C}_{\text{sum,G}}^{\text{FB}}(K) \leq A + (K-2)D + C - E, \quad (215)$$

which implies that the symmetric feedback capacity satisfies

$$\mathcal{C}_{\text{sym,G}}^{\text{FB}}(K) \leq D + \frac{(A + C - 2D - E)}{K}. \quad (216)$$

We now simplify this upper bound to compare it with the lower bound obtained in (197).

We note that

$$2A = \log(1 + \mathsf{SNR} + \mathsf{INR} + 2\sqrt{\mathsf{SNR} \cdot \mathsf{INR}}) \quad (217)$$

$$\leq \log(1 + 4\mathsf{INR}) \quad (218)$$

$$\leq \log(4) + \log(1 + \mathsf{INR}) \quad (219)$$

$$= 2 + 2E, \quad (220)$$

and

$$2C = \log(1 + \mathsf{SNR} + \mathsf{INR}) \quad (221)$$

$$\leq \log(1 + 2\mathsf{INR}) \quad (222)$$

$$\leq \log(2) + \log(1 + \mathsf{INR}) \quad (223)$$

$$= 1 + 2E, \quad (224)$$

which together imply that

$$A + C - 2D - E \leq (3/2) + 2E - 2D - E \quad (225)$$

$$= (E - 2D) + 3/2. \quad (226)$$

Hence, from (197) and (216), the symmetric feedback capacity satisfies

$$
\begin{aligned}
\left[D + \frac{(E - 2D)}{K} \right] &- \frac{(K \log(3) + 1)}{2K} \\
&\leq \mathcal{C}_{\text{sym,G}}^{\text{FB}}(K) \\
&\leq \left[D + \frac{(E - 2D)}{K} \right] + \frac{3}{2K} \quad (227)
\end{aligned}
$$

which implies that the gap is given as

$$\Delta = \frac{4 + K \log(3)}{2K} \tag{228}$$

$$\leq \frac{2 + K}{K} \tag{229}$$

which is at most 2-bits. We note here that the gap of 2-bits can be reduced further to 1-bit by modifying the power allocation in our coding scheme. The resulting gap analysis is however complicated and is not pursued here.

Moreover, from (227), we have

$$\mathbf{GDoF}^{\mathrm{FB}}(\alpha, K) = 1 + \frac{(\alpha - 2)}{K}, \quad \alpha \geq 2. \tag{230}$$

VII. CONCLUSIONS

In this paper, we have considered the K-user cyclic Z-interference channel with noiseless feedback. The symmetric feedback capacity of the linear deterministic CZIC has been completely characterized for all interference regimes. Using insights from the linear model, the symmetric feedback capacity for the Gaussian CZIC has been characterized within a constant number of bits for all interference regimes. As a consequence of the constant bit gap result, the symmetric feedback degrees of freedom for the Gaussian CZIC have also been characterized. It has been shown that the capacity gain obtained via feedback decreases as the number of users increases. The resulting $\mathbf{GDoF}^{\mathrm{FB}}(\alpha, K)$ for $K > 2$ users is a skewed V-curve, as a function of the interference parameter α. Moreover as $K \to \infty$, the resulting skewed V-curve converges to the well known W-curve corresponding to the no-feedback $\mathbf{GDoF}$. As part of future work, we believe that the characterization of the approximate feedback capacity region of the K-user Gaussian CZIC is an interesting problem. Moreover, we believe that new coding schemes and novel upper bounds would be required to achieve this goal.

VIII. APPENDIX

A. Proof of Theorem 4

We show that the normalized symmetric feedback capacity is upper bounded as follows:

$$\mathcal{C}_{\mathrm{sym,LD}}^{\mathrm{FB}}(\alpha, K) \leq \max\left(1 - \frac{\alpha}{2}, \frac{\alpha}{2}\right). \tag{231}$$

To prove (231), we first prove the following upper bound on the sum of the rates of users 1 and 2:

$$T(R_1 + R_2)$$
$$= H(W_1) + H(W_2) \tag{232}$$
$$= H(W_1|W_3, \ldots, W_K) + H(W_2|W_1, W_3, \ldots, W_K) \tag{233}$$
$$\leq I(W_1; Y_1^T|W_3, \ldots, W_K)$$
$$\quad + I(W_2; Y_2^T, Y_1^T|W_1, W_3, \ldots, W_K) + \epsilon_T \tag{234}$$
$$= I(W_1; Y_1^T|W_3, \ldots, W_K)$$
$$\quad + H(Y_2^T, Y_1^T|W_1, W_3, \ldots, W_K) + \epsilon_T \tag{235}$$

$$= H(Y_1^T|W_3, \ldots, W_K)$$
$$\quad + H(Y_2^T|Y_1^T, W_1, W_3, \ldots, W_K) + \epsilon_T \tag{236}$$
$$\leq H(Y_1^T) + H(Y_2^T|Y_1^T, W_1, W_3, \ldots, W_K) + \epsilon_T \tag{237}$$
$$\leq T \max(m, n) + H(Y_2^T|Y_1^T, W_1, W_3, \ldots, W_K) + \epsilon_T \tag{238}$$
$$\leq T \max(m, n) + T(n - m)^+ + \epsilon_T, \tag{239}$$

where (233) follows from the fact that the messages $(W_1, \ldots, W_K)$ are all mutually independent, (234) follows from Fano's inequality [15], (235) follows from the deterministic nature of the channel model and (237) follows from the fact that conditioning reduces entropy.

Before proving (239) we first prove the following claim:

Claim 1. $(X_{1t}, X_{3t}, \ldots, X_{Kt})$ *is a deterministic function of* $(Y_1^{t-1}, W_1, W_3, \ldots, W_K)$.

Proof: First note that from (3), we have

$$X_{1t} = f_{1t}\left(W_1, Y_1^{t-1}\right), \tag{240}$$

and

$$X_{Kt} = f_{Kt}\left(W_K, Y_K^{t-1}\right) \tag{241}$$
$$= f_{Kt}\left(W_K, X_1^{t-1}, X_K^{t-1}\right), \tag{242}$$

which together imply that

$$(X_{1t}, X_{Kt}, X_1^{t-1}, X_K^{t-1}) = f\left(W_1, W_K, Y_1^{t-1}\right). \tag{243}$$

Repeating this argument for $k = K - 1, \ldots, 3$, the proof of the claim is straightforward. ∎

We now bound the second term in (238) as follows:

$$H(Y_2^T|Y_1^T, W_1, W_3, \ldots, W_K)$$
$$\leq \sum_{t=1}^{T} H(Y_{2t}|Y_{1t}, W_1, W_3, \ldots, W_K, Y_1^{t-1}) \tag{244}$$
$$= \sum_{t=1}^{T} H(Y_{2t}|Y_{1t}, X_{1t}, W_1, X_{3t}, W_3, \ldots, X_{Kt}, W_K, Y_1^{t-1}) \tag{245}$$
$$\leq \sum_{t=1}^{T} H(Y_{2t}|Y_{1t}, X_{1t}, X_{3t}) \tag{246}$$
$$\leq \sum_{t=1}^{T} H(X_{2t}|Y_{1t}, X_{1t}, X_{3t}) \tag{247}$$
$$\leq T(n - m)^+, \tag{248}$$

where (245) follows from Claim 1 and (248) follows from the fact that (X_{1t}, Y_{1t}) completely determine at least m levels of X_{2t}. This completes the proof of (239). Dividing (239) by nT and taking the limit $T \to \infty$, we have $\epsilon_T \to 0$, which yields

$$\frac{R_1 + R_2}{n} \leq \max\left(\frac{m}{n}, 1\right) + \left(1 - \frac{m}{n}\right)^+ \tag{249}$$
$$= \max(\alpha, 1) + (1 - \alpha)^+ \tag{250}$$
$$= \max(2 - \alpha, \alpha). \tag{251}$$

In a similar manner it can be shown that for any $1 \leq j \leq K$,

$$\frac{R_j + R_{(j+1)\mathbf{mod}(K)}}{n} \leq \max(2 - \alpha, \alpha). \tag{252}$$

Adding all such K upper bounds, we obtain

$$\frac{2(R_1 + \ldots + R_K)}{n} \leq K \max\left(2 - \alpha, \alpha\right), \qquad (253)$$

and hence,

$$\mathcal{C}_{\text{sym,LD}}^{\text{FB}}(\alpha, K) \leq \max\left(1 - \frac{\alpha}{2}, \frac{\alpha}{2}\right). \qquad (254)$$

This upper bound on the normalized symmetric feedback capacity is independent of K and is the same as the normalized symmetric capacity *without feedback* when $\alpha \in [2/3, 2]$. Hence, for this interference regime, feedback does not increase the symmetric capacity. Also note that the range of α in deriving these bounds is immaterial and hence from a symmetric feedback capacity point of view, the feedback capacity for $K = 2$ users always serves as an upper bound for any $K > 2$.

B. Proof of Theorem 5

In this section we provide the proof for Theorem 5 for the special case in which $K = 3$ and $\pi = (1, 2, 3)$ (identity permutation). The generalization to arbitrary (K, π) is straight-forward.

For the 3-user interference channel with local feedback, we have the following upper bound on the sum-rate:

$$T(R_1 + R_2 + R_3)$$
$$= H(W_1) + H(W_2) + H(W_3) \qquad (255)$$
$$= H(W_1) + H(W_2|W_1) + H(W_3|W_1, W_2) \qquad (256)$$
$$\leq I(W_1; Y_1^T) + I(W_2; Y_2^T, Y_1^T|W_1)$$
$$\quad + I(W_3; Y_3^T, Y_2^T, Y_1^T|W_1, W_2) + \epsilon_T \qquad (257)$$
$$= H(Y_1^T) + H(Y_2^T, Y_1^T|W_1) + H(Y_3^T, Y_2^T, Y_1^T|W_1, W_2)$$
$$\quad - H(Y_1^T|W_1) - H(Y_2^T, Y_1^T|W_1, W_2)$$
$$\quad - H(Y_3^T, Y_2^T, Y_1^T|W_1, W_2, W_3) + \epsilon_T \qquad (258)$$
$$= H(Y_1^T) + H(Y_2^T|Y_1^T, W_1) + H(Y_3^T|Y_2^T, Y_1^T, W_1, W_2)$$
$$\quad - H(Y_3^T, Y_2^T, Y_1^T|W_1, W_2, W_3) + \epsilon_T \qquad (259)$$
$$\leq \sum_{t=1}^{T}\Big[H(Y_{1t}|Y_1^{t-1}) + H(Y_{2t}|Y_{1t}, Y_2^{t-1}, Y_1^{t-1}, W_1)$$
$$\quad + H(Y_{3t}|Y_{2t}, Y_{1t}, Y_3^{t-1}, Y_2^{t-1}, Y_1^{t-1}, W_1, W_2)$$
$$\quad - H(Y_{3t}, Y_{2t}, Y_{1t}|W_1, W_2, W_3, Y_1^{t-1}, Y_2^{t-1}, Y_3^{t-1})\Big] + \epsilon_T \qquad (260)$$
$$\leq \sum_{t=1}^{T}\Big[H(Y_{1t}) + H(Y_{2t}|Y_{1t}, X_{1t})$$
$$\quad + H(Y_{3t}|Y_{2t}, Y_{1t}, X_{2t}, X_{1t})$$
$$\quad - H(Y_{3t}, Y_{2t}, Y_{1t}|X_{1t}, X_{2t}, X_{3t})\Big] + \epsilon_T \qquad (261)$$
$$\leq T \max_{p(x_1, x_2, x_3)} \Big[H(Y_1) + H(Y_2|Y_1, X_1)$$
$$\quad + H(Y_3|Y_2, Y_1, X_2, X_1)$$
$$\quad - H(Y_3, Y_2, Y_1|X_1, X_2, X_3)\Big] + \epsilon_T, \qquad (262)$$

where (256) follows from the independence of the messages, (257) follows from Fano's inequality [15], and (259) follows from the fact that the negative term corresponding to the kth

mutual information is canceled by a part of the positive term in the $(k + 1)$th mutual information, for $k = 1, \ldots, (K - 1)$. Finally, dividing (262) by T and letting $T \to \infty$, we have the proof of Theorem 5.

C. Proof of (151)

We first obtain a bound on the sum of the rates of users 1 and 2:

$$T(R_1 + R_2)$$
$$= H(W_1) + H(W_2) \qquad (263)$$
$$= H(W_1|W_3, \ldots, W_K) + H(W_2|W_1, W_3, \ldots, W_K) \qquad (264)$$
$$\leq I(W_1; Y_1^T, Z_3^T, \ldots, Z_K^T|W_3, \ldots, W_K)$$
$$\quad + I(W_2; Y_2^T, Y_1^T, Z_3^T, \ldots, Z_K^T|W_1, W_3, \ldots, W_K) + \epsilon_T \qquad (265)$$
$$= h(Y_1^T, Z_3^T, \ldots, Z_K^T|W_3, \ldots, W_K)$$
$$\quad + h(Y_2^T|Y_1^T, Z_3^T, \ldots, Z_K^T, W_1, W_3, \ldots, W_K)$$
$$\quad - h(Y_2^T, Y_1^T, Z_3^T, \ldots, Z_K^T|W_1, W_2, W_3 \ldots, W_K) + \epsilon_T \qquad (266)$$
$$= h(Y_1^T, Z_3^T, \ldots, Z_K^T|W_3, \ldots, W_K)$$
$$\quad + h(Y_2^T|Y_1^T, Z_3^T, \ldots, Z_K^T, W_1, W_3, \ldots, W_K)$$
$$\quad - \sum_{t=1}^{T} h(Y_{2t}, Y_{1t}, Z_{3t}, \ldots, Z_{Kt}|W_1, \ldots, W_K, Y_2^{t-1}, Y_1^{t-1},$$
$$\qquad\qquad\qquad Z_3^{t-1}, \ldots, Z_K^{t-1}) + \epsilon_T \qquad (267)$$
$$= h(Y_1^T, Z_3^T, \ldots, Z_K^T|W_3, \ldots, W_K) + \epsilon_T$$
$$\quad + h(Y_2^T|Y_1^T, Z_3^T, \ldots, Z_K^T, W_1, W_3, \ldots, W_K)$$
$$\quad - \sum_{t=1}^{T} h(Y_{2t}, Y_{1t}, Z_{3t}, \ldots, Z_{Kt}|X_{1t}, X_{2t}, X_{3t}, \ldots, X_{Kt}) \qquad (268)$$
$$= h(Y_1^T, Z_3^T, \ldots, Z_K^T|W_3, \ldots, W_K)$$
$$\quad + h(Y_2^T|Y_1^T, Z_3^T, \ldots, Z_K^T, W_1, W_3, \ldots, W_K)$$
$$\quad - \sum_{t=1}^{T} h(Z_{1t}, Z_{2t}, Z_{3t}, \ldots, Z_{Kt}) + \epsilon_T \qquad (269)$$
$$\leq h(Y_1^T, Z_3^T, \ldots, Z_K^T)$$
$$\quad + h(Y_2^T|Y_1^T, Z_3^T, \ldots, Z_K^T, W_1, W_3, \ldots, W_K)$$
$$\quad - \sum_{t=1}^{T} h(Z_{1t}, Z_{2t}, Z_{3t}, \ldots, Z_{Kt}) + \epsilon_T \qquad (270)$$
$$\leq h(Y_1^T) + h(Z_3^T, \ldots, Z_K^T)$$
$$\quad + h(Y_2^T|Y_1^T, Z_3^T, \ldots, Z_K^T, W_1, W_3, \ldots, W_K)$$
$$\quad - \sum_{t=1}^{T} h(Z_{1t}, Z_{2t}, Z_{3t}, \ldots, Z_{Kt}) + \epsilon_T \qquad (271)$$
$$\leq TA + \sum_{t=1}^{T} h(Z_{3t}, \ldots, Z_{Kt})$$
$$\quad + h(Y_2^T|Y_1^T, Z_3^T, \ldots, Z_K^T, W_1, W_3, \ldots, W_K)$$
$$\quad - \sum_{t=1}^{T} h(Z_{1t}, Z_{2t}, Z_{3t}, \ldots, Z_{Kt}) + \epsilon_T \qquad (272)$$

$$\leq TA + \sum_{t=1}^{T} h(Y_{2t}|X_{1t}, Y_{1t}, X_{3t}) - \sum_{t=1}^{T} h(Z_{2t}) + \epsilon_T \quad (273)$$

$$= TA$$
$$+ \sum_{t=1}^{T} h(\sqrt{\mathsf{SNR}}X_{2t} + Z_{2t}|X_{1t}, \sqrt{\mathsf{INR}}X_{2t} + Z_{1t}, X_{3t})$$
$$- \sum_{t=1}^{T} h(Z_{2t}) + \epsilon_T \quad (274)$$

$$\leq TA + \sum_{t=1}^{T} h(\sqrt{\mathsf{SNR}}X_{2t} + Z_{2t}|\sqrt{\mathsf{INR}}X_{2t} + Z_{1t})$$
$$- \sum_{t=1}^{T} h(Z_{2t}) + \epsilon_T \quad (275)$$

$$\leq TA + T(C - E) + \epsilon_T \quad (276)$$
$$= T(A + C - E) + \epsilon_T, \quad (277)$$

where (264) follows from the independence of the messages, (265) follows from Fano's inequality, (267) follows from the chain rule, and (268) follows from the following argument:

$$X_{1t} \text{ is a function of } (W_1, Y_1^{t-1})$$
$$X_{2t} \text{ is a function of } (W_2, Y_2^{t-1}),$$
$$X_{Kt} \text{ is a function of } (W_K, X_1^{t-1}, Z_K^{t-1}),$$
$$X_{(K-1)t} \text{ is a function of } (W_K, X_K^{t-1}, Z_{K-1}^{t-1}),$$
$$\vdots \quad (278)$$
$$X_{4t} \text{ is a function of } (W_4, X_5^{t-1}, Z_4^{t-1}),$$
$$X_{3t} \text{ is a function of } (W_3, X_4^{t-1}, Z_3^{t-1}).$$

This argument allows us to write $(X_{1t}, X_{2t}, \ldots, X_{Kt})$ in the conditioning in the last term in (268) and then use the memoryless property of the channel to arrive at (269).

The same argument also allows us to write (Y_{1t}, X_{1t}, X_{3t}) in the conditioning of the third term in (273) and subsequently drop all the remaining random variables from the conditioning. We remark here that this argument is similar to Claim 1 used in the proof of Theorem 4 for the linear deterministic model.

Finally, normalizing (277) by T and taking the limit $T \to \infty$, so that $\epsilon_T \to 0$, we have

$$R_1 + R_2 \leq A + C - E. \quad (279)$$

In a similar manner, it can be shown that for any $1 \leq j \leq K$, we have

$$R_j + R_{(j+1)} \leq A + C - E. \quad (280)$$

Adding all such K bounds, we obtain

$$2(R_1 + \ldots + R_K) \leq K(A + C - E), \quad (281)$$

which yields

$$\mathcal{C}_{\mathrm{sum,G}}^{\mathrm{FB}}(K) \leq \frac{K}{2}(A + C - E). \quad (282)$$

Hence, we have proved the analog of the type-I upper bound for the K-user Gaussian CZIC.

D. Marginal Range of Parameters Excluded in Section VI

In the coding scheme and performance analysis presented for various regimes of parameter in Section VI, we have inherently always assumed that SNR and INR (and possibly their ratio) are greater than certain constants, so that the desired rates are non-negative. Although this is a valid assumption for the range of parameters of primary interest, we prove the bounded gap from capacity result for arbitrary parameters for completeness. In this section we focus on the range of parameters excluded from the discussions in Section VI, and show the bounded gap result. In sake of brevity, we present this analysis only for the very weak interference regime $(0 \leq \alpha \leq 1/2)$. The analysis for other ranges of α is very similar, and is omitted.

a) Very Weak Interference $0 \leq \alpha \leq 1/2$: In this regime we have $\mathsf{INR}^2 \leq \mathsf{SNR}$. Recall the rate allocation presented in (74),

$$R_1' = \frac{1}{2}\log^+\left(\frac{\mathsf{INR}+1}{3}\right), \quad k = 1, \ldots, K,$$

$$R_2' = \frac{1}{2}\log^+\left(\frac{\frac{\mathsf{SNR}}{\mathsf{INR}}+1}{2\mathsf{INR}+1}\right), \quad k = 1, \ldots, K, \; j = 1, \ldots, K,$$

$$R_3' = \frac{1}{2}\log^+\left(\frac{\mathsf{INR}+1}{2}\right), \quad k = 1, \ldots, K, \; j = 1, \ldots, K. \quad (283)$$

Consider the following four cases:

I	INR ≥ 2	SNR $\geq 2\mathsf{INR}^2$
II	INR ≥ 2	$\mathsf{INR}^2 \leq \mathsf{SNR} < 2\mathsf{INR}^2$
III	INR < 2	SNR $\geq 2\mathsf{INR}^2$
IV	INR < 2	$\mathsf{INR}^2 \leq \mathsf{SNR} < 2\mathsf{INR}^2$

Case I. The conditions in the first case guarantee that all the rates in (74) are positive, and so the analysis in Section VI-A is valid.

In the following we analyze the remaining three cases which were excluded in Section VI-A.

Case II. In case II, $R_2' = 0$, and hence, the total achievable rate would be

$$R_{\mathrm{sym}} = \frac{R_1 + KR_3}{K}$$
$$= \frac{1}{2}\log(1 + \mathsf{INR}) + \frac{1}{2K}\log(1 + \mathsf{INR}) - \frac{K + \log 3}{2K}. \quad (284)$$

However, note that under the conditions of case II, from (102) we have

$$\mathcal{C}_{\mathrm{sym,G}}^{\mathrm{FB}}(K)$$
$$\leq (B - E) + \frac{E}{K}$$
$$\leq \frac{1}{2}\log\left(1 + 3\mathsf{INR}^2 + 2\mathsf{INR} + 2\sqrt{2\mathsf{INR}^3}\right)$$
$$\quad - \frac{1}{2}\log(1 + \mathsf{INR}) + \frac{1}{2K}\log(1 + \mathsf{INR})$$
$$\leq \frac{1}{2}\log 4\left(1 + \mathsf{INR}\right)^2 - \frac{1}{2}\log(1 + \mathsf{INR}) + \frac{1}{2K}\log(1 + \mathsf{INR})$$
$$\leq \frac{1}{2}\log(1 + \mathsf{INR}) + \frac{1}{2K}\log(1 + \mathsf{INR}) + 1. \quad (285)$$

Therefore, the gap between the upper bound and the achievable rate can be upper bounded as

$$\Delta \leq 1 + \frac{K + \log 3}{2K} < 2. \tag{286}$$

Case III. Next, we should examine the conditions in case III. In this case $R_1' = 0$ and R_3' is upper bounded by a constant. So we have

$$R_{\text{sym}} > R_2 = \frac{1}{2} \log \left(\frac{\frac{\text{SNR}}{\text{INR}} + 1}{2\text{INR} + 1} \right)$$
$$\geq \frac{1}{2} \log(1 + \text{SNR}) - \frac{1}{2} \log 10. \tag{287}$$

Under this condition the upper bound in (102) reduces to

$$\mathcal{C}_{\text{sym,G}}^{\text{FB}}(K)$$
$$\leq (B - E) + \frac{E}{K}$$
$$\leq \frac{1}{2} \log \left(1 + \text{INR} + \frac{\text{SNR} + 2\sqrt{\text{INR} \cdot \text{SNR}}}{1 + \text{INR}} \right) + \frac{1}{6} \log 3$$
$$\leq \frac{1}{2} \log(1 + \text{SNR}) + \frac{1}{2} \log 3 + + \frac{1}{6} \log 3, \tag{288}$$

where we used the facts that $K \geq 3$ and $\text{INR} < 2$ in the second inequality. Therefore,

$$\Delta \leq \frac{1}{2} \log 10 + \frac{2}{3} \log 3 \leq 3. \tag{289}$$

Case IV. Finally, in the last case $R_1' = R_2' = 0$, and R_3' is a constant, and we do not claim any positive rate based on the proposed coding scheme. However, under this condition, the upper bound in (102) would be

$$\mathcal{C}_{\text{sym,G}}^{\text{FB}}(K) \leq (B - E) + \frac{E}{K}$$
$$\leq \frac{1}{2} \log \left(1 + 3\text{INR}^2 + 2\text{INR} + 2\sqrt{2\text{INR}^3} \right)$$
$$\leq \frac{1}{2} \log 25 < \frac{5}{2}, \tag{290}$$

and hence, the gap is bounded by $5/2$.

E. Verification of Rate Bounds in Section VI

In this section we verify the constraints on the rates allocated to sub-messages in the coding scheme used for the Gaussian network. These constraints are due to decodability of the messages at different terminals based on the proposed decoding strategies. We verify the constraints for very weak interference regime, $0 \leq \alpha \leq 1/2$, and omit the details for other cases for sake of brevity. More precisely, will will show that the rates proposed in (74) satisfy the inequalities in (76)–(80), provided that $\text{INR} \geq 2$ and $\text{SNR} \geq 2\text{INR}^2$.

It is trivial to see that R_1 in (74) satisfies (76). Comparing R_2 in (74) to the right-hand side (RHS) of (78) reveals that both expressions have identical denominators, while the nominator of (78) has an extra additive INR term, which makes it larger than the proposed rate. Similarly, R_3 in (74) is always smaller than the RHS of (80).

It remains to verify (77). To this end, it suffices to show that

$$\frac{\text{INR} + 1}{3} \leq \frac{\text{SNR} + \text{INR} + 1}{\frac{\text{SNR}}{\text{INR}} + \text{INR} + 1},$$

or equivalently,

$$3(\text{SNR} + \text{INR} + 1) - \left(\text{SNR} + \text{INR}^2 + 2\text{INR} + \frac{\text{SNR}}{\text{INR}} + 1 \right) \geq 0,$$

which can be further simplified to

$$(\text{SNR} - \text{INR}^2) + \left(\text{SNR} - \frac{\text{SNR}}{\text{INR}} \right) + \text{INR} + 2 \geq 0.$$

Note that the latter is obvious due to the regime assumptions for the values of SNR and INR.

REFERENCES

[1] O. Somekh, B. M. Zaidel, and S. S. (Shitz), "Sum rate characterization of joint multiple cell-site processing," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4473–4497, Dec. 2007.

[2] A. D. Wyner, "Shannon-theoretic approach to a Gaussian cellular multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 40, no. 6, pp. 1713–1727, Nov. 1994.

[3] L. Zhou and W. Yu, "On the capacity of the K-user cyclic Gaussian interference channel," *IEEE Trans. Inf. Theory,* to appear.

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

[5] Y. Liu and E. Erkip, "On the sum capacity of K-user cascade Gaussian Z-interference channel," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, St. Petersburg, Russia, 2011, pp. 1382–1386.

[6] T. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. Inf. Theory*, vol. 27, no. 1, pp. 49–60, January 1981.

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

[8] M. Gastpar and G. Kramer, "On noisy feedback for interference channels," in *Proc. Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, CA, USA, Oct. 29-Nov. 1 2006, pp. 216–220.

[9] D. Tuninetti, "K-user interference channels: General outer bound and sum-capacity for certain Gaussian channels," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, St. Petersburg, Russia, 2011, pp. 1166–1170.

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

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

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

[13] R. Tandon and H. V. Poor, "On the feedback capacity of K-user cyclic interference channel," in *Proc. 49th Annual Allerton Conference on Communications, Control and Computing*, Monticello, IL, September 2011, pp. 855–862.

[14] S. Mohajer, R. Tandon, and H. V. Poor, "On the feedback capacity of the fully connected K-user interference channel," *IEEE Trans. Inf. Theory,* to appear.

[15] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* New York: Wiley, 1991.

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 the areas of 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 Communications Conference.

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 post-doctoral 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.

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 Networking* (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.