Abstract—In this paper, the Nash equilibrium (NE) region of the two-user Gaussian interference channel (IC) with perfect output feedback is characterized to within 2 bits/s/Hz. The relevance of the NE-region is that it provides the set of rate-pairs that are achievable and stable in the IC when both transmitter-receiver pairs autonomously tune their own transmit/receive configurations seeking an optimal individual transmission rate. Therefore, any rate tuple outside the NE region is not stable as there always exists at least one link able to increase its own transmission rate by updating its own transmit/receive configuration. The main conclusions of this paper are: (i) The NE region achieved with feedback is strictly larger than the NE region without feedback. More importantly, all the rate pairs uniquely achievable using feedback are at least weakly Pareto superior to those achievable without feedback. (ii) The use of feedback allows the achievability of all the strictly Pareto optimal rate pairs of the (approximate) capacity region of the Gaussian IC with feedback even when the network is fully decentralized.

Index Terms—Interference channels, feedback communications, Gaussian channels, wireless networks, distributed information systems.

I. INTRODUCTION

In point-to-point communications, perfect output feedback does not increase the capacity either in the discrete or the continuous memoryless channel [1], [2], [3]. At most, feedback increases the capacity by a bounded number of bits per channel use in channels with memory. This is the case for colored additive Gaussian noise [4], [5] and stationary first-order moving average Gaussian noise [6]. The same can be said for some multiuser channels in which the capacity region is broadened only by a limited number of bits per channel use. This is the case for the memoryless Gaussian multiple access channel (MAC) [7], [8], [9], [10]. In the discrete memoryless broadcast channel (BC), there exists evidence that feedback increases the capacity region [11], [12], [13]. However, in particular cases such as the physically degraded BC, the opposite has been formally proven [14]. Feedback substantially enlarges the capacity region of the 2-user memoryless Gaussian interference channel (IC) [15]. The same effect is observed in some special cases with a larger number of users, e.g., in the symmetric $K$-user cyclic $Z$-interference channel [16] and the fully connected-user IC [17]. The two-user linear deterministic IC with partial feedback has been considered in [18]. In the particular case of the 2-user memoryless Gaussian IC, in the weak interference regime and the very strong interference regime, the gain provided by feedback is arbitrarily large when the signal to noise ratio (SNR) and the interference to noise ratio (INR) grow to infinity at a constant ratio. One of the reasons why feedback provides such a surprising benefit relies on the fact that it creates an alternative path to the existing point-to-point paths. For instance, in the two-user IC, feedback creates a path from transmitter 1 (resp. transmitter 2) to receiver 1 (resp. receiver 2) in which signals that are received at receiver 2 (resp. receiver 1) are fed back to transmitter 2 (resp. transmitter 1) which decodes the messages and re-transmits them to receiver 1 (resp. receiver 2). This implies a type of cooperation in which transmitters engage each other to transmit each other’s messages.

In decentralized multiuser channels, the benefits of feedback are less well understood and the existing results from the centralized perspective do not apply immediately. In a decentralized network, each transmitter-receiver link acts autonomously and tunes its individual transmit/receive configuration aiming to optimize a given performance metric. Therefore, in decentralized networks, a competitive scenario arises in which the individual improvement of one link often implies the detriment of the other links due to the mutual interference. From this point of view, in decentralized networks, the notion of capacity region is shifted to the notion of equilibrium region. Such a region varies depending on the associated notion of equilibrium, e.g., Nash Equilibrium [19], correlated equilibrium [20], satisfaction equilibrium [21], etc. In particular, when each individual link aims to selfishly optimize its individual transmission rate by tuning its transmit/receive configuration, the equilibrium must be understood in terms of the Nash equilibrium (NE). Once an NE is achieved, none of the links has a particular interest in unilaterally deviating from the actual transmit/receive configuration. Essentially, any deviation from an NE implies a loss in the individual rate of the deviating transmitter. Therefore, any rate tuple outside the Nash-region is not stable as there always exists at least one link that is able to increase its own transmission rate by updating its own transmit/receive configuration.

An approximate characterization of the NE region of the decentralized Gaussian IC is presented in [22]. This characterization implies two important points. First, in all the interference regimes, the NE region is non-empty, which verifies some of the existing results in [23], [24] and [25]. Second, the
individual rates achievable at an NE are both lower and upper bounded. The lower bound corresponds to the rate achieved by treating interference as noise, whereas the upper bound requires partial decoding of the interference. Interestingly, it is shown that in the strong and very strong interference regimes, the NE region equals the capacity region. Conversely, in all the other regimes, the NE region is a subregion of the capacity region and often, it does not contain all the strictly Pareto optimal rate pairs, e.g., the rate pairs on the boundary of the sum-capacity.

From the reasoning above, it might appear natural that feedback does not bring any benefit to the NE-region. This is mainly because feedback can be seen as a cooperation strategy in most of the centralized multiuser channels. Indeed, in the centralized Gaussian IC (G-IC) [15], feedback does not benefit the link that implements it but rather the other links in terms of transmission rate. Thus, from a selfish point of view, no user would be interested in using feedback. However, this paper shows the opposite. Even in the strictly competitive scenario in which both links are selfish, the use of feedback appears to be at least, non-disadvantageous to the user implementing it and thus, transmitters might opt to use it in some cases. This observation leads to two of the most important conclusions of this work: (i) The NE region with feedback is strictly larger than the NE region with no feedback. More importantly, all the rate pairs achieved only by using feedback are either strictly or weakly Pareto superior to the rate pairs achieved without feedback; and (ii) The use of feedback allows the achievability of all Pareto optimal rate pairs of the approximate capacity region of the IC with feedback.

The remainder of this paper unfolds as follows. In Sec. II, the decentralized IC with feedback is formally introduced and its equivalent game theoretic model is presented. In Sec. III, the Nash region of the linear deterministic IC (LD-IC) with feedback (LD-IC-FB) is fully characterized. In Sec. IV, using the intuition obtained from the LD-IC-FB, the NE region of the G-IC with feedback is approximated to within two bits. This approximation inherits the two-bit precision of the approximation of the capacity region of this channel [15]. Finally, in Sec. V, using metrics such as the price of anarchy and the price of stability, it is shown that under proper equilibrium selection, there is no loss in the sum rate due to the anarchical behavior of both links in the IC with feedback. Otherwise, if such a loss exists, it monotonically increases with the INR and monotonically decreases with the SNR of the individual links. This work is concluded by Sec. VI.

II. Problem Formulation

Consider the fully decentralized two-user interference channel with perfect output feedback in Fig. 1. Transmitter $i$, with $i \in \{1, 2\}$, communicates with receiver $i$ during $T$ consecutive blocks subject to the interference produced by transmitter $j \in \{1, 2\} \setminus \{i\}$. During block $t$, transmitter $i$ sends $M_i$ information bits $b_{i1}^{(t)}, \ldots, b_{iM_i}^{(t)}$ by transmitting the codeword $x_{i1}^{(t)} = (x_{i11}^{(t)}, \ldots, x_{i1N_i})^T \in \mathcal{X}_i$, where $\mathcal{X}_i$ denotes the codebook of transmitter $i$. All information bits are independent and identically distributed (i.i.d.) following a uniform probability distribution. The input to receiver $i$ is denoted by $y_{i1}^{(t)} = (y_{i11}^{(t)}, \ldots, y_{i1N_i})^T$. The signals $y_{i1}^{(t)}$ and $y_{i2}^{(t)}$ during block $t$ can be written as follows:

\[
y_{i1}^{(t)} = h_{11}x_{i1}^{(t)} + h_{12}x_{i2}^{(t)} + z_{i1}^{(t)} \quad \text{and} \quad (1)
y_{i2}^{(t)} = h_{21}x_{i1}^{(t)} + h_{22}x_{i2}^{(t)} + z_{i2}^{(t)}, \quad (2)
\]

where $z_{i1}^{(t)} = (z_{i11}^{(t)}, \ldots, z_{i1N_i})^T$ represents the noise observed at each of the $N_i$ channel uses. For all $l \in \{1, \ldots, N_i\}$, the noise terms $z_{i1l}^{(t)}$ are independent circularly symmetric complex Gaussian (CSCG) random variables with zero means and unit variances. The channel coefficient from transmitter $j$ to receiver $i$ is denoted by $h_{ij}$ and it is a complex number with norm $|h_{ij}|^2 = g_{ij}$. The elements of the input vector $x_{i1}^{(t)} = (x_{i11}^{(t)}, \ldots, x_{i1N_i})^T$ are complex CSCG random variables with zero means and subject to the variance constraint

\[
\frac{1}{N_i} \sum_{l=1}^{N_i} |x_{i1l}^{(t)}|^2 \leq P_i, \quad (3)
\]

with $P_i = 1$ the average transmit power of transmitter $i$.

A perfect feedback link from receiver $i$ to transmitter $i$ allows at the end of each block $t$, the observation of the sequence $y_{i1}^{(t)}$ at transmitter $i$. Therefore, the encoder of transmitter $i$, during block $t$, can be modeled as a deterministic mapping $f_i^{(t)}$ such that $x_{i1}^{(t)} = f_i^{(t)}(y_{i1}^{(t)}, y_{i1}^{(t-1)}, \ldots, y_{i1}^{(0)}) \in \mathcal{X}_i$, where $k \in \{1, \ldots, 2^{M_i}\}$ is the index of the message to be transmitted. At the end of the complete transmission, after block $T$, receiver $i$ uses the sequences $y_{1}^{(t)}, \ldots, y_{T}^{(t)}$ to generate estimates $\hat{b}_{i1}^{(t)}, \forall (t, i) \in \{1, \ldots, M_i\} \times \{1, \ldots, T\}$. The average bit error probability of transmitter $i$ during block $t$, denoted by $p_{i1}^{(t)}$, is calculated as follows:

\[
p_{i1}^{(t)} = \frac{1}{M_i} \sum_{l=1}^{M_i} \mathbf{1}\{\hat{b}_{i1l}^{(t)} \neq b_{i1l}^{(t)}\}. \quad (4)
\]

The Gaussian IC with feedback in Fig. 1 can be fully described by four parameters: (a) the signal to noise ratios $SNR_i = |h_{ii}|^2$; and (b) the interference to noise ratios $INR_{ij} = |h_{ij}|^2$. The rate pair $(R_1, R_2) \in \mathbb{R}_+^2$ is achievable.
if there exists at least one pair of codebooks \( \mathcal{X}_1 \) and \( \mathcal{X}_2 \) with codewords of length \( N_1 \) and \( N_2 \), respectively, with the corresponding encoding functions \( f_1 \) and \( f_2 \) such that the average bit error probability can be made arbitrarily small by letting the block lengths \( N_1 \) and \( N_2 \) grow to infinity.

The aim of transmitter \( i \) is to autonomously choose its transmit configuration \( s_i \) in order to maximize its achievable rate \( R_i \). More specifically, the transmit configuration \( s_i \) can be described in terms of the number of information bits per block \( M_i \), the block length \( N_i \), the codebook \( X_i \), the encoding functions \( f_i \), etc. Note that the rate achieved by receiver \( i \) depends on both configurations \( s_1 \) and \( s_2 \) due to the mutual interference naturally arising in the interference channel. This reveals the competitive interaction between both links in the decentralized interference channel. The following section models this interaction using tools from game theory.

### A. Game Formulation

The competitive interaction of the two transmitters in the interference channel, as described in the previous section, can be modeled by the following game in normal-form:

\[
\mathcal{G} = (\mathcal{K}, \{\mathcal{A}_k\}_{k \in \mathcal{K}}, \{u_k\}_{k \in \mathcal{K}}).
\]  

(5)

The set \( \mathcal{K} = \{1, 2\} \) is the set of players, that is, the set of transmitter-receiver pairs. The sets \( \mathcal{A}_1 \) and \( \mathcal{A}_2 \) are the sets of actions of player 1 and 2, respectively. An action of a player \( i \), which is denoted by \( s_i \in \mathcal{A}_i \), is basically its transmit/receive configuration as described above. The utility function of player \( i \) is \( u_i : \mathcal{A}_1 \times \mathcal{A}_2 \rightarrow \mathbb{R}_+ \) and it is defined as the achieved rate of transmitter \( i \),

\[
u_i(s_1, s_2) = \begin{cases} R_i(s_1, s_2), & \text{if } \forall t \in \{1, \ldots, T\}, p_i^{(t)} < \epsilon, \\ 0, & \text{otherwise}, \end{cases}
\]

(6)

where \( \epsilon > 0 \) is an arbitrarily small number and \( R_i(s_1, s_2) \) denotes a transmission rate achievable with the configurations \( s_1 \) and \( s_2 \) such that \( p_i^{(t)} < \epsilon \). Often, the rate \( R_i(s_1, s_2) \) is written as \( R_i \) for the sake of simplicity. However, every non-negative rate is associated with a particular pair of transmit configurations \( s_1 \) and \( s_2 \). It is worth noting that there might exist several transmit configurations that achieve the same rate pair \( (R_1, R_2) \).

A particular class of action profiles \( s = (s_1, s_2) \in \mathcal{A}_1 \times \mathcal{A}_2 \) which are particularly important in the analysis of this game are referred to as \( \eta \)-Nash equilibria (\( \eta \)-NE) and satisfy the following conditions.

**Definition 1 (\( \eta \)-Nash equilibrium):** In the game \( \mathcal{G} = (\mathcal{K}, \{\mathcal{A}_k\}_{k \in \mathcal{K}}, \{u_k\}_{k \in \mathcal{K}}) \), an action profile \( (s_1^*, s_2^*) \) is an \( \eta \)-Nash equilibrium if \( \forall i \in \mathcal{K} \) and \( \forall s_i \in \mathcal{A}_i \),

\[
u_i(s_1^*, s_2^*) \leq \nu_i(s_1^*, s_2^*) + \eta.
\]

(7)

From Def. 1, it becomes clear that if \( (s_1^*, s_2^*) \) is an \( \eta \)-Nash equilibrium, then none of the transmitters can increase its own transmission rate more than \( \eta \) bits per block by changing its own transmit configuration and keeping the average bit error probability arbitrarily close to zero. Thus, at a given \( \eta \)-NE, every transmitter achieves a utility (transmission rate) that is \( \eta \)-close to its maximum achievable rate given the transmit configuration of the other transmitter. Note that if \( \eta = 0 \), then the classical definition of Nash equilibrium is obtained [19]. The relevance of the notion of equilibrium is that at any NE, every transmitter configuration is optimal with respect to the configuration of the other transmitters. The following investigates the set of rate pairs that can be achieved at an NE. This set of rate pairs is known as the Nash region.

**Definition 2 (Nash Region):** An achievable rate pair \( (R_1, R_2) \) is said to be in the Nash region of the game \( \mathcal{G} = (\mathcal{K}, \{\mathcal{A}_k\}_{k \in \mathcal{K}}, \{u_k\}_{k \in \mathcal{K}}) \) if there exists an action profile \( (s_1^*, s_2^*) \) that is an \( \eta \)-Nash equilibrium for an arbitrarily small \( \eta \) and the following hold:

\[
u_1(s_1^*, s_2^*) = R_1 \quad \text{and} \quad \nu_2(s_1^*, s_2^*) = R_2.
\]

(8)

The following section studies the NE region of the game \( \mathcal{G} \) in (5) using a deterministic approximation in order to obtain some initial insight.

### III. Linear Deterministic Interference Channel with Feedback

The linear deterministic approximation of the G-IC was introduced in [26]. In general, the linear deterministic model deemphasizes the effect of background noise and focuses on the signal interactions. Furthermore, as shown in [27], it provides valuable insights that can be used to apply to the corresponding Gaussian model.

The linear deterministic IC is described by four parameters: \( n_{11}, n_{22}, n_{12}, \) and \( n_{21} \), where \( n_{ii} \) captures the signal strength from transmitter \( i \) to receiver \( i \), and \( n_{ij} \) captures the interference strength from transmitter \( j \) to receiver \( i \). The input-output relationship is given as follows:

\[
y_1^{(t)} = S^{q-n_{11}} x_1^{(t)} + S^{q-n_{12}} x_2^{(t)}, \quad \text{and} \quad y_2^{(t)} = S^{q-n_{21}} x_1^{(t)} + S^{q-n_{22}} x_2^{(t)}.
\]

(9)

where \( x_1^{(t)}, y_1^{(t)} \in \{0, 1\}^q \), with \( q = \max_{(i,j) \in \{1,2\}^2} n_{ij} \). Additions and multiplications are over a binary field, and \( S \) is a \( q \times q \) lower shift matrix of the form

\[
S = \begin{bmatrix}
0 & 0 & 0 & \cdots & 0 \\
1 & 0 & 0 & \cdots & 0 \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
0 & \cdots & 0 & 1 & 0 \\
1 & 0 & 0 & \cdots & 0
\end{bmatrix}.
\]

The parameter \( n_{ii} \) corresponds to \( \log_2(\text{SNR}_i) \) and \( n_{ji} \) corresponds to \( \log_2(\text{INR}_{ji}) \), where \( \text{SNR}_i \) and \( \text{INR}_{ji} \) in the corresponding Gaussian interference channel described above. For a detailed discussion about the connections between the LD-IC and the G-IC, the reader is referred to [28].

### A. Preliminaries

In the following, some of the existing results used to fully characterize the Nash region of the LD-IC-FB are briefly presented.
1) Capacity of the LD-IC without Feedback: The capacity region of the two-user LD-IC without feedback is denoted by $\mathcal{C}_{\text{LDIC}}$ and it is fully characterized by Lemma 4 in [28].

Lemma 1 (Lemma 4 in [28]): The capacity region $\mathcal{C}_{\text{LDIC}}$ of the LD-IC without feedback corresponds to the set of non-negative rate pairs $(R_1, R_2)$ satisfying

$$R_i \leq n_{ii}, \text{ with } i \in \{1, 2\},$$

$$R_1 + R_2 \leq (n_{11} - n_{12})^+ + \max(n_{22}, n_{12}),$$

$$R_1 + R_2 \leq (n_{22} - n_{21})^+ + \max(n_{11}, n_{21}),$$

$$R_1 + R_2 \leq \max(n_{21}, (n_{11} - n_{12})^+) + \max(n_{12}, (n_{22} - n_{21})^+),$$

$$2R_1 + R_2 \leq \max(n_{11}, n_{21}) + (n_{11} - n_{12})^+) + \max(n_{12}, (n_{22} - n_{21})^+),$$

$$R_1 + 2R_2 \leq \max(n_{22}, n_{12}) + (n_{22} - n_{21})^+ + \max(n_{21}, (n_{11} - n_{12})^+).$$

(10)

Note that the capacity region shown in Lemma 1 is a particular case of the capacity region presented in [29] that applies to a larger class of deterministic interference channels.

2) Nash Region of the LD-IC without Feedback: The Nash region of the linear deterministic interference channel without feedback is denoted by $\mathcal{N}_{\text{LDIC}}$ and it is fully characterized in [22] in terms of the set

$$\mathcal{B}_{\text{LDIC}} = \{(R_1, R_2) \in \mathbb{R}^2 : L_i \leq R_i \leq U_i, \forall i \in \{1, 2\}\},$$

(11)

where, $\forall i \in \{1, 2\},$

$$L_i = (n_{ii} - n_{ij})^+,$$

$$U_i = \left\{ \begin{array}{ll}
             n_{ii} & \text{if } n_{ij} \leq n_{ii},
            \min\left((n_{ij} - L_j)^+, n_{ii}\right) & \text{if } n_{ij} > n_{ii},
         \end{array} \right. $$

(12)

(13)

Therefore, the following lemma holds.

Lemma 2 (Theorem 1 in [22]): The Nash region of the two-user LD-IC without feedback $\mathcal{N}_{\text{LDIC}}$ is

$$\mathcal{N}_{\text{LDIC}} = \mathcal{B}_{\text{LDIC}} \cap \mathcal{C}_{\text{LDIC}}.$$  

(14)

3) Capacity of the LD-IC with Feedback: The capacity region of the linear deterministic interference channel with feedback is denoted by $\mathcal{C}_{\text{LDIC/FB}}$ and it is fully characterized by Corollary 1 in [15].

Lemma 3 (Corollary 1 in [15]): The capacity region $\mathcal{C}_{\text{LDIC/FB}}$ of the two-user LD-IC with feedback corresponds to the set of non-negative rate pairs $(R_1, R_2)$ satisfying

$$R_1 \leq \min\left(\max(n_{11}, n_{12}), \max(n_{11}, n_{21})\right),$$

$$R_2 \leq \min\left(\max(n_{22}, n_{21}), \max(n_{22}, n_{12})\right),$$

$$R_1 + R_2 \leq \min\left(\max(n_{22}, n_{21}) + (n_{11} - n_{21})^+,\right.$$

$$\max(n_{11}, n_{12}) + (n_{22} - n_{12})^+) \right\}. $$

(15)

B. Main Result

The Nash region of the LD-IC with feedback is denoted by $\mathcal{N}_{\text{LDIC/FB}}$. In order to define the set $\mathcal{N}_{\text{LDIC/FB}}$, first consider the following set:

$$\mathcal{B}_{\text{LDIC/FB}} = \{(R_1, R_2) \in \mathbb{R}^2 : R_i \geq L_i, \forall i \in \{1, 2\}\},$$

(16)

where $L_i$ is defined in (12). The main result for the linear deterministic IC with feedback is stated in terms of the set $\mathcal{B}_{\text{LDIC/FB}}$.

Theorem 1: For a two-user linear deterministic IC with feedback,

$$\mathcal{N}_{\text{LDIC/FB}} = \mathcal{B}_{\text{LDIC/FB}} \cap \mathcal{C}_{\text{LDIC/FB}}.$$  

(17)

The proof of Theorem 1 is presented in Sec. III-D. However, before describing the details of the proof, this result is illustrated using a simple example. Consider a symmetric linear deterministic IC with feedback in which, $n = n_{11} = n_{22}$ and $m = n_{12} = n_{21}$, with normalized cross gain $\alpha = \frac{m}{n}$. The regions $\mathcal{C}_{\text{LDIC}}, \mathcal{N}_{\text{LDIC}}, \mathcal{C}_{\text{LDIC/FB}},$ and $\mathcal{N}_{\text{LDIC/FB}}$ are plotted in Fig. 2 for different interference regimes, i.e., very weak interference $(0 \leq \alpha \leq \frac{1}{2})$, weak interference $(\frac{1}{2} < \alpha \leq \frac{3}{4})$, moderate interference $(\frac{3}{4} < \alpha \leq 1)$, strong interference $(1 < \alpha \leq 2)$ and very strong interference $(\alpha \geq 2)$, respectively. Note that all the rate pairs $(R_1, R_2)$, such that $R_1 > L_1$ and $R_2 > L_2$, are achievable at an NE either with feedback or without feedback. However, in all the interference regimes, the use of feedback increases the number of rate pairs that are achievable at the NE with respect to the case without feedback. More importantly, the new achievable rate pairs are either Pareto optimal or weakly Pareto superior to the rate pairs achievable without feedback at the NE. More precisely, $\forall (R_1^*, R_2^*) \in \mathcal{N}_{\text{LDIC}}$ there always exists a rate pair $(R_1^*, R_2^*) \in \mathcal{N}_{\text{LDIC/FB}}$ such that $R_i^* \geq R_i^*$, $\forall i \in \{1, 2\}$. This observation underlines the main benefits of using feedback when all users act selfishly and can be summarized by the following inclusion that holds with strict inequality for all interference regimes:

$$\mathcal{N}_{\text{LDIC}} \subset \mathcal{N}_{\text{LDIC/FB}}.$$  

(18)

The previous observation also shows that in the interference channel, increasing the space of actions by letting users choose whether to use feedback or not does not decrease either their individual rates or the aggregated achievable transmission rate at the equilibrium. This statement might appear obvious, however, it has been shown that increasing the set of actions of some players might reduce their individual performance or even the global sum-performance. In particular, this effect has been observed in the parallel IC and parallel MAC when transmitters are allowed to use a larger number of dimensions to transmit [30], [25]. This effect is often understood as a type of Braess’s paradox [31].

The main advantage of using feedback when users act selfishly is that it allows the achievability of all sum-rate maximizing pairs of $\mathcal{C}_{\text{LDIC/FB}}$ at the equilibrium in all the interference regimes. It is worth noting that in the case $\alpha > 1$ with feedback, $\mathcal{C}_{\text{LDIC/FB}} \subseteq \mathcal{B}_{\text{LDIC/FB}}$ and thus, all the achievable rates of the IC with feedback are also achievable.
Fig. 2. Illustration of $C_{\text{LDIC}}$ (red dotted line), $N_{\text{LDIC}}$ (solid blue line), $C_{\text{LDIC}}/\text{FB}$ (green dotted line), and $N_{\text{LDIC}}/\text{FB}$ (magenta solid line) in all interference regimes.

at the NE, i.e., $N_{\text{LDIC}}/\text{FB} = C_{\text{LDIC}}/\text{FB}$. This contrasts with the case without feedback, in which only a subset of the sum-rate maximizing pairs in $C_{\text{LDIC}}$ are achievable at the equilibrium in most of the interference regimes. For instance, when $0 \leq \alpha \leq \frac{1}{2}$ and $\frac{1}{2} < \alpha \leq \frac{2}{3}$ (see Fig. 2) only one pair $(R_1, R_2)$ of the infinitely many Pareto optimal pairs is achieved at the equilibrium.

The following presents a few examples to provide some intuition into the impact of feedback on the Nash equilibrium region of the IC.

C. Examples

Consider the scenario of very weak interference, for instance, let $\alpha = \frac{1}{3}$, with $m = 2$ and $n = 6$. From Theorem 1, it follows that the Nash region is $N_{\text{LDIC}}/\text{FB} = \{(R_1, R_2) \in \mathbb{R}^2 : \forall i R_i \geq 4, R_1 + R_2 \leq 10\}$. In Fig. 2, the region $N_{\text{LDIC}}/\text{FB}$ corresponds to the convex hull of the points $(4, 4)$, $(6, 4)$ and $(4, 6)$. The pair $(4, 4)$ is achieved without feedback, the rate pair $(6, 4)$ or $(4, 6)$ is achieved when one of the transmitters uses feedback and the rate pair $(5, 5)$ is achieved when both players use feedback.

1) Achievability of $(4, 4)$: The rate pair $(4, 4)$ is achievable when none of the transmitters uses feedback (See Fig. 2). Note that when one of the transmitters sends only private messages to its corresponding receiver, the highest achievable rate of the other transmitter is achieved by sending only private messages to its corresponding receiver. Here, any attempt of a transmitter to increase its rate by using its $m$ lowest levels would bound its probability of error away from zero since those levels are subject to the interference of the $m$ highest levels of the other transmitter. That is, when both players send new bits at every block using their $n - m$ highest levels (see Fig. 3), this configuration is an NE independently of whether one or both transmitters implement feedback. In this case, the bits obtained at the transmitter via feedback are not useful to improve its own coding/decoding scheme.

Remark 1: The choice of using all the highest levels to transmit new bits at every block $t$, or equivalently not using feedback, can be associated with a greedy behavior. However, when both players adopt this coding scheme, the outcome is a Nash equilibrium. Nonetheless, it is the worst Nash equilibrium in terms of their individual rates.
As shown in the next example, if one of the transmitters uses feedback, it obtains the same rate as in the previous case while allowing the other transmitter to achieve a higher rate.

2) Achievability of (6, 4) and (4, 6): The rate pairs (6, 4) and (4, 6) are achievable at an NE when one of the transmitters uses feedback. Consider for instance that transmitter 1 uses all its $n$ levels to transmit new information at each block $t$, that is, it does not use feedback. Under this condition, the maximum rate achievable by transmitter 2 is 4 bits per block (see the capacity region $C_{\text{LDIC/FB}}$). Note that the rate of 4 bits per block can be achieved by transmitter 2 by simply using its $n - m$ highest levels or using its $n - m$ lowest levels with feedback to resolve the interference produced by the top levels of transmitter 1. In both cases, any attempt by transmitter 2 to use its $m$ highest levels to send new bits instead of those obtained via feedback would constrain transmitter 1 to achieve a rate of 6 bits per block. This is basically because at least $m$ bits would not be decoded reliably at receiver 1. Thus, the rate pair (6, 4) is an NE if there exists a coding scheme for transmitter 2 that does not use its $m$ highest levels and achieves at least a rate of 4 bits per block. This coding scheme is presented in Fig. 4.

Note that transmitter 2 uses its $n - m$ lowest levels to transmit new information bits at each block, and thus, at least $m$ of these bits are subject to the interference of the $m$ highest levels of transmitter 1. However, by using feedback, transmitter 2 can transmit over its $m$ highest levels during block $t$, the $m$ bits received by its $m$ lowest levels in block $t - 1$. At block $t$, these retransmitted bits ($a_1$ and $a_2$ in Fig. 4) are received in the $m$ lowest levels of receiver 1, however, they do not represent any interference as they have been decoded in the previous block $t - 1$ and thus, they can be subtracted. At receiver 2, these bits are received interference-free at the $m$ highest levels during block $t$ and can be used to decode the $m$ bits received in the $m$ lowest levels in the previous block $t - 1$. Thus, this coding scheme achieves a rate of 4 bits per block without using the $m$ highest levels of transmitter 2. This shows that the rate pair (6, 4) can be achieved at an NE. By exchanging the identity of the players in the analysis above, the achievability of the rate pair (4, 6) at an NE is verified.

Remark 2: When player $i$ implements a strategy that does not use $\ell \leq m$ of its highest levels for transmitting new bits but rather those obtained via feedback, transmitter $i$ grants $\ell$ additional bits per block to transmitter $j$ with respect to the case in which feedback is not used. That is, transmitter $j$ achieves a rate $R_j = (n - m)^+ + \ell$ and transmitter $i$ achieves $R_i = (n - \ell)^+ \geq L_i$. Note that as long as $0 \leq \ell \leq m$, both transmitters achieve a rate that is at least equal to $L_i$, which is the minimum rate at an NE. Moreover, it is easy to verify that neither of the transmitters can increase its transmission rate by unilateral deviation and thus, these rate pairs are achievable at an NE.

The case in which both transmitters implement feedback is presented in the following example.

3) Achievability of (5, 5): The rate pair (5, 5) is achieved at an NE when both transmitters use feedback. In this case, both transmitters use their highest levels to transmit the bits obtained via feedback instead of new bits at every channel use. The coding scheme that achieves this rate pair is presented in Fig. 5. Therein, it can be verified that any attempt by either of the transmitters to increase its individual rate by sending new bits at each block (instead of those obtained via feedback) raises its own probability of error. This verifies that the rate pair (5, 5) is achievable at an NE.

Remark 3: The underlying conclusion of this example is that a transmitter using feedback does not negatively affect its own transmission rate but significantly benefits the other transmitter at the equilibrium. Thus, the use of feedback can be easily associated with an altruistic behavior.

D. Proofs

To prove Theorem 1, the first step is to show that a rate pair $(R_1, R_2)$, with $R_i \leq (n_{ii} - n_{ij})^+$ for at least one $i \in \{1, 2\}$ is not an $\eta$-equilibrium for an arbitrarily small $\eta$. That is,

$$\mathcal{N}_{\text{LDIC/FB}} \subseteq \mathcal{C}_{\text{LDIC/FB}} \cap \mathcal{B}_{\text{LDIC/FB}}.$$  \hspace{1cm} (19)

The second step is to show that any point in $\mathcal{C}_{\text{LDIC/FB}} \cap \mathcal{B}_{\text{LDIC/FB}}$ is an $\eta$-equilibrium $\forall \eta > 0$. That is,

$$\mathcal{N}_{\text{LDIC/FB}} \supseteq \mathcal{C}_{\text{LDIC/FB}} \cap \mathcal{B}_{\text{LDIC/FB}}.$$  \hspace{1cm} (20)

which proves Theorem 1.
and thus, the most significant levels and possible combinations of the IC without feedback. Let into their common messages as done in [22] for the case of the consists of allowing users to introduce some random symbols of the feedback coding scheme presented in [15]. The novelty second part of the proof of Theorem 1, consider a modification follows that the action profile improves rate s (equilibrium, with symmetric LD-IC with feedback, with Fig. 5. Coding scheme for achieving the rate pair 1

1) Non-equilibrium Rate Pairs: The first part of the proof of Theorem 1 is completed by the following lemma.

**Lemma 4:** A rate pair \((R_1, R_2) \in \mathcal{C}_{LDIC/FB}\), with either \(R_1 < (n_{11} - n_{12})^+\) or \(R_2 < (n_{22} - n_{21})^+\) is not an \(\eta\)-equilibrium, with \(\eta \geq 0\) and arbitrarily small.  

**Proof:** Let \((s_1, s_2)\) be an action profile such that users 1 and 2 achieve the rate pair \(R_1(s_1, s_2)\) and \(R_2(s_1, s_2)\), respectively. Assume, without loss of generality, that \(R_1(s_1, s_2) < (n_{11} - n_{12})^+\). Then, note that there exists at least one action \(s_1^1\) such that transmitter 1 uses its top levels, which are interference free, and thus it is always able to achieve a rate \(R_1(s_1^1, s_2) \geq (n_{11} - n_{12})^+\). Note also that the utility improvement \(R_1(s_1^1, s_2) - R_1(s_1, s_2) > 0\) is always possible independently of the current action \(s_2\) of user 2. Thus, it follows that the action profile \((s_1, s_2)\) is not an \(\eta\)-equilibrium, for an arbitrarily small \(\eta\). This completes the proof.

2) Achievable Equilibrium Rate Pairs: To continue with the second part of the proof of Theorem 1, consider a modification of the feedback coding scheme presented in [15]. The novelty consists of allowing users to introduce some random symbols into their common messages as done in [22] for the case of the IC without feedback. Let \(X_i\) be the input alphabet of link \(i\), that is, the set of symbols obtained from the different combinations of its \(n_{ii}\) levels. The set \(X_i\) is constructed as the Cartesian product of the sets \(X_ip\) and \(X_ic\), where \(X_ip\) contains all the possible combinations of the \((n_{ii} - n_{ij})^+\) least significant levels and \(X_ic\) contains all the possible combinations of the \(n_{ij}\) most significant levels. Thus, at each block \(t\) transmitter \(i\) sends the codeword \(x_i(t) = (x_{ic}(t), x_{ip}(t)) \in X_ic \times X_ip\). The least significant \((n_{ii} - n_{ij})^+\) levels \(x_{ip}(t)\) are seen only by receiver \(i\), and thus \(x_{ip}(t)\) plays the role of a private message. Conversely, the most significant \(n_{ij}\) levels \(x_{ic}(t)\) are seen by both receivers and thus, \(x_{ic}(t)\) plays the role of a common message. It is worth noting that thanks to the action of feedback, transmitter \(i\) can obtain at the end of each block \(t\) the common message \(x_{ic}(t)\) sent by transmitter \(j\). This is basically because \(x_{ic}(t)\) is observed in the \(n_{ij}\) least significant levels of receiver \(i\). Note that even when the min \((n_{ij}, n_{ii})\) least significant levels of \(x_{ic}(t)\) are mixed with the min \((n_{ij}, n_{ii})\) least significant levels of \(x_{ip}(t)\), the latter are perfectly known at transmitter \(i\) and thus, \(x_{ic}(t)\) can be obtained interference-free. At the beginning of block \(t\), transmitter \(i\) generates the symbol \(x_{ic}(t)\) using the common message index \(m_{ic}(t) \in \{1, \ldots, 2^nR_{ic}\}\) and a randomly generated index \(m_{ip}(t) \in \{1, \ldots, 2^nR_{ip}\}\). Transmitter \(i\) uses the mapping \(f_{ic}: \{1, \ldots, 2^nR_{ic}\} \times \{1, \ldots, 2^nR_{ip}\} \rightarrow X_{ic}\) and thus \(f_{ic}(m_{ic}(t), m_{ip}(t)) = x_{ic}(t)\). The private part \(x_{ip}(t)\) is generated using the private message index \(m_{ip}(t) \in \{1, \ldots, 2^nR_{ip}\}\); the common codeword \(x_{ic}(t) \in X_{ic}\); and the two previous common codewords \(x_{ic}(t-1) \in X_{ic}\) and \(x_{ic}(t-1) \in X_{ic}\), where the latter is obtained by feedback. Hence, transmitter \(i\) uses the mapping \(f_{ip}: X_{ic} \times X_{jc} \times X_{ic} \times \{1, \ldots, 2^nR_{ip}\} \rightarrow X_{ip}\) and thus \(f_{ip}(x_{ic}(t-1), x_{jc}(t-1), x_{ic}(t), m_{ip}(t)) = x_{ip}(t)\). Thus, the pair of codewords \((x_{ic}(t-1), x_{jc}(t-1))\) determines a center of a cloud of codewords, the codeword \(x_{ic}(t)\) determines a smaller cloud of codewords inside the previous cloud and finally, the private message index \(m_{ip}(t)\) determines a private codeword inside the smaller cloud.

The random indices \(m_{ic}(t)\) are assumed to be known at the receiver \(i\) and thus, they do not convey any new information. That is, if transmitter \(i\) achieves a rate tuple \((R_{ip}, R_{ic}, R_{ir})\), its actual rate is \(R_i = R_{ic} + R_{ip}\).

This coding scheme is referred to as a randomized Han-Kobayashi coding scheme with feedback and it is thoroughly described in Appendix A. The achievable region of this coding scheme is presented by the following lemma.

**Lemma 5:** The achievable region of the randomized Han-Kobayashi coding scheme with feedback in the linear deterministic IC is the set of tuples \((R_{1c}, R_{1r}, R_{1p}, R_{2c}, R_{2r}, R_{2p})\) that satisfy the following conditions:

\[
\begin{align*}
R_{1c} + R_{1r} & \leq n_{21} \\
R_{2p} & \leq (n_{22} - n_{12})^+ \\
R_{1c} + R_{1p} + R_{2c} + R_{2r} & \leq \max(n_{11}, n_{12}) \\
R_{2c} + R_{2r} & \leq n_{12} \\
R_{1p} & \leq (n_{11} - n_{21})^+ \\
R_{2c} + R_{2p} + R_{1c} + R_{1r} & \leq \max(n_{22}, n_{21}) \\
\end{align*}
\]

The proof of Lemma 5 is presented in Appendix A.

The set of inequalities in (21) can be written in terms of the transmission rates \(R_1 = R_{1p} + R_{1c}\) and \(R_2 = R_{2p} + R_{2c}\), which yields the following conditions:

\[
\begin{align*}
R_{1r} & \leq n_{21} \\
R_1 + R_{1r} & \leq \max(n_{11}, n_{21}) \\
R_1 + R_{2r} & \leq \max(n_{11}, n_{12}) \\
R_1 + R_2 + R_{2r} & \leq \max(n_{11}, n_{12}) + (n_{22} - n_{12})^+ \\
R_{2r} & \leq n_{12} \\
R_2 + R_{2r} & \leq \max(n_{22}, n_{12}) \\
R_2 + R_{1r} & \leq \max(n_{22}, n_{21}) \\
R_2 + R_1 + R_{1r} & \leq \max(n_{22}, n_{21}) + (n_{11} - n_{21})^+ \\
\end{align*}
\]

When \(R_1r = R_2r = 0\), the region characterized by (22), at least in terms of the pairs \((R_1, R_2)\), corresponds to the region \(\mathcal{C}_{LDIC/FB}\) (Lemma 3). Therefore, the relevance of Lemma 5 relies on the implication that any rate pair \((R_1, R_2) \in \mathcal{C}_{LDIC/FB}\) would not be an \(\eta\)-equilibrium.
where, \( \vec{Y}_1 = (Y_{1,1}, \ldots, Y_{1,n'}) \) be the vector of inputs to receiver during \( n' \) consecutive channel uses. Hence, an upper bound for \( \tilde{R}_1 \) is obtained from the following inequalities:

\[
\tilde{R}_1 \leq \max (n_{i1}, n_{i2}) - (R_{1c} + R_{r1}) + \frac{2}{3} \eta.
\]  

Proof: Without loss of generality, let \( i = 1 \) be the deviating user in the following analysis. After the deviation, the new coding scheme used by transmitter 1 can be of any type. Indeed, with such a new coding scheme, the deviating transmitter might or might not use the signal observations obtained via feedback to generate its codewords. It can also use or not use random symbols and it might possibly have a different block length \( n' \neq N_1 \). Let \( W_1 \) and \( W_{1r} \) be the message and the random message indices sent by transmitter 1 using the new coding scheme, respectively. Let also \( W_{2c} \) and \( W_{2r} \) be the random variables representing the indices of the common and common random messages sent by transmitter 2 using its corresponding randomized Han-Kobayashi scheme with feedback, respectively. Let also \( Y_1 = (Y_1,1, \ldots, Y_1,n') \) be the vector of inputs to receiver 1 during \( n' \) consecutive channel uses. Hence, an upper bound for \( \tilde{R}_1 \) is obtained from the following inequalities:

\[
n' \tilde{R}_1 = H(W_1)
\]

where, \( \vec{Y}_1 = (Y_{1,1}, \ldots, Y_{1,n'}) \) be the vector of symbols sent by transmitter 1 and the vector of common symbols sent by transmitter 2 during \( n' \) consecutive channel uses, respectively. Here, transmitter 1 uses a new coding scheme different from the initial randomized Han-Kobayashi scheme with feedback while transmitter 2 remains using the initial coding scheme. Then, the following holds:

\[
n'(R_{2c} + R_{2r}) = H(W_{2c}, W_{2r})
\]

where (c) follows from the mutual independence between \( W_2 \), \( W_{1r} \) and \( W_{1r} \); and (d) follows from the fact that the encoding function of the second user (which is kept fixed) is such that \( (W_{2c}, W_{2r}) \) can be decoded from \( Y_1 \).

Finally, adding (24) and (25) yields the following upper bound:

\[
\tilde{R}_1 \leq \max (n_{11}, n_{12}) - (R_{2c} + R_{2r}) + \delta(n'),
\]  

where always exists a block length \( n' \) such that \( \delta(n') = \delta_1(n') + \delta_2(n') \) can be made arbitrarily small and thus, \( \delta(n') < \frac{2}{3} \eta \). The same can be proved for the other transmitter-receiver pair by assuming \( i = 2 \), which completes the proof.

Lemma 6 reveals the relevance of the random symbols \( m(t) \) and \( x(t) \) used in the construction of the common words \( x_{1c} \) and \( x_{2c} \) during block \( t \), respectively. Even though the random symbols used by transmitter \( j \) do not increase the achievable rate \( R_j \), they strongly limit the rate improvement transmitter \( i \) can obtain by unilaterally deviating from the initial randomized Han-Kobayashi scheme with feedback. This observation can be used to show that the randomized Han-Kobayashi scheme with feedback can be an \( \eta \)-NE, when both \( R_{1r} \) and \( R_{2r} \) are properly chosen. For instance, for any achievable rate pair \( (R_1, R_2) \in B_{\text{LDIC/FB}} \), there exists a randomized Han-Kobayashi scheme with feedback that achieves the rate tuple

\[
R = (R_{1c} - \eta, R_{1r} - \eta, R_{1p} - \eta, R_{2c} - \eta, R_{2r} - \eta)
\]

with \( R_1 = R_{1p} + R_{1c} - \frac{3}{2} \eta \) and \( R_2 = R_{2p} + R_{2c} - \frac{3}{2} \eta \) and \( \eta \) arbitrarily small. Denote by \( \tilde{R}_{1,\text{max}} \) and \( \tilde{R}_{2,\text{max}} \), the maximum rates achieved by transmitter 1 and 2 when either of them unilaterally changes its transmit configuration. Then, when the rates \( R_{1r} \) and \( R_{2r} \) are chosen such that \( \tilde{R}_{1,\text{max}} - R_{1r} \leq \eta \) and \( \tilde{R}_{2,\text{max}} - R_{2r} \leq \eta \), i.e.

\[
\max (n_{11}, n_{12}) - (R_{2c} + R_{2r}) \geq R_1 \quad \text{and} \quad \max (n_{22}, n_{21}) - (R_{1c} + R_{1r}) \geq R_2,
\]  

any improvement obtained by either transmitter deviating from such an initial scheme is bounded by \( \eta \). However, as shown later, the conditions above can be satisfied at most with strict equality given the conditions in (21). The following lemma formalizes this observation.

Lemma 7: Let \( \eta > 0 \) be an arbitrarily small number and let the rate tuple \( R = (R_{1c} - \eta, R_{1r} - \eta, R_{1p} - \eta, R_{2c} - \eta, R_{2r} - \eta) \).
be achievable with the randomized Han-Kobayashi coding scheme with feedback and satisfy $\forall i \in \{1, 2\}$,

$$R_{ip} + R_{ic}\geq L_i + \frac{1}{3} \eta \text{ and (29)}$$

Then, the rate pair $(R_1, R_2)$, with $R_i = R_{ic} + R_{ip} - \frac{1}{3} \eta$ is a utility pair achieved at an $\eta$-Nash equilibrium.

Proof: Let $(s_1^*, s_2^*) \in A_1 \times A_2$ be an action profile, in which the individual strategy $s_i^*$ is a randomized Han-Kobayashi scheme with feedback satisfying both (29) and (30). Then, from the assumptions of the lemma, $(s_1^*, s_2^*)$ is an $\eta$-NE and $u_1(s_1^*, s_2^*) = R_{ic} + R_{ip} - \frac{1}{3} \eta$ and $u_2(s_1^*, s_2^*) = R_{2c} + R_{2r} - \frac{1}{3} \eta$.

Consider that such a strategy profile $(s_1^*, s_2^*)$ is not an $\eta$-Nash equilibrium. Then, from Def. 1, there exists at least one $i \in \{1, 2\}$ and at least one strategy $s_i \in A_i$ such that the utility $u_i$ is improved by at least $\eta$ bits per block when player $i$ deviates from $s_i^*$ to $s_i$. Without loss of generality, let $i = 1$ be the deviating user and denote by $\tilde{R}_1$ the rate achieved after the deviation. Then,

$$u_1(s_1, s_2^*) = \tilde{R}_1 \geq R_1 + \eta. \quad (31)$$

However, from Lemma 6, it follows that

$$\tilde{R}_1\leq \max(n_{11}, n_{12}) - (R_{2c} + R_{2r}) + \frac{2}{3} \eta. \quad (32)$$

and from the assumption in (30), with $i = 1$, i.e.,

$$R_{2c} + R_{2r} = \max(n_{11}, n_{12}) - (R_{ic} + R_{ip}) + \frac{2}{3} \eta; \quad (33)$$

it follows that

$$\tilde{R}_1\leq R_{ic} + R_{ip} = R_1. \quad (34)$$

The result in (34) contradicts condition (31) and shows that there exists no other coding scheme that brings an individual utility improvement higher than $\eta$. It is interesting to note that from (29), with $i = 1$, it follows that

$$R_{ic} + R_{ip} \geq (n_{11} - n_{12}) + \frac{1}{3} \eta. \quad (35)$$

Then, combining both inequalities (33) and (35), it yields $R_{2c} + R_{2r} < n_{12} + \frac{1}{3} \eta$. This verifies that there always exists a rate $R_{2r}$ that simultaneously satisfies both (29) and (30) and the corresponding conditions in (21). The same can be proved for the other player by assuming $i = 2$. This proves the “if” part of the lemma. Finally, note that if conditions (29) or (30) are not simultaneously met, then an improvement is always feasible as stated in Lemma 4 and Lemma 6. This proves the “only if” part of the lemma and completes the proof.

The following lemma shows that all the rate pairs $(R_1, R_2) \in C_{LDIC/FB} \cap B_{LDIC/FB}$ can be achieved by at least one $\eta$-NE.

Lemma 8: Let $\eta > 0$ be an arbitrarily small number. Then, for all rate pairs $(R_1, R_2) \in C_{LDIC/FB} \cap B_{LDIC/FB}$, there always exists at least one $\eta$-NE strategy profile $(s_1^*, s_2^*) \in A_1 \times A_2$, such that $u_1(s_1^*, s_2^*) = R_1$ and $u_2(s_1^*, s_2^*) = R_2$.

Proof: From Lemma 7, it is known that a strategy profile $(s_1^*, s_2^*)$ in which each player’s transmission scheme is the randomized Han-Kobayashi scheme with feedback satisfying conditions (29) and (30) is an $\eta$-NE and achieves any rate tuple $(R_{1c}, R_{1r}, R_{1p}, R_{2c}, R_{2r}, R_{2p})$. Thus, from the conditions in (21), (29) and (30), the following holds

$$\begin{cases}
R_{1c} + R_{1p} &\geq (n_{11} - n_{12})^+,
R_{1c} + R_{1r} &\leq n_{21},
R_{2p} &\leq (n_{22} - n_{12})^+,
R_{1c} + R_{1p} + R_{2c} + R_{2r} &\geq \max(n_{11}, n_{12}),
R_{2c} + R_{2r} &\leq n_{12},
R_{1p} &\leq (n_{13} - n_{21})^+,
R_{2c} + R_{2p} + R_{1c} + R_{1r} &\leq \max(n_{22}, n_{21}).
\end{cases} \quad (36)$$

The region characterized by (36) can be written in terms of $R_1 = R_{1c} + R_{1p}$ and $R_2 = R_{2c} + R_{2p}$. This yields

$$\begin{cases}
R_1 &\geq (n_{11} - n_{12})^+,
R_1 + R_2 + R_{2r} &\leq \max(n_{11}, n_{12}) + (n_{22} - n_{12})^+,
R_2 &\geq (n_{22} - n_{21})^+,
R_2 + R_1 + R_{1r} &\leq \max(n_{22}, n_{21}) + (n_{11} - n_{21})^+.
\end{cases} \quad (37)$$

Note that depending on the choices of $R_1$ and $R_{2r}$, the conditions described in (37) fully span the region $C_{LDIC/FB} \cap B_{LDIC/FB}$. This implies that for any pair $(R_1, R_2) \in C_{LDIC/FB} \cap B_{LDIC/FB}$, there always exists at least one $\eta$-NE strategy profile that allows transmitters 1 and 2 to achieve the rates $R_1$ and $R_2$, respectively. In this case, such an equilibrium strategy is the randomized Han-Kobayashi coding scheme with feedback whose random components are generated at a rate satisfying condition (30). This completes the proof.

Finally, the proof of Theorem 1 is immediate from Lemma 4 which proves that $N_{LDIC/FB} \subseteq C_{LDIC/FB} \cap B_{LDIC/FB}$; Lemma 5 which proves that the randomized Han-Kobayashi scheme with feedback achieves all the rate pairs $(R_1, R_2) \in C_{LDIC/FB}$; Lemma 7 which proves that when the rates of the random components $R_{1s}$ and $R_{2s}$ are properly chosen, the randomized Han-Kobayashi scheme with feedback is an $\eta$-NE; and finally Lemma 8 which shows that for all rate pairs in $C_{LDIC/FB} \cap B_{LDIC/FB}$ there exists at least one $\eta$-NE strategy profile that achieves it. More specifically, there always exists a randomized Han-Kobayashi scheme with feedback that is an $\eta$-NE and achieves such a rate pair. This verifies that $C_{LDIC/FB} \cap B_{LDIC/FB} \subseteq N_{LDIC/FB}$ and completes the proof of Theorem 1.

IV. GAUSSIAN INTERFERENCE CHANNEL WITH FEEDBACK

This section characterizes the Nash region of the G-IC with feedback (G-IC-FB) following the intuition gained from its equivalent LD-IC-FB model. This result is given in terms of existing inner and outer bounds of the capacity region of the G-IC-FB. Such existing results are briefly described hereunder.

A. Preliminaries

The following definition provides a formal description of a class of bounds known as “approximation to within $b$ units”.

Definition 3 (Approximation to within $\xi$ units): A closed and convex region $X \subseteq \mathbb{R}_+^b$ is approximated to within $\xi$ units
if there exist two sets \( \mathcal{X} \) and \( \overline{\mathcal{X}} \) such that \( \mathcal{X} \subseteq \mathcal{X} \subseteq \overline{\mathcal{X}} \) and \( \forall \mathbf{x} = (x_1, \ldots, x_n) \in \overline{\mathcal{X}} \) then \(((x_1-\xi)^+, \ldots, (x_n-\xi)^+) \in \mathcal{X} \).

Using Def. 3 existing results can be easily described.

1) Capacity Region of the Gaussian IC: The capacity region of the G-IC is denoted by \( C_{GIC} \). An exact characterization of \( C_{GIC} \) is known only for the case of the very weak interference regime \([32],[33],[34]\) and the very strong interference regime \([35],[36]\). In all the other regimes, the capacity region is approximated to within one bit \([37]\) (see Def. 3). The approximation in \([37]\) is given in terms of two regions: (a) A region \( \mathcal{R} \) that is achievable with a "simplified" Han-Kobayashi scheme; and (b) an outer bound of the capacity region, denoted by \( \overline{\mathcal{R}} \). The full descriptions of both \( \mathcal{R} \) and \( \overline{\mathcal{R}} \) are available in \([37]\).

2) Nash Region of the Gaussian IC: The Nash region of the G-IC without feedback is denoted by \( \mathcal{N}_{FB} \) and it has been approximated to within 1 bit in \([22]\). This approximation is given in terms of two other regions: (a) \( \mathcal{R}_{GIC} \) that acts as an inner bound; and (b) \( \overline{\mathcal{R}}_{GIC} \) that acts as an outer bound. Here,

\[
\begin{align*}
\mathcal{N}_{FB} &= \{(R_1, R_2) : B_1 \leq R_i \leq A_i, \ i \in \{1, 2\}\}, \\
\mathcal{N}_{GIC} &= \{(R_1, R_2) : B_1 \leq R_i \leq \max \{A_i - 1, B_i\} \ i \in \{1, 2\}\},
\end{align*}
\]

and,

\[
\begin{align*}
B_i &= \log \left(1 + \frac{\text{SNR}_i}{1 + \text{INR}_{ij}}\right) \quad \text{and} \\
A_i &= \min \left(\log \left(1 + \text{SNR}_i + \text{INR}_{ij}\right), \right.
\left.- \log \left(1 + \frac{[\text{SNR}_j - \max \{\text{INR}_{ji}, \text{SNR}_j/\text{INR}_{ij}\}]^+}{1 + \text{INR}_{jj} + \max \{\text{INR}_{ji}, \text{SNR}_j/\text{INR}_{ij}\}}\right), \right.
\left.\log \left(1 + \text{SNR}_i\right)\right) \quad \text{for} \quad \text{all} \quad i \in \{1, 2\}.
\end{align*}
\]

Note that \( B_i \) is the rate achieved by the transmitter-receiver pair \( i \) when it saturates the power constraint in \((3)\) and treats interference as noise. Following this notation, the Nash region of the two-user G-IC can be written as in the following lemma.

**Lemma 9 (Theorem 2 in \([22]\)):** The Nash region of the two-user G-IC \( \mathcal{N}_{GIC} \) is approximated to within 1 bit by the regions \( \mathcal{R} \cap \mathcal{R}_{GIC} \) and \( \overline{\mathcal{R}} \cap \overline{\mathcal{R}}_{GIC} \) and thus,

\[
\mathcal{R} \cap \mathcal{R}_{GIC} \subseteq \mathcal{N}_{GIC} \subseteq \mathcal{R} \cap \overline{\mathcal{R}}_{GIC}.
\]

3) Capacity Region of the Gaussian IC with Feedback: The most precise approximation of \( C_{GIC/FB} \) is given by Suh and Tse in \([15]\). This approximation to within two bits (Def. 3) is given in terms of two regions: (a) a region \( \mathcal{R}_{FB} \) achievable with a simplified Han-Kobayashi scheme with feedback that uses block Markov encoding and backward decoding (Theorem 2 in \([15]\)); and (b) an outer-bound region (Theorem 3 in \([15]\)), denoted in the following by \( \overline{\mathcal{R}}_{FB} \). The set of pairs \((R_1, R_2) \in \mathcal{R}_{FB} \) satisfy the following set of inequalities for a given \( \rho \in [0, 1] \):

\[
\begin{align*}
R_1 &\leq \log \left(1 + \text{SNR}_1 + \text{INR}_{12}\right) \\
&+ 2\rho \sqrt{\text{SNR}_1 \cdot \text{INR}_{12}}, \quad \text{(41)} \\
R_1 &\leq \log \left(1 + (1 - \rho)\text{INR}_{21}\right) \\
&+ \log \left(2 + \frac{\text{SNR}_1}{\text{INR}_{21}}\right) - 2, \quad \text{(42)} \\
R_2 &\leq \log \left(1 + \text{SNR}_2 + \text{INR}_{21}\right) \\
&+ 2\rho \sqrt{\text{SNR}_2 \cdot \text{INR}_{21}}, \quad \text{(43)} \\
R_2 &\leq \log \left(1 + (1 - \rho)\text{INR}_{12}\right) \\
&+ \log \left(2 + \frac{\text{SNR}_2}{\text{INR}_{12}}\right) - 2, \quad \text{(44)} \\
R_1 + R_2 &\leq \log \left(2 + \frac{\text{SNR}_1}{\text{INR}_{21}}\right) + \log \left(1 + \text{SNR}_2 + \right.

\text{INR}_{21} + 2\rho \sqrt{\text{SNR}_2 \cdot \text{INR}_{21}} - 2, \quad \text{(45)} \\
R_1 + R_2 &\leq \log \left(2 + \frac{\text{SNR}_2}{\text{INR}_{12}}\right) + \log \left(1 + \text{SNR}_1 + \right.

\text{INR}_{12} + 2\rho \sqrt{\text{SNR}_1 \cdot \text{INR}_{12}} - 2. \quad \text{(46)}
\end{align*}
\]

The region \( \overline{\mathcal{R}}_{FB} \) is an outer bound of the capacity region, i.e., \( C_{GIC/FB} \subseteq \overline{\mathcal{R}}_{FB} \). The region \( \overline{\mathcal{R}}_{FB} \) is the set of pairs \((R_1, R_2) \) that satisfy the following set of inequalities, with \( \rho \in [0, 1] \):

\[
\begin{align*}
R_1 &\leq \log \left(1 + \text{SNR}_1 + \text{INR}_{12}\right) \\
&+ 2\rho \sqrt{\text{SNR}_1 \cdot \text{INR}_{12}}, \quad \text{(47)} \\
R_1 &\leq \log \left(1 + (1 - \rho^2)\text{INR}_{21}\right) \\
&+ \log \left(1 + (1 - \rho^2)\text{SNR}_1\right), \quad \text{(48)} \\
R_2 &\leq \log \left(1 + \text{SNR}_2 + \text{INR}_{21}\right) \\
&+ 2\rho \sqrt{\text{SNR}_2 \cdot \text{INR}_{21}}, \quad \text{(49)} \\
R_2 &\leq \log \left(1 + (1 - \rho^2)\text{INR}_{12}\right) \\
&+ \log \left(1 + (1 - \rho^2)\text{SNR}_2\right), \quad \text{(50)} \\
R_1 + R_2 &\leq \log \left(1 + \frac{(1 - \rho^2)\text{SNR}_1}{1 + (1 - \rho^2)\text{INR}_{21}}\right) \\
&+ \log \left(1 + \text{SNR}_2 + \text{INR}_{21}\right) + 2\rho \sqrt{\text{SNR}_2 \cdot \text{INR}_{21}} \quad \text{and} \quad \text{(51)} \\
R_1 + R_2 &\leq \log \left(1 + \frac{(1 - \rho^2)\text{SNR}_2}{1 + (1 - \rho^2)\text{INR}_{12}}\right) \\
&+ \log \left(1 + \text{SNR}_1 + \text{INR}_{12}\right) + 2\rho \sqrt{\text{SNR}_1 \cdot \text{INR}_{12}}. \quad \text{(52)}
\end{align*}
\]

The approximation of \( C_{GIC/FB} \) is described in terms of \( \mathcal{R}_{FB} \) and \( \overline{\mathcal{R}}_{FB} \) by the following lemma.
Lemma 10 (Theorem 4 in [15]): The capacity region $C_{GIC/FB}$ is approximated to within two bits by the regions $\overline{R}_{FB}$ and $\underline{R}_{FB}$.

B. Main Results

In this subsection, the NE region of the GIC-FB $N_{GIC/FB}$ is approximated to within two bits. This approximation is given in terms of three regions: $\overline{R}_{FB}$, $\underline{R}_{FB}$ and $B_{GIC/FB}$, where the open region $B_{GIC/FB}$ is

$$B_{GIC/FB} = \{(R_1, R_2) : R_i \geq B_i, \ i \in \{1, 2\}\} \tag{53}$$

with $B_i$ given in (38). Using these elements, the main result is given in the following theorem.

Theorem 2 (Nash Region of the G-IC w. Feedback): The Nash region $N_{GIC/FB}$ of the Gaussian interference channel with perfect output feedback satisfies that

$$\overline{R}_{FB} \cap B_{GIC/FB} \subseteq N_{GIC/FB} \subseteq \underline{R}_{FB} \cap B_{GIC/FB}. \tag{54}$$

It is worth noting that Theorem 2 is analogous to Theorem 1. Indeed, the Nash equilibrium region $N_{DIC/FB}$ presented in Theorem 1 can be shown to be at a constant gap from the Nash equilibrium region $N_{GIC/FB}$ presented in Theorem 2, when both $n_{ii} = \log_2(\text{SNR}_i)$ and $n_{ij} = \log_2(\text{INR}_{ij})$. Therefore, the discussion presented in the previous section provides significant insight to interpret Theorem 2. The relevance of Theorem 2 relies on two important implications: (a) If the pair of configurations $(s_1, s_2)$ is an NE, then players 1 and 2 always achieve a rate equal or higher than $B_1$ and $B_2$ in (38), respectively; and (b) There always exists a Nash equilibrium configuration pair $(s_1, s_2)$ that achieves a rate pair $(R_1(s_1, s_2), R_2(s_1, s_2))$ that is at most 2 bits/s/Hz per user away from the outer bound of the capacity region.

Implication (a) refers to the fact that at an NE, the worst individual rates are those obtained by saturating the power constraint (3), treating the interference as noise and neglecting the output observations obtained by feedback. This implies that the use of feedback in the decentralized G-IC does not induce rate pairs that can be Pareto dominated by rate pairs in the NE region without feedback, i.e., by implementing feedback, none of the transmitter-receiver pairs degrades its individual transmission rate with respect to the case without feedback. This observation might appear obvious, however, it has been shown that in decentralized systems, when radio devices are granted with higher action spaces, the individual rates or even the total sum rate can be substantially degraded at the equilibria [25], [30].

Implication (b) refers to the fact that the use of feedback in the G-IC allows the achievability of all rate pairs of the capacity region as long as the individual rates are higher than $B_1$ and $B_2$ bits/s/Hz, respectively. This includes all the rate pairs that are at most two bits away from the strictly Pareto optimal boundary. Following both implication (a) and (b), it holds that

$$N_{GIC} \subseteq N_{GIC/FB} \subseteq C_{GIC/FB}. \tag{55}$$

with strict inclusion between $N_{GIC}$ and $N_{GIC/FB}$ in all the interference regimes.

A final comment on Theorem 2, which applies also to Theorem 1, is that it does not provide any insight into the uniqueness of the $\eta$-NE strategy pair $(s_1, s_2)$ that generates the rate pair $(R_1(s_1, s_2), R_2(s_1, s_2))$. Indeed, it is possible that another equilibrium configuration pair $(s'_1, s'_2)$ generates the same rate pair, i.e., $(R_1(s_1, s_2), R_2(s_1, s_2)) = (R_1(s'_1, s'_2), R_2(s'_1, s'_2))$. Theorem 2 does not characterize the set of NE configurations but the set of rate pairs observed at any of the possible NE of the decentralized G-IC with feedback. The characterization of all the configuration pairs $(s_1, s_2)$ that are NEs is still an open problem.

C. Proofs

The proof of Theorem 2 closely follows along the same lines as the proof of Theorem 1. In the first part of this proof, it is shown that a rate pair $(R_1, R_2)$, with $R_i \leq \log_2(1 + \text{SNR}_i \frac{x_{ic}}{1+\text{INR}_{ij}})$ for at least one $i \in \{1, 2\}$, is not an $\eta$-equilibrium, with $\eta > 0$ and arbitrarily small, that is,

$$N_{GIC/FB} \subseteq \overline{R}_{FB} \cap B_{GIC/FB}. \tag{56}$$

In the second part of the proof, it is shown that any point in $\overline{R}_{FB} \cap B_{GIC/FB}$ is an $\eta$-equilibrium, with $\eta > 0$ and arbitrarily small, that is,

$$\overline{R}_{FB} \cap B_{GIC/FB} \subseteq N_{GIC/FB}. \tag{57}$$

which proves Theorem 2.

1) Non-Equilibrium Rate Tuples: The first part of the proof of Theorem 2 is presented in the following lemma.

Lemma 11: A rate pair $(R_1, R_2) \in C_{GIC/FB}$, with either $R_1 < \log_2(1 + \frac{\text{SNR}_1}{1+\text{INR}_{ij}})$ or $R_2 < \log_2(1 + \frac{\text{SNR}_2}{1+\text{INR}_{ij}})$ is not an $\eta$-equilibrium, with $\eta > 0$, arbitrarily small.

Proof: Let $(s_1, s_2)$ be an action profile such that users achieve the rate pair $R_1 = R_1(s_1, s_2)$ and $R_2 = R_2(s_1, s_2)$, respectively, and assume $(s_1, s_2)$ is an $\eta$-NE. Hence, from Def. 1, it holds that none of the players can increase its payoff by unilaterally deviating from $(s_1, s_2)$. Without loss of generality, let at least $R_1(s_1, s_2) < \log_2(1 + \frac{\text{SNR}_1}{1+\text{INR}_{ij}})$. Then, note that independently of the transmit configuration of player 2, player 1 can always use a transmission configuration $s'_1$ in which transmitter 1 saturates the average power constraint (3), interference is treated as noise and the output samples obtained by feedback from its own receiver are not used to generate the codewords during the complete transmission. Thus, player 1 is always able to achieve the rate $R(s'_1, s_2) = \log_2(1 + \frac{\text{SNR}_1}{1+\text{INR}_{ij}})$, which implies that a utility improvement $\eta = R(s'_1, s_2) - R(s_1, s_2) > 0$ is always possible. Thus, the assumption that the rate pair $(s_1, s_2)$ is an NE does not hold since $\eta$ is bounded away from zero. Hence, an NE pair does not exist outside the region $\overline{R}_{FB} \cap B_{GIC/FB}$, which implies that $N_{GIC/FB} \subseteq \overline{R}_{FB} \cap B_{GIC/FB}$ and completes the proof.

2) Achievable Equilibrium Rate Tuples: Consider the randomized Han-Kobayashi scheme with feedback introduced in Sec. III-B, for the case of the LD-IC model. This coding scheme can be extended to the Gaussian case by letting $X_{ip}$ and $\lambda_{ic}$ be the set of private and common codewords of length $n$ symbols for transmitter $i$ such that $\forall x_{ic} \in X_{ic}$, $\frac{1}{n}E[x_{ip}^{2}] \leq \lambda_{i,c}$ and $\forall x_{ip} \in X_{ip}$, $\frac{1}{n}E[x_{ip}^{2}] \leq \lambda_{i,p}$. 


The terms $\lambda_{ip}$ and $\lambda_{ic}$ are the fractions of power assigned to the common and private codewords, i.e., $\lambda_{ic} + \lambda_{ip} \leq 1$. As suggested in [15], the fraction $\lambda_{ip}$ is chosen such that the interference produced at receiver $j$ is at the level of the noise, i.e., $\lambda_{ip} \text{INR}_{ij} \leq 1$ and thus, $\forall i \in \{1, 2\}$,

$$\lambda_{i,p} = \begin{cases} \min \left( \frac{1}{\text{INR}_{ii}}, 1 \right) & \text{if INR}_{ij} < \text{SNR}_i, \\ 0 & \text{otherwise}. \end{cases} \quad (58)$$

This choice of the power allocation reproduces the main assumption of the linear deterministic model in which the private messages do not appear in the other receiver as they are seen at a lower or equal level than the noise. More interestingly, note that by using this power allocation, transmitter $i$ uses message splitting only in the weak interference regime ($1 < \text{INR}_{ij} < \text{SNR}_i$). The reasoning behind is that in the strong interference regime ($\text{INR}_{ij} > \text{SNR}_i$), transmitter $i$ uses the alternative path provided by feedback to communicate with receiver $i$, i.e., the link transmitter $i$ - receiver $j$ - transmitter $j$ - receiver $i$, and thus, no private message is used, i.e., $\lambda_{ip} = 0$.

Conversely, in the very weak interference regime, $\text{INR}_{ij} < 1$, no common message is used and transmitter $i$ privileges the private messages, i.e., $\lambda_{ip} = 1$, under the assumption that there is no significant loss of performance by treating the interference as noise. Therefore, in the following, unless explicitly mentioned, the focus is on the weak interference regime and upper regimes, in which at least $\text{INR}_{12} \geq 1$ and $\text{INR}_{21} \geq 1$.

As in the LD-IC case, at each block $t$ transmitter $i$ sends the symbol $x_{i}^{(t)} = x_{ic}^{(t)} + x_{ip}^{(t)}$. The common codeword $x_{ic}^{(t)}$ is generated by using the common message index $m_{ic}^{(t)} \in \{1, \ldots, 2^{nR_{ic}}\}$ and a randomly generated index $m_{ir}^{(t)} \in \{1, \ldots, 2^{nR_{ir}}\}$. Transmitter $i$ uses the mapping $f_{ic} : \{1, \ldots, 2^{nR_{ic}}\} \times \{1, \ldots, 2^{nR_{ir}}\} \rightarrow X_{ic}$ and thus, $f_{ic}(m_{ic}^{(t)}, m_{ir}^{(t)}) = x_{ic}^{(t)}$. The private part $x_{ip}^{(t)}$ is generated using the private message index $m_{ip}^{(t)} \in \{1, \ldots, 2^{nR_{ip}}\}$, the common codeword $x_{ic}^{(t)}$, and the two previous common codewords $x_{ic}^{(t-1)}$ and $x_{ic}^{(t-2)}$. The random indices $m_{ir}^{(t)}$ are uniformly drawn from some predefined and finite sets, are assumed to be known at the receiver $i$ and thus, they do not convey any new information to receiver $i$. That is, if transmitter $i$ achieves a rate tuple $(R_{ip}, R_{ic}, R_{ir})$, its actual rate is $R_i = R_{ic} + R_{ip}$.

This scheme is thoroughly described in Appendix B and its achievable region is determined by the following lemma.

**Lemma 12:** The achievable region of the randomized Han-Kobayashi coding scheme with feedback in the G-IC is the set of tuples $(R_{ic}, R_{ir}, R_{ip}, R_{2c}, R_{2r}, R_{2p})$ that satisfy, $\forall i \in \{1, 2\}$ and $\forall \rho \in [0, 1]$,

$$R_{ic} + R_{ir} \leq \log (1 + (1 - \rho)\text{INR}_{ij}) - 1, \quad (59)$$

$$R_{ip} \leq \log \left(\frac{2 + \text{SNR}_i}{\text{INR}_{ij}}\right) - 1 \quad \text{and} \quad (60)$$

$$R_{ic} + R_{ip} + R_{jc} + R_{jr} \leq \log \left(1 + \text{SNR}_i + \text{INR}_{ij}\right) + 2\rho \sqrt{\text{SNR}_i \text{INR}_{ij}} - 1. \quad (61)$$

The proof of Lemma 12 is presented in Appendix B. The set of inequalities in Lemma 12 can be written in terms of $R_1 = R_{1c} + R_{1p}$ and $R_2 = R_{2c} + R_{2p}$. This yields the following set of conditions:

$$R_{1r} \leq \log \left(1 + (1 - \rho)\text{INR}_{21}\right) - 1,$$

$$R_{1} + R_{1r} \leq \log \left(1 + (1 - \rho)\text{INR}_{21}\right) + 2\rho \sqrt{\text{SNR}_1 \text{INR}_{12}} - 1,$$

$$R_{2r} \leq \log \left(1 + (1 - \rho)\text{INR}_{12}\right) - 1,$$

$$R_{2} + R_{2r} \leq \log \left(1 + (1 - \rho)\text{INR}_{12}\right) + 2\rho \sqrt{\text{SNR}_2 \text{INR}_{21}} - 1,$$

$$R_{1} + R_{2} + R_{2r} \leq \log \left(1 + \text{SNR}_1 + \text{INR}_{12}\right) + 2\rho \sqrt{\text{SNR}_1 \text{INR}_{12}} - 2$$

$$R_{2} + R_{1} + R_{1r} \leq \log \left(1 + \text{SNR}_2 + \text{INR}_{21}\right) + 2\rho \sqrt{\text{SNR}_2 \text{INR}_{21}} - 2. \quad (62)$$

It is worth noting that this set of inequalities spans the achievable regions $\mathcal{R}_{FB}$ of the simplified Han-Kobayashi scheme with feedback (see (41) - (46)) when both $R_{1r} = 0$ and $R_{2r} = 0$. This implies that any rate pair $(R_1, R_2) \in \mathcal{R}_{FB}$ is also achievable by the randomized Han-Kobayashi scheme with feedback (see Appendices A and B) as long as the rates $R_{1r}$ and $R_{2r}$ are properly chosen.

The following lemma determines the maximum rate improvement that can be achieved by a transmitter that unilaterally deviates from a strategy pair in which both transmitters use the randomized Han-Kobayashi scheme with feedback. The statement of the lemma as well as its proof are analogous to Lemma 6 in the LD-IC case with feedback.

**Lemma 13:** Let $\eta \geq 0$ be an arbitrarily small number and let the rate tuple $R = (R_{1c}, R_{1p}, R_{2c}, R_{2r}, R_{2p})$ be achievable with the randomized Han-Kobayashi coding scheme with feedback such that $R_{1} = R_{1p} + R_{1c} - \frac{1}{2} \eta$ and $R_{2} = R_{2p} + R_{2c} - \frac{1}{2} \eta$. Then, any unilateral deviation of player $i$ by using any other coding scheme leads to a transmission rate $R_i$ that satisfies

$$R_{i} \leq \log (1 + \text{SNR}_i + \text{INR}_{ij} + 2\sqrt{\text{SNR}_i \text{INR}_{ij}}1_{(\rho \neq 0)}) - 1 - (R_{jc} + R_{jr}) + \frac{2}{3} \eta, \quad (63)$$

where $\rho$ is the Pearson correlation coefficient between the symbols of the codewords $x_{1}^{(t)}$ and $x_{2}^{(t)}$, $t \in \{0, \ldots, T\}$. 


Proof: From Lemma 12, it is known that for all rate tuples \((R_1, R_2) \in \mathbb{R}_{FP}\), there always exists a randomized Han Kobayashi scheme with feedback that achieves a rate tuple \( \mathbf{R} = (R_{1c} - \frac{q}{2}, R_{1p} - \frac{q}{2}, R_{2c} - \frac{q}{2}, R_{2p} - \frac{q}{2}) \), with \( R_1 = R_{1p} + R_{1c} - \frac{q}{2} \eta \) and \( R_2 = R_{2p} + R_{2c} - \frac{q}{2} \eta \) arbitrarily small. Let both transmitters use the corresponding Han-Kobayashi scheme following the code construction in (101) and the power allocation in (98) - (100) Without loss of generality, let transmitter 1 change its coding scheme while transmitter 2 remains using the same coding scheme. The new coding scheme adopted by transmitter 1 might or might not use feedback and/or any random symbols. Indeed, it might even use a different code length \( n' \neq N_1 \). Let \( W_1, W_{2c} \), and \( W_{2b} \) be random variables representing the index of the message sent by transmitter 1 and the index of the common and common random messages sent by transmitter 2, respectively. After the deviation of transmitter 1, if its new coding scheme uses random symbols, the random message index is denoted by \( W_{1r} \), and it is assumed to be known at both transmitter 1 and receiver 1. Let \( \mathbf{X}_1 \) be the vector of symbols sent by transmitter 1 during \( n' \) consecutive channel uses with the coding scheme chosen after the deviation. Let also \( \mathbf{Y}_1 = (Y_{1,1}, \ldots, Y_{1,n'}) \) be the vector of inputs to receiver 1 during \( n' \) consecutive channel uses. Hence, an upper bound for \( R_1 \) is obtained from the following inequalities:

\[
\begin{align*}
n' \tilde{R}_1 &= H(W_1) \\
&= H(W_1 | W_{1r}) \\
&= I(W_1; \mathbf{Y}_1 | W_{1r}) + h(W_1 | W_{1r}, \mathbf{Y}_1) \\
&\leq I(W_1; \mathbf{Y}_1 | W_{1r}) + n' \delta_1(n') \\
&= h(\mathbf{Y}_1 | W_{1r}) - h(\mathbf{Y}_1 | W_{1r}, W_{1c}) + n' \delta_1(n') \\
&= \sum_{m=1}^{n'} h(Y_{1,m} | Y_{1,1}, \ldots, Y_{1,m-1}, W_{1r}) \\
&\quad - h(Y_{1,m} | W_{1r}, W_{1c}) + n' \delta_1(n') \\
&\leq n' \cdot \log (1 + SNR_1 + INR_{12}) \\
&\quad + 2 \tilde{\rho} \sqrt{SNR_1 INR_{12}} - h(\mathbf{Y}_1 | W_{1r}, W_{1c}) + n' \delta_1(n'), \\
&\leq n' \cdot \log (1 + SNR_1 + INR_{12}) \\
&\quad + 2 \sqrt{SNR_1 INR_{12}} \mathbb{P}(\rho \neq 0) - h(\mathbf{Y}_1 | W_{1r}, W_{1c}) \\
&\quad + n' \delta_1(n'),
\end{align*}
\]

where, (a) follows from Fano’s inequality, as the rate \( \tilde{R}_1 \) is achievable from the assumptions of the lemma, and thus, \( W_1 \) can be reliably decodable from \( \mathbf{Y}_1 \); (b) follows from the fact that \( h(Y_{1,m} | Y_{1,1}, \ldots, Y_{1,m-1}, W_{1r}) \leq h(Y_{1,m}) \leq \log (1 + SNR_1 + INR_{12} + 2 \tilde{\rho} \sqrt{SNR_1 INR_{12}}), \forall m \in \{1, \ldots, n'\} \), where \( \tilde{\rho} \) is the Pearson correlation coefficient between \( \hat{X}_{1,m} \) and \( X_{2,m} \), i.e.,

\[
\tilde{\rho}_m = \frac{\mathbb{E}[\hat{X}_{1,m} X_{2,m}]}{\sqrt{\mathbb{E}[\hat{X}_{1,m}^2] \mathbb{E}[X_{2,m}^2]}}.
\]

Therefore, the following holds

\[
\begin{align*}
\sum_{m=1}^{n'} h(Y_{1,m} | Y_{1,1}, \ldots, Y_{1,m-1}, W_{1r}) \\
&\leq \sum_{m=1}^{n'} h(Y_{1,m}) \\
&\leq \sum_{m=1}^{n'} \log (1 + SNR_1 + INR_{12}) \\
&\leq n' \log (1 + SNR_1 + INR_{12} + 2 \tilde{\rho} \sqrt{SNR_1 INR_{12}}),
\end{align*}
\]

where, \( \tilde{\rho} = \frac{1}{n'} \sum \tilde{\rho}_m \). Finally, (c) follows from the fact that \( \tilde{\rho} > 0 \) if and only if a strictly positive power is allocated to the symbol \( U \) in (101) at each channel use, i.e., \( \rho \neq 0 \) (see Appendix B). Consider first the case when \( \rho = 0 \) and later, the case when \( \rho > 0 \).

When \( \rho = 0 \), before any deviation, transmitter \( i \) allocates all the transmit power to the symbols \( U_i \) and \( X_{ip} \) at each channel use, and thus \( X_i = U_i + X_{ip} \), with \( i \in \{1, 2\} \). The symbols \( U_i \) are deterministically generated by \( W_{ic} \) and \( W_{ir} \) by the encoding function \( f_{ic} \). The symbol \( X_{ip} \) is deterministically generated by \( W_{ip} \), \( U_i \) and \( U \) by the encoding function \( f_{ip} \). Therefore, after any deviation of player 1, the average Pearson’s correlation coefficient is \( \tilde{\rho} = 0 \) due to the independence of \( U_2 \) and \( X_{2p} \) with any possible \( X_1 \) that can be generated by transmitter 1 at any given channel use.

When \( \rho > 0 \), before any deviation, transmitter \( i \) generates codewords of the form \( X_i = U_i + X_{ip} \), with \( i \in \{1, 2\} \). The symbols \( U_i \) are deterministically generated by \( W_{ic} \) and \( W_{ir} \) by the encoding function \( f_{ic} \). Therefore, after any deviation of player 1, the average Pearson’s correlation coefficient \( \tilde{\rho} \) can be increased by transmitter 1 with respect to the initial value \( \rho \) by simply allocating more power to the symbols \( U \) at every channel use.

To refine this upper bound, the term \( h(Y_1 | W_1) \) in (64) can be lower bounded. Let \( \mathbf{X}_{2c} = U + U_2 \) and \( \mathbf{X}_{2p} \) be the vector of common and private symbols sent by transmitter 2 during \( n' \) consecutive channel uses using the randomized Han-Kobayashi scheme with feedback. The noise component at each of the channel uses at receiver 1 are represented by the vector \( \mathbf{Z}_1 \). Hence, it follows that

\[
h(Y_1^{[n']} | W_1, W_{1r}) = \sum_{m=1}^{n'} h(Y_{1,m} | W_1, W_{1r}, Y_{1,1}, \ldots, Y_{1,m-1})
\]

\[
\geq \sum_{m=1}^{n'} h(Y_{1,m} | W_1, W_{1r}, Y_{1,1}, \ldots, Y_{1,m-1}, \hat{X}_{1,m})
\]

\[
\geq \sum_{m=1}^{n'} h(\sqrt{INR_{12}}(X_{2c,m} + X_{2p,m}) + Z_{1,m})
\]

\[
= \sum_{m=1}^{n'} \log (1 + INR_{12})
\]

\[
\geq n'(R_{2c} + R_{2r} + 1).
\]
\{1, \ldots, n\}', \tilde{X}_{1,m} is deterministically obtained from \(W_1, W_1\) and \(Y_{1,1}, \ldots, Y_{1,m}\), given the coding function of transmitter 1; (e) follows from the signal construction in (1) and the independence of \(X_{2c,m} \) and \(X_{2p,m}\) with \(W_1, W_1\) and \(Y_{1,1}, \ldots, Y_{1,m-1}\) for all \(m \in \{1, \ldots, n\}\); and finally, (f) follows from (59). Substituting (67) into (64) and considering that there always exists a block length \(n'\) such that \(\delta_1(n')\) can be made arbitrarily small and thus, \(\delta_1(n') < \frac{2}{3} \eta\), the following holds:

\[
\tilde{R}_1 \leq \log \left( 1 + \frac{1}{1+SNR_{i} + INR_{i,j} + 2\sqrt{SNR_1 INR_{12} 1}_{\rho \neq 0} } \right) - (R_{2c} + R_{2r}) - 1 - \frac{2}{3} \eta. \tag{68}
\]

The same can be proved for the other transmitter-receiver pair by assuming \(i = 2\) and this completes the proof.

The conclusion from Lemma 13 is analogues to the conclusion obtained from Lemma 6. For instance, for any achievable rate pair \((R_1, R_2) \in B_{GIC/FB} \cap \bar{B}_{FB}\), there exists a randomized Han-Kobayashi scheme with feedback that achieves the rate tuple \(R = (R_{1c} - \frac{\eta}{6}, R_{1r} - \frac{\eta}{6}, R_{1p} - \frac{\eta}{6}, R_{c} - \frac{\eta}{6}, R_{2r} - \frac{\eta}{6}, R_{2p} - \frac{\eta}{6})\), with \(R_1 = R_{1p} + R_{1c} - \frac{2}{3} \eta\) and \(R_2 = R_{2p} + R_{2c} - \frac{1}{3} \eta\) and \(\eta\) arbitrarily small. Let \(\tilde{R}_{1,\text{max}}\) be the maximum rate achievable by transmitter \(i\) by deviating from such a randomized Han-Kobayashi scheme with feedback. Hence, \(\forall i \in \{1,2\}\),

\[
\tilde{R}_{1,\text{max}} = \log \left( 1 + \frac{1}{1+SNR_{i} + INR_{i,j} + 2\sqrt{SNR_1 INR_{12} 1}_{\rho \neq 0} } \right) - (R_{2c} + R_{2r}) + \frac{2}{3} \eta. \tag{69}
\]

Note that when the rates \(R_{1c}\) and \(R_{2c}\) are chosen such that the rates \(R_1\) and \(R_2\) satisfy \(\tilde{R}_{1,\text{max}} - \tilde{R}_1 \leq \eta\) and \(\tilde{R}_{2,\text{max}} - R_2 \leq \eta\), that is,

\[
\left( \log \left( 1 + \frac{1}{1+SNR_{i} + INR_{i,j} + 2\sqrt{SNR_1 INR_{12} 1}_{\rho \neq 0} } \right) - (R_{2c} + R_{2r}) \right) \leq \tilde{R}_1 \quad \text{and} \quad \left( \log \left( 1 + \frac{1}{1+SNR_{i} + INR_{i,j} + 2\sqrt{SNR_1 INR_{12} 1}_{\rho \neq 0} } \right) - (R_{1c} + R_{1r}) \right) \leq \tilde{R}_2,
\]

then any rate improvement obtained by either transmitter deviating from the initial coding scheme is bounded by \(\eta\). However, (70) and (71) are not the only conditions for the rate pair \((R_1, R_2)\) to be an \(\eta\)-equilibrium. From Lemma 11, the rate pair \((R_1, R_2)\) must also satisfy the conditions

\[
R_1 \geq \log \left( 1 + \frac{1}{1+INR_{12}} \right), \quad \text{and} \quad R_2 \geq \log \left( 1 + \frac{1}{1+INR_{21}} \right). \tag{72}
\]

Therefore, to satisfy (70) at least with equality given (72), the following must hold:

\[
R_{2c} + R_{2r} \leq \log \left( 1 + \frac{2\sqrt{SNR_{1} INR_{12} 1}_{\rho \neq 0} }{1 + SNR_{1} + INR_{12} } \right) + \log \left( 1 + INR_{12} \right) - 1. \tag{74}
\]

However, condition (74) cannot always be satisfied. For instance, any \(R_{2c} + R_{2r} > \log \left( 1 + (1-\rho)INR_{12} \right) - 1\) is outside the achievable region of the randomized Han-Kobayashi scheme with feedback. Nonetheless, condition (74) can always be satisfied when \(\rho = 0\). The following lemma formalizes this observation.

**Lemma 14:** Let \(\eta \geq 0\) be an arbitrarily small number and let the rate tuple \(R = (R_{1c} - \frac{\eta}{6}, R_{1r} - \frac{\eta}{6}, R_{1p} - \frac{\eta}{6}, R_{c} - \frac{\eta}{6}, R_{2r} - \frac{\eta}{6}, R_{2p} - \frac{\eta}{6})\) be achievable with the randomized Han-Kobayashi coding scheme with feedback. Let also \(\rho = 0\) and \(\forall i \in \{1,2\}\),

\[
R_{ip} + R_{ic} \geq \log \left( 1 + \frac{SNR_{i}}{1+INR_{ij}} \right) + \frac{1}{3} \eta. \tag{75}
\]

\[
R_{ic} + R_{ip} + R_{jc} + R_{jr} = \log \left( 1 + \frac{SNR_{i} + INR_{ij}}{1} \right) - 1 + \frac{2}{3} \eta. \tag{76}
\]

Then, the rate pair \((R_1, R_2)\), with \(R_1 = R_{ic} + R_{ip} - \frac{1}{3} \eta\) is a utility pair achieved at an \(\eta\)-Nash equilibrium.

**Proof:** This proof follows the same steps as in the proof of Lemma 7. Let \(s_i^* \in A_i\) be a transmit/receive configuration in which communication takes place using the randomized Han-Kobayashi scheme with feedback, \(\rho = 0\) in the signal construction in (101) and \(R_{1s}\) and \(R_{2s}\) chosen according to conditions (76), with \(i = 1\) and \(i = 2\), respectively. From the assumptions of the lemma such strategy profile \((s_1^*, s_2^*)\) is an \(\eta\)-NE and \(u_i(s_1^*, s_2^*) = R_{ic} + R_{ip} + \frac{1}{3} \eta\).

Now, consider that such a strategy profile \((s_1^*, s_2^*)\) is not an \(\eta\)-equilibrium. Then, from Def. 1, there exists at least one \(i \in \{1,2\}\) and at least one strategy \(s_i \neq s_i^* \in A_i\) such that the utility \(u_i\) is improved by at least \(\eta\) bits per block when player \(i\) deviates from \(s_i^*\) to \(s_i\). Without loss of generality, let \(i = 1\) be the deviating user in the following analysis. Then, the following must hold:

\[
\text{\(u_1(s_1, s_2^*) = \tilde{R}_1 \geq R_1 + \eta. \tag{77}\)
\]

However, from Lemma 13 given that \(\rho = 0\), it follows that

\[
\tilde{R}_1 \leq \log \left( 1 + \frac{1}{1+INR_{12}} \right) - (R_{2c} + R_{2r}) + \frac{2}{3} \eta. \tag{78}\]

From assumption (76), the previous upper bound can be rewritten with strict inequality as

\[
\tilde{R}_1 < R_{1c} + R_{ip} = R_1. \tag{79}\]

This result contradicts condition (77) and shows that there does not exist another coding scheme that brings an individual utility improvement higher than \(\eta\). The same can be proved for the other player by assuming \(i = 2\) and this completes the proof.

The following lemma shows that all the rate pairs \((R_1, R_2) \in R_{FB} \cap B_{GIC/FB}\) can be achieved at an \(\eta\)-NE by a randomized Han-Kobayashi scheme with feedback, and thus, \(R_{FB} \cap B_{GIC/FB} \subseteq N_{GIC/FB}\).

**Lemma 15:** For all rate pairs \((R_1, R_2) \in C_{GIC/FB} \cap B_{GIC/FB}\), there always exists at least one Nash equilibrium strategy profile \((s_1^*, s_2^*) \in A_1 \times A_2\), such that \(u_1(s_1^*, s_2^*) = R_1\) and \(u_2(s_1^*, s_2^*) = R_2\).

**Proof:** From Lemma 14, it is known that the strategy profile \((s_1^*, s_2^*)\) in which each player’s transmission scheme
is the randomized Han-Kobayashi scheme with $\rho = 0$ and feedback satisfying both conditions (75) and (76) is an $\eta$-NE. Therefore, the rate tuple $(R_{1c}, R_{1r}, R_{1p}, R_{2c}, R_{2r}, R_{2p})$ achieved by the transmit configuration $(s^{*}_1, s^{*}_2)$ satisfies, for any $\eta \geq 0$ arbitrarily small and $\forall i \in \{1, 2\}$,
\[
R_{1c} + R_{1r} \leq \log (1 + \text{INR}_{ji}) - 1, \\
R_{1p} \leq \log \left(2 + \frac{\text{SNR}_i}{\text{INR}_{ji}}\right) - 1, \\
R_{1c} + R_{1p} + R_{1c} + R_{1r} = \log \left(1 + \frac{\text{SNR}_i + \text{INR}_{ji}}{1 + \text{INR}_{ji}}\right) - 1 \\
R_{1p} + R_{1c} \geq \log \left(1 + \frac{\text{SNR}_i}{1 + \text{INR}_{ji}}\right).
\]

The region characterized by the inequalities above can be written in terms of $R_1 = R_{1c} + R_{1p}$ and $R_2 = R_{2c} + R_{2p}$. This yields,
\[
\begin{align*}
R_1 & \geq \log \left(1 + \frac{\text{SNR}_i}{\text{INR}_{ji}}\right), \\
R_2 & \geq \log \left(1 + \frac{\text{SNR}_i}{\text{INR}_{ji}}\right), \\
R_1 + R_2 + R_{1r} & \leq \log \left(1 + \text{SNR}_2 + \text{INR}_{21}\right), \\
R_1 + R_2 + R_{2r} & \leq \log \left(1 + \frac{\text{SNR}_2 + \text{INR}_{21}}{2 + \text{SNR}_1 + \text{INR}_{12}}\right) - 2 \\
& \quad + \log \left(2 + \frac{\text{SNR}_2}{\text{SNR}_{12}}\right) - 2.
\end{align*}
\]

(80)
Note that the region described by (80) corresponds exactly to the region $\mathcal{R}_{FB} \cap \mathcal{B}_{GIC/FB}$, with $\rho = 0$ in (41)-(46). This implies that for any pair $(R_1, R_2) \in \mathcal{R}_{FB} \cap \mathcal{B}_{GIC/FB}$, with $\rho = 0$ in (41)-(46), there always exists at least one $\eta$-NE strategy profile that achieves such a rate pair. In this case, such an equilibrium strategy is a randomized Han-Kobayashi coding scheme with $\rho = 0$ and feedback when $R_{1r}$ and $R_{2r}$ are chosen to satisfy the condition (76). This completes the proof.

It is worth highlighting that for all rate pairs $(R_1, R_2) \in \mathcal{R}_{FB} \cap \mathcal{B}_{GIC/FB}$, it also holds that $(R_1 + 2, R_2 + 2) \in \mathcal{R}_{FB} \cap \mathcal{B}_{GIC/FB}$. Hence, the proof of Theorem 2 is immediate from Lemma 11 which proves that $\mathcal{N}_{GIC/FB} \subseteq \mathcal{R}_{FB} \cap \mathcal{B}_{GIC/FB}$; Lemma 12 which proves that the randomized Han-Kobayashi scheme with feedback achieves all the rate pairs $(R_1, R_2) \in \mathcal{R}_{FB}$; Lemma 14 which proves that when $\rho = 0$ and the rates of the random components $R_{1r}$ and $R_{2r}$ are properly chosen, all the rate pairs $(R_1, R_2) \in \mathcal{R}_{FB} \cap \mathcal{B}_{GIC/FB}$ are achieved at an $\eta$-NE; and finally, Lemma 15 that shows that $\mathcal{R}_{FB} \cap \mathcal{B}_{GIC/FB} \subseteq \mathcal{N}_{GIC/FB}$ and completes the proof of Theorem 2.

V. EFFICIENCY OF THE NASH EQUILIBRIA

The price of anarchy (PoA) [38] and the price of stability (PoS) [39] are both measures of the efficiency of the set of equilibria of a game. Basically, the PoA measures the loss of global performance due to decentralization by comparing the maximum sum utility achieved under global control with the minimum sum utility achieved at the Nash equilibrium. The PoS measures also the loss of global performance due to decentralization by comparing the maximum sum utility achieved under the global control with the maximum sum utility achieved at the Nash equilibrium.

In the following, both the PoA and the PoS of the game $G$ in (5) are investigated.

A. Price of Anarchy

The PoA of the game $G$ is
\[
\text{PoA}(G) = \frac{\max_{(s_1, s_2) \in \mathcal{A}} \sum_{i=1}^{2} R_i(s_1, s_2)}{\min_{(s_1, s_2) \in \mathcal{A}_{NE}} \sum_{i=1}^{2} R_i(s_1^*, s_2^*)},
\]
where $\mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2$ is the set of all possible configuration pairs and $\mathcal{A}_{NE} \subset \mathcal{A}$ is the set of NE configuration pairs. From Theorem 2 and the existing inner and outer bounds $\mathcal{R}_{FB}$ and $\mathcal{R}_{FB}$, the following proposition holds.

Proposition 1 (PoA of the G-IC with Feedback): The PoA of the G-IC with feedback satisfies the following condition:
\[
1 < \frac{\min \{\tau_{12}, \tau_{21}\}}{B_1 + B_2} = \frac{\text{PoA}(G)}{B_1 + B_2} < \frac{\min \{v_{12}, v_{21}\}}{B_1 + B_2},
\]
where
\[
\tau_{ij} = \log \left(2 + \frac{\text{SNR}_i}{\text{INR}_{ij}}\right) + \log \left(1 + \frac{\text{SNR}_j + \text{INR}_{ij} + 2\rho \sqrt{\text{SNR}_j \cdot \text{INR}_{ij}}}{1 + \frac{\text{SNR}_i}{\text{INR}_{ij}} - 2}\right) - 2 \\
v_{ij} = \log \left(1 + \frac{(1 - \rho^2)\text{SNR}_i}{1 + (1 - \rho^2)\text{INR}_{ij}}\right) + \log \left(1 + \frac{\text{SNR}_j + \text{INR}_{ij} + 2\rho \sqrt{\text{SNR}_j \cdot \text{INR}_{ij}}}{1 + \frac{\text{SNR}_i}{\text{INR}_{ij}} - 2}\right).
\]

and $\rho \in [0, 1]$.

It is worth noting that the PoA is determined with a precision of two bits since, $\min \{v_{12}, v_{21}\} - \min \{\tau_{12}, \tau_{21}\} \leq 2$ in (82). More importantly, the PoA is strictly larger than 1. This implies that the system can be stabilized at an NE in which the achieved rate pair is not on the boundary of the sum-rate. Therefore, there always exists a loss of performance due to decentralization (anarchical behavior of both transmitter-receiver links) when the equilibrium is not properly selected. Indeed, when both SNR and INR arbitrarily increase at a given rate $\alpha = \frac{\text{SNR}}{\text{INR}} \leq \infty$, the PoA grows unboundedly towards infinity, i.e.,
\[
\lim_{\frac{\text{INR}}{\text{SNR}} \rightarrow \alpha} \text{PoA}(G) = \infty.
\]

More specifically, at high SNR and INR, if the system is stable at the worst NE in terms of sum-rate, i.e., the rate pair $(B_1, B_2)$, the loss in terms of sum-rate grows unboundedly. In Fig. 6, the PoA for the symmetric G-IC with feedback is plotted as a function of SNR = SNR$_1 = $ SNR$_2$ and INR = INR$_{12} = $ INR$_{21}$. Note that the PoA is monotonically increasing with INR and monotonically decreasing with the SNR. This observation confirms that the main reason for the lost of performance at the (worst) NE derives from the effect of interference, which is treated as noise in this case and feedback is not used.
B. Price of Stability

The PoS of the game $\mathcal{G}$ is

$$\text{PoS} (\mathcal{G}) = \max_{(s_1, s_2) \in \mathcal{A}} \sum_{i=1}^{2} R_i(s_1, s_2) \quad \text{or} \quad \max_{(s_1^*, s_2^*) \in \mathcal{A}\mathcal{S}\mathcal{P}} \sum_{i=1}^{2} R_i(s_1^*, s_2^*). \quad (87)$$

From Theorem 2 and the inner and outer bounds $\mathcal{R}_{FB}$ and $\mathcal{R}_{FB}$, the following proposition holds.

**Proposition 2 (PoS of the G-IC with feedback):** The PoS of the G-IC with feedback is $\text{PoS} (\mathcal{G}) = 1$.

The price of stability is equal to one, independently of both the SNR$_1$ and SNR$_2$. This implies that, despite the anarchical behavior of both links, it is always possible to observe an NE in which no loss of performance is observed. Indeed, thanks to the use of feedback, the sum-rate maximizing rate pairs are also achievable at the NE, which guarantees PoS $= 1$.

VI. CONCLUSIONS

In this paper, the Nash equilibrium region of the two-user decentralized Gaussian interference channel with feedback has been approximated to within 2 bits. Using this result, the following two observations have been extensively discussed to highlight the benefits of feedback in the decentralized interference channel: (i) The NE region achieved with feedback is strictly larger than the NE region without feedback. More importantly, the rate pairs uniquely achievable using feedback are at least weakly Pareto superior to those achievable without feedback. (ii) The use of feedback allows the achievability of all the strictly Pareto optimal rate pairs of the (approximated) capacity region of the IC with feedback even when the network is fully decentralized.

One of the directions for future work would be studying the multiplicity of the transmit/receive configurations at each rate pair in the NE region. From the results presented here, only the existence of one NE configuration per NE rate pair can be claimed. However, the possibility that many NE configurations generate the same NE rate pair remains open. Another question that remains unsolved is: How to achieve an NE in the case in which players do not know the set of strategies of all the other players and only local information is available?

The intuition obtained from the results presented here leads to imply that other tools different from those brought from information theory and game theory would be needed to solve these questions. For instance, tools from machine learning and multi-agent learning theory might help to solve these open problems.

Finally, using the price of anarchy and the price of stability, it has been shown that there is no loss of performance due to the anarchical behavior of both links as long as proper equilibrium selection is performed. Otherwise, it has been shown that there always exists a loss of performance in terms of sum-rate and such a loss monotonically increases with the interference to noise ratio and monotonically decreases with the signal to noise ratio. This observation highlights the importance of mechanisms to perform equilibrium selection in decentralized networks.

**APPENDIX A**

**PROOF OF LEMMA 5**

This appendix provides a description of the randomized Han-Kobayashi scheme with feedback used in Sec. III and provides a proof of Lemma 5.

**Codebook Generation:** Fix a joint probability distribution $p(u, u_1, u_2, x_1, x_2) = p(u_1)p(u_2)p(x_1|u_1)p(x_2|u_2)$. Generate $2^{n(R_{1}\text{c} + R_{1}\text{r} + R_{2}\text{r} + R_{2}\text{r})}$ i.i.d. length-$n$ codewords $u(s, r) = \sum_{m=1}^{n} p(u_{m}(s, r), w_{m}(s, r))$, with $s \in \{1, \ldots, 2^{n(R_{1}\text{c} + R_{1}\text{r})}\}$ and $r \in \{1, \ldots, 2^{n(R_{2}\text{r} + R_{2}\text{r})}\}$.

For encoder 1, generate for each codeword $u(s, r)$, $2^{n(R_{1}\text{c} + R_{1}\text{r})}$ i.i.d. length-$n$ codewords $u_1(s, r, k) = (u_1, ..., u_1)$ according to $p(u_1, u_2, x_1, x_2) = \sum_{m=1}^{n} p(u_{m}(s, r), k)|u_{m}(s, r))$, with $k \in \{1, \ldots, 2^{n(R_{1}\text{c} + R_{1}\text{r})}\}$. For each pair of codewords $(u(s, r), u_1(s, r, k))$, generate $2^{n(R_{1}\text{c})}$ i.i.d. length-$n$ codewords $x_1(s, r, k, l) = (x_1, ..., x_1)$ according to $p(x_1, x_2) = \sum_{m=1}^{n} p(x_{m}(s, r, k, l))|u_{m}(s, r), u_{1}(s, r, k))$, with $l \in \{1, \ldots, 2^{n(R_{\text{c}})}\}$.

For encoder 2, generate for each codeword $u(s, r)$, $2^{n(R_{2}\text{c} + R_{2}\text{r})}$ i.i.d. length-$n$ codewords $u_2(s, r, q) = (u_2, ..., u_2, u_2)$ according to $p(u_1, u_2, x_1, x_2) = \sum_{m=1}^{n} p(u_{m}(s, r, q))|u_{m}(s, r))$, with $q \in \{1, \ldots, 2^{n(R_{2}\text{c} + R_{2}\text{r})}\}$. For each pair of codewords $(u(s, r), u_2(s, r, q))$, generate $2^{n(R_{2}\text{c})}$ i.i.d. length-$n$ codewords $x_2(s, r, q, z) = (x_2, ..., x_2)$ according to $p(x_1, x_2) = \sum_{m=1}^{n} p(x_{m}(s, r, k, l))|u_{m}(s, r), u_{1}(s, r, k))$, with $z \in \{1, \ldots, 2^{n(R_{\text{c}})}\}$.

**Encoding:** Let $m_{ic}(t) \in \{1, \ldots, 2^{n(R_{\text{c}})}\}$, $m_{ir}(t) \in \{1, \ldots, 2^{n(R_{\text{r}})}\}$
and $m_{ip}^{(t)} \in \{1, \ldots, 2^{nR_{ip}}\}$ be the index of the common message, the index of the common random message, and the index of the private message sent by transmitter $i$ during block $t$, respectively. For the ease of notation, let $w_{ic}^{(t)} = (m_{ic}^{(t)}, m_{ir}^{(t)}) \in \{1, \ldots, 2^{n(R_{ic} + R_{ir})}\}$ be the (joint) index of the common message. Consider Markov encoding with a length of $T$ blocks. At block $t \in \{1, \ldots, T\}$, transmitter $i$ sends the codeword $d_{ic}^{(t)} = x_{1} \left( w_{1c}^{(t-1)}, w_{2c}^{(t-1)}, w_{ic}^{(t)}, m_{ip}^{(t)} \right)$, where $w_{ic}^{(0)} = s^{t}$, $w_{2c}^{(0)} = r^{t}$, $(T \leftarrow (T)$ is defined, and $w_{2c}^{(T)} = l^{t}$. The 4-tuple $(s^{t}, k^{t}, r^{t}, l^{t}) \in \{1, \ldots, 2^{n(R_{ic} + R_{ir})}\}^{2}$ is predefined and known at both receivers and transmitters. It is worth noting that the message index $w_{ic}^{(t-1)}$ is obtained by transmitter $1$ from the feedback of $y_{1}^{(t-1)}$ at the end of block $t - 1$. That is, for $t > 1$, $w_{ic}^{(t-1)} = w_{ic}^{(t-2)} \in \{1, \ldots, 2^{n(R_{ic} + R_{ir})}\}$, which is the only index that satisfies

$$
\left( u_{1c}^{(w_{1c}^{(t-2)}, w_{2c}^{(t-2)}), u_{1c}^{(w_{1c}^{(t-2)}, w_{2c}^{(t-2)}), w_{ic}^{(t-1)}}, x_{1}^{(w_{1c}^{(t-2)}, w_{2c}^{(t-2)}), w_{ic}^{(t-1)}}, m_{ip}^{(t)}}, u_{2}^{(w_{1c}^{(t-2)}, w_{2c}^{(t-2)}), w_{ic}^{(t-1)}}, y_{1}^{(t-1)} \right) \in \mathcal{A}_{c}^{(n)}, \tag{88}
$$

where $\mathcal{A}_{c}^{(n)}$ represents the set of jointly typical sequences under the assumption that $w_{ic}^{(t-1)}, w_{ic}^{(t-2)}$ have been decoded without errors at transmitter $1$. Transmitter $2$ follows a similar encoding scheme. **Decoding:** Both receivers decode their messages at the end of block $T$ in a backward decoding fashion. For each block $t \in \{1, \ldots, T\}$, receiver $1$ obtains the message indices $(\hat{w}_{1c}^{(T-t)}, \hat{w}_{2c}^{(T-t)}, \hat{m}_{ip}^{(T-t-1)}) \in \{1, \ldots, 2^{n(R_{ic} + R_{ir})}\} \times \{1, \ldots, 2^{n(R_{ic} + R_{ir})}\} \times \{1, \ldots, 2^{nR_{IP}}\}$. The tuple $(\hat{w}_{ic}^{(T-t)}, \hat{w}_{2c}^{(T-t)}, \hat{m}_{ip}^{(T-t-1)})$ is the unique tuple that satisfies

$$
\left( u_{1c}^{(\hat{w}_{1c}^{(T-t)}, \hat{w}_{2c}^{(T-t)}), u_{1c}^{(\hat{w}_{1c}^{(T-t)}, \hat{w}_{2c}^{(T-t)}), \hat{w}_{ic}^{(T-t-1)}}, x_{1}^{(\hat{w}_{1c}^{(T-t)}, \hat{w}_{2c}^{(T-t)}), \hat{w}_{ic}^{(T-t-1)}}, \hat{m}_{ip}^{(T-t-1)}}, u_{2}^{(\hat{w}_{1c}^{(T-t)}, \hat{w}_{2c}^{(T-t)}), \hat{w}_{ic}^{(T-t-1)}}, y_{1}^{(t)} \right) \in \mathcal{A}_{c}^{(n)}, \tag{88}
$$

where $\mathcal{A}_{c}^{(n)}$ represents the set of jointly typical sequences. Receiver $2$ follows a similar decoding scheme. **Probability of Error Analysis:** An error might occur during the coding phase at the beginning of block $t$ if the common random message index $w_{ic}^{(t-1)}$ is not correctly decoded at transmitter $i$. For instance, this error might occur at transmitter $1$ because: (i) there does not exist an index $\tilde{r} \in \{1, \ldots, 2^{n(R_{ic} + R_{ir})}\}$ that satisfies (88), or (ii) several indices simultaneously satisfy (88). From the asymptotic equipartition property (AEP) [40], the probability of an error due to (i) tends to zero when $n$ grows to infinity. The probability of error due to (ii) can be made arbitrarily close to zero when $n$ grows to infinity, if

$$
R_{ic} + R_{ir} \leq I(U_{i}; Y_{j}|X_{j}, U) \tag{89}
$$

An error might occur during the (backward) decoding step $t$ if the messages $w_{1c}^{(t+1)}, w_{2c}^{(t+1)}$ and $m_{ip}^{(t)}$ are not decoded correctly given that the message indices $w_{1c}^{(t)}$ and $w_{2c}^{(t)}$ were correctly decoded in the previous decoding step $t - 1$. These errors might arise for two reasons: (i) there does not exist a pair $(\tilde{s}, \tilde{r}, \tilde{k})$ that satisfies (89), or (ii) there exist several pairs $(\tilde{s}, \tilde{r}, \tilde{k})$ that simultaneously satisfy (89). From the AEP, the probability of an error due to (i) tends to zero as $n$ tends to infinity. Consider the error due to (ii) and define the following event during the decoding interval of block $t$,

$$
E_{srk}^{(t)} = \left\{ \left( u_{1}(s, r), u_{1}(s, r, w_{1c}^{(t)}), x_{1}(s, r, w_{1c}^{(t)}, k), u_{2}(s, r, w_{2c}^{(t)}), y_{1}^{(t)} \right) \in \mathcal{A}_{c}^{(n)} \right\}. \tag{90}
$$

Assume also that at phase $t$ the indices $(\tilde{s}, \tilde{r}, \tilde{k})$ are $(1, 1, 1)$ without loss of generality, due to the symmetry of the code. Then, the probability of error due to (ii) during phase $t$, $p_{e}^{(t)}$, can be bounded as follows:

$$
p_{e}^{(t)} = \text{Pr} \left( \bigcup_{(s, r, k) \neq (1, 1, 1)} E_{srk}^{(t)} \right) + \sum_{s \neq 1, r \neq 1, k \neq 1} \text{Pr} \left( E_{srk}^{(t)} \right) + \sum_{s = 1, r \neq 1, k \neq 1} \text{Pr} \left( E_{srk}^{(t)} \right) + \sum_{s \neq 1, r = 1, k = 1} \text{Pr} \left( E_{srk}^{(t)} \right) + \sum_{s = 1, r = 1, k = 1} \text{Pr} \left( E_{srk}^{(t)} \right)
$$

$$
\leq 2^{n(R_{ic} + R_{ir} + R_{2c} + R_{2r} + R_{1p} - I(U_{2}I_{2}; Y_{1})) + 4c}
+ 2^{n(R_{ic} + R_{ic} + R_{1p} - I(U_{2}I_{2}; X_{1})) + 4c}
+ 2^{n(R_{1c} + R_{1r} + R_{1p} - I(U_{2}I_{2}; X_{1})) + 4c}
+ 2^{n(R_{1c} + R_{1r} + R_{1p} - I(U_{2}I_{2}; X_{1})) + 4c}
+ 2^{n(R_{2c} + R_{2r} - I(U_{2}I_{2}; Y_{1})) + 4c}
\tag{91}
$$

Now, from (89) and (92), given that $I(U_{i}; Y_{j}|X_{j}, U) < I(U_{i}, U_{j}; X_{j})$, the probability of error due to (ii) can be made arbitrarily small if the following conditions hold:

$$
\begin{align*}
R_{2c} + R_{2r} &\leq I(U_{2}; Y_{1}|X_{1}, U) \\
R_{1p} &\leq I(X_{1}; Y_{1}|U_{1}, U_{2}) \\
R_{1c} + R_{1r} + R_{1p} + R_{2c} + R_{2r} &\leq I(U_{2}, U_{j}; X_{1}) \tag{93}
\end{align*}
$$

The common random message index $m_{ip}^{(t)}$ used to generate the common message index $w_{ic}^{(t)} \forall t \in \{1, \ldots, T\}$, are assumed to be known at both transmitter $i$ and receiver $i$. Therefore, the set of inequalities in (94) reduces to the following:

$$
\begin{align*}
R_{2c} + R_{2r} &\leq I(U_{2}; Y_{1}|X_{1}, U) \\
R_{1p} &\leq I(X_{1}; Y_{1}|U_{1}, U_{2}) \tag{94}
\end{align*}
$$

The same analysis is carried out for transmitter $2$ and thus,

$$
\begin{align*}
R_{1c} + R_{1r} &\leq I(U_{1}; Y_{j}|X_{j}, U) \\
R_{2p} &\leq I(X_{2}; Y_{2}|U_{1}, U_{2}) \tag{95}
\end{align*}
$$

$$
\begin{align*}
R_{2c} + R_{2p} + R_{1c} + R_{1r} &\leq I(U_{1}, U_{2}; X_{2}; Y_{2}) \tag{95}
\end{align*}
$$
Suppose that transmitter tuples achievable in the Gaussian interference channel by the deterministic model, it follows that the entire capacity region of $C_{LDIC/FB}$. In terms of the linear deterministic model, it follows that the rate-pairs achievable with the proposed randomized coding scheme described above reduces to the coding scheme presented in [15] and the achievable region corresponds to the entire capacity region of $C_{LDIC/FB}$. Hence, the following inequalities hold:

$$\begin{align*}
R_{ic} + R_{ir} &\leq n_{ij}, \\
R_{ip} &\leq (n_{j} - n_{ij})^{+}, \\
R_{ic} + R_{ip} + R_{jc} + R_{jr} &\leq \max(n_{ii}, n_{ij}),
\end{align*}$$

which completes the proof of Lemma 5.

**APPENDIX B**

**PROOF OF LEMMA 12**

The code generation presented in Appendix A is general and thus, it applies for both the LD-IC model and the G-IC model. Therefore, from the analysis of the error probability, the rate tuples achievable in the Gaussian interference channel by the randomized Han-Kobayashi scheme with feedback satisfy the inequalities in (94) and (95). That is,

$$\begin{align*}
R_{ic} + R_{ir} &\leq I(U; U_{j}|X_{j}, U) \\
R_{ip} &\leq I(X_{i}; X_{i}|U_{1}, U_{2}) \\
R_{ic} + R_{ip} + R_{jc} + R_{jr} &\leq I(U, U_{i}, X_{i}; Y_{i}).
\end{align*}$$

Suppose that transmitter $i$ uses the Gaussian input distribution

$$\begin{align*}
U_{i} &\sim \mathcal{CN}(0, \rho), \\
U_{j} &\sim \mathcal{CN}(0, \lambda_{ic}) \text{ and} \\
X_{ip} &\sim \mathcal{CN}(0, \lambda_{ip}),
\end{align*}$$

with $\rho + \lambda_{ic} + \lambda_{ip} = 1$ and let $\lambda_{ip}$ be determined by (58). Let also

$$X_{i} = U + U_{i} + X_{ip},$$

and assume that $U$, $U_{1}$, $U_{2}$, $X_{ip}$, and $X_{2p}$ are mutually independent. Then the following holds:

$$\begin{align*}
I(U; U_{j}, X_{i}; Y_{j}) &= h(Y_{j}) - h(Y_{j}|U, U_{j}, X_{i}) \\
&= \log(1 + SNR_{i} + INR_{ij} + \lambda_{ip}SNR_{i} + \lambda_{ip}INR_{ij} + 1) \\
&\leq \log(1 + SNR_{i} + INR_{ij}) + 2\rho \sqrt{SNR_{i}INR_{ij}} \\
&\leq \log(1 + \lambda_{ip}INR_{ij}).
\end{align*}$$

The inequality in (a) is explained as follows. Let $\Re(\cdot)$ and $\Im(\cdot)$ represent the real and imaginary parts of a given complex number. Hence,

$$\begin{align*}
h_{ii} &= \Re(h_{ii}) + (\sqrt{-1}) \Im(h_{ii}) \text{ and} \\
h_{ij} &= \Re(h_{ij}) + (\sqrt{-1}) \Im(h_{ij}).
\end{align*}$$

Then,

$$\begin{align*}
\rho (h_{ij}^{*} h_{ii} + h_{ii}^{*} h_{ij}) &= 2\rho \Re(h_{ii}) \Re(h_{ij}) \\
&\leq 2\rho |\Re(h_{ii}) \Re(h_{ij})| \\
&= 2\rho \sqrt{|\Re(h_{ii})^{2} (\Re(h_{ij}))^{2}|} \\
&= 2\rho \sqrt{SNR_{i}INR_{ij}},
\end{align*}$$

which verifies the inequality in (a).

The term $I(U; Y_{j}|U, X_{i})$ satisfies the condition

$$\begin{align*}
I(U; Y_{j}|U, X_{i}) &= h(Y_{j}|U, X_{i}) - h(Y_{j}|U, U_{i}, X_{i}) \\
&= \log(1 + (1 - \rho)INR_{ji}) \\
&\leq \log(1 + \lambda_{ip}INR_{ij}).
\end{align*}$$

Similarly, the term $I(X_{i}; Y_{i}|U, U_{i}, U_{2})$ satisfies the condition

$$\begin{align*}
I(X_{i}; Y_{i}|U, U_{i}, U_{2}) &= h(Y_{i}|U_{1}, U_{2}) - h(Y_{i}|U_{1}, U_{2}, X_{i}) \\
&= \log(\lambda_{ip}SNR_{i} + \lambda_{ip}INR_{ij} + 1) \\
&\leq \log(1 + \lambda_{ip}INR_{ij}).
\end{align*}$$

Finally it follows from (102), (105) and (106) that $\forall i \in \{1, 2\}$,

$$\begin{align*}
R_{ic} + R_{ir} &\leq \log(1 + (1 - \rho)INR_{ji}) \\
R_{ip} &\leq \log(\lambda_{ip}SNR_{i} + \lambda_{ip}INR_{ij} + 1) \\
R_{ic} + R_{ip} + R_{jc} + R_{jr} &\leq \log(1 + SNR_{i} + INR_{ij}) \\
&\leq \log(1 + \lambda_{ip}INR_{ij}).
\end{align*}$$

In the specific scenario in which $INR_{ij} \geq 1$, the inequalities above can be written as follows:

$$\begin{align*}
R_{ic} + R_{ir} &\leq \log(1 + (1 - \rho)INR_{ji}) - 1 \\
R_{ip} &\leq \log(2 + \frac{SNR_{i}}{INR_{ij}}) - 1 \\
R_{ic} + R_{ip} + R_{jc} + R_{jr} &\leq \log(1 + SNR_{i} + INR_{ij}) + 2\rho \sqrt{SNR_{i}INR_{ij}} - 1,
\end{align*}$$

and this concludes the proof.

**REFERENCES**


