Harnessing Bursty Interference in Multicarrier Systems with Output Feedback

Shaunak Mishra, *Student Member, IEEE,* I-Hsiang Wang, *Member, IEEE,* and Suhas Diggavi, *Fellow, IEEE*

***Abstract*—We study parallel 2-user interference channels when the interference is bursty and feedback is available from the respective receivers. Presence of interference in each subcarrier is modeled as a memoryless Bernoulli random state. The states across subcarriers are drawn from an arbitrary joint distribution with the same marginal probability for each subcarrier and instantiated i.i.d. over time. For the linear deterministic setup with symmetric interference in each subcarrier, we give a complete characterization of the capacity region. For the analogous setup with Gaussian noise, we give outer bounds and a tight generalized degrees of freedom characterization. We propose a novel helping mechanism which enables subcarriers in very strong interference regime to help in recovering interfered signals for subcarriers in strong and weak interference regimes. Depending on the interference and burstiness regime, the inner bounds either employ the proposed helping mechanism to code across subcarriers or treat the subcarriers separately. The outer bounds demonstrate a connection to a subset entropy inequality by Madiman and Tetali.**

***Index Terms*—Interference channel, bursty interference, feedback, and multicarrier systems.**

I. Introduction

Understanding the capacity of interference limited wireless networks is a topic of great practical interest as interference management continues to be a major challenge in achieving higher data rates. In this context, though the capacity characterization of the two user interference channel is still an open question, that of the two user Gaussian interference channel is characterized to within 1 bits/s/Hz [2]. In addition, when there is channel output feedback from the receivers, the capacity region is characterized to within 2 bits/s/Hz [3]. These approximate capacity results are important milestones in understanding interference networks, and provide valuable insight into the design of optimal interference management schemes [1].

In the system model for most works studying the capacity of interference channels (including [2] and [3]), the interfering link is assumed to be always present. However, in practice, the temporal nature of interference in wireless networks tends to be *bursty*. Such burstiness can be attributed to multiple factors: the bursty nature of data traffic, distributed medium access control mechanisms, and decentralized networking protocols. As an example, consider an OFDM based cellular setup (with full frequency reuse), and two users in neighboring cells which are close to the cell boundary. If the subcarrier allocation decisions of neighbouring base stations are uncoordinated, the two users at the cell boundary may be assigned an overlapping set of subcarriers, leading to interference. However, as the allocation decisions for an user change over time, the resulting interference tends to be bursty. In view of the above observations, considering a system model where interference is always present may be pessimistic; it fails to address the possibility of exploiting bursty interference for improvements in the achievable rate.

To exploit the burstiness of interference, [4] and [5] studied a single carrier setup using a degraded message set approach. This approach guarantees a base rate when the carrier faces interference. In addition to the base rate, an incremental rate is provided whenever the carrier is interference free. A multicarrier version of [4] is studied in [12], where the degraded message set approach in [4] is extended to guarantee opportunistic rate increments as the number of interfered subcarriers decreases. However, [4], [5] and [12] do not consider: 1) the possibility of coding across several instantiations of bursty interference, and 2) feedback from receivers. Feedback from receivers is not only a resource which is available in practice, but is also known to provide an unbounded gain in capacity for the (non-bursty) Gaussian interference channel [3]. To study benefits of feedback, [11] considered a single carrier setup with bursty interference and output feedback from the receivers. In [11], bursty interference is modeled using a Bernoulli random state (instantiated i.i.d. over time), and a complete capacity characterization is given for the linear deterministic setup [7]. In addition, schemes in [11] also employ coding across several instantiations of bursty interference. Results in [11] show that, depending on the regime of interference and the level of burstiness, there are significant gains in the capacity compared to the non-bursty setup.

In this paper, we study the multicarrier version of [11], *i.e.*, a setup with output feedback in multicarrier systems with bursty interference. Our motivation stems from the positive results in [11], and the widespread usage of multicarrier systems like OFDM. Since [11] developed optimal single carrier schemes, a natural question arises in the multicarrier version: is it always optimal to treat each subcarrier *separately* and just copy the optimal scheme in [11] on each subcarrier? As the following example illustrates, such a separation may not be always optimal.

Toy example: Consider two parallel symmetric 2-user linear deterministic interference channels (LDICs) [7] as shown

S. Mishra and S. Diggavi are with the Department of Electrical Engineering, University of California, Los Angeles, CA 90095 USA (e-mail: shaunakmishra@ucla.edu, suhas@ee.ucla.edu).

I.-H. Wang is with the Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan (e-mail: ihwang@ntu.edu.tw).

A shorter version of this work was presented at the IEEE International Symposium on Information Theory (ISIT) 2014.

in Figure 1(a) (details of the linear deterministic channel model are described in Section II). The first subcarrier has one direct link ($n_1 = 1$) and one interfering link ($k_1 = 1$, hence $\alpha_1 = \frac{k_1}{n_1} = 1$) and the second subcarrier has one direct link and three interfering links ($\alpha_2 = \frac{k_2}{n_2} = 3$). Causal output feedback is available from the receivers to the respective transmitters. Bernoulli random states $S_1[t]$ and $S_2[t]$ indicate the presence of interference in the first and second subcarrier respectively and are instantiated i.i.d. (over time) from an arbitrary joint distribution $\mathbb{P}_{S_1S_2}$. For this example, we assume the expectation of both the states to be $p = \frac{1}{2}$. Our goal here is to find the maximum achievable symmetric rate. Using the optimal single carrier schemes in [11], we can achieve symmetric rate 0.667 from the first subcarrier and symmetric rate 1.25 from the second subcarrier. Summing these rates, we can achieve a total symmetric rate 1.917. Now, we will show that rate 2.0 is achievable by coding *across* the subcarriers rather than treating the subcarriers separately. We use a block based pipelined scheme (block length N_B) as follows. The transmitters always send fresh symbols in the first subcarrier (a-symbols for Rx_1 and b-symbols for Rx_2 as shown in Figure 1). In the first subcarrier, for sufficiently large N_B, with high probability (w.h.p.) only pN_B a-symbols in a block get interfered at Rx_1 (and pN_B b-symbols at Rx_2). At the end of a block, due to feedback from Rx_1, Tx_1 knows exactly which of its transmitted a-symbols caused interference at Rx_2 (since the same state variable $S_1[t]$ holds for both the receivers). For the next block, Tx_1 creates N_B linear combinations of these pN_B a-symbols (which caused interference at Rx_2 in the previous block) and sends these N_B linear combinations as $c_2[t]$ (in the second subcarrier as shown in Figure 1 (b)) over the next N_B time slots. Due to bursty interference, w.h.p. only pN_B of these linear combinations appear at Rx_2; but this is sufficient to decode pN_B a-symbols constituting the linear combinations. Using these a-symbols Rx_2 can now recover all the interfered b-symbols in the previous block and hence achieve rate 1 from the first subcarrier (same for Rx_1 due to symmetry). For the remaining levels in the second subcarrier, the following is done: lowest levels are not used ($c_3[t] = d_3[t] = 0$), and the transmitters send fresh symbols in the highest level (as shown in Figure 1 (b)) which appear interference free at the receivers (as the lowest levels are not used). This leads to an additional rate 1 from the second subcarrier. Adding rates from the two subcarriers, we achieve symmetric rate 2. This is in fact the symmetric capacity; an easy consequence of the outer bounds developed in this paper.

The above example demonstrates a *helping* mechanism; the second subcarrier *helped* the first subcarrier in recovering interfered symbols in a pipelined fashion. In this paper, we generalize this idea for an arbitrary collection of subcarriers with the following constraint: interference states across subcarriers are drawn from an arbitrary joint distribution (instantiated i.i.d. over time) and the marginal probability of interference is the same for each subcarrier. The main idea behind the generalization is to use specific *levels* in very strongly interfered subcarriers to recover interfered signals for strongly and weakly interfered subcarriers in a pipelined fashion as shown in the toy example. Output feedback was crucial in coding across subcarriers and recovering interfered signals in our toy example. In a similar spirit, [6] leverages delayed channel state information to code across parallel (heterogeneous) broadcast channels resulting in a scheme which outperforms treating the channels separately. Another aspect captured by the toy example is the importance of burstiness; subcarriers in the above example are separable (due to our results and [11]) when interference is always present. Hence, the proposed helping mechanism owes its relevance to bursty interference.

Our main contributions are as follows:

- In the linear deterministic setup, we have a complete capacity region characterization. In the setup with Gaussian noise, we have a tight generalized degrees of freedom (GDoF) characterization and provide outer bounds on the capacity region.
- The inner and outer bounds are non-trivial extensions of single carrier results [11]. We identify regimes where treating subcarriers separately is optimal. For the remaining regimes, we employ coding across subcarriers (helping mechanism) to achieve tight results. The outer bounds involve a subset entropy inequality by Madiman and Tetali [9].

The remainder of this paper is organized as follows. Section II deals with the notation and setup. Section III summarizes the main results of this paper. This is followed by Section IV on inner bounds for the linear deterministic setup, and Section V on the GDoF characterization for the setup with Gaussian noise. Section VI deals with the outer bounds, and we conclude the paper with a discussion in Section VII.

II. Notation and Setup

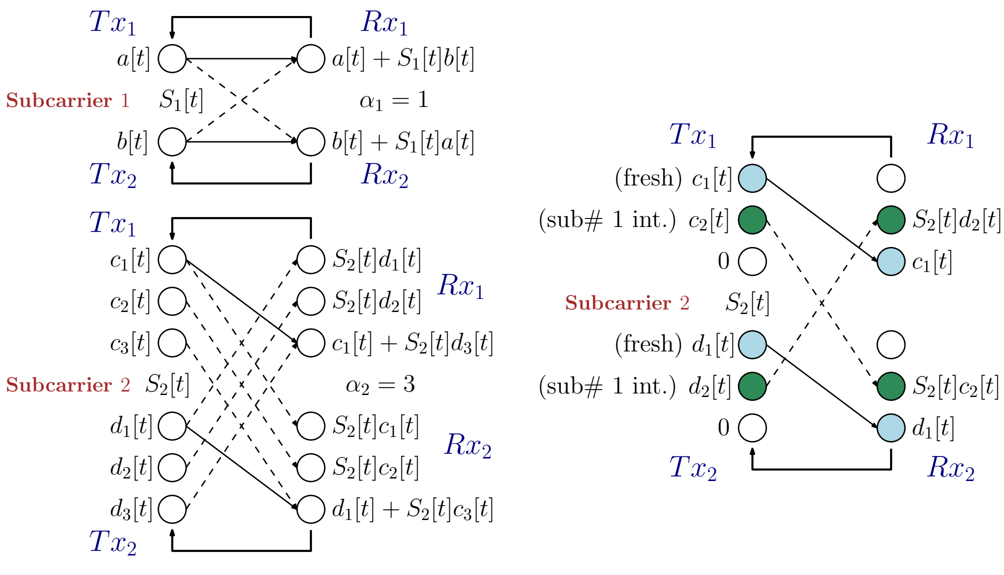
We consider a system with two base stations (transmitters) Tx_1 and Tx_2, and two users (receivers) Rx_1 and Rx_2. For $i \in \{1,2\}$, Tx_i has message $W^{(i)}$ for Rx_i. There are M parallel channels from Tx_i to Rx_i (subcarriers indexed by $j \in \{1,2,\dots M\}$). In this paper, we consider two setups for the subcarrier channel: the first one is based on the linear deterministic model [7], [8] (LD setup[1]), and the second one is based on the Gaussian interference channel (GN setup). The subcarrier channel model for both the setups, followed by the bursty interference model and rate requirements are described below.

Subcarrier channel model: In the LD setup, each subcarrier is modeled by a 2-user (symmetric) LDIC [7], [8] with a bursty interfering link (explained below) and feedback from respective receivers. At discrete time index $t \in \{1,2,\dots N\}$, the transmitted signal in subcarrier j of Tx_i is $\mathbf{x}_j^{(i)}[t] \in \mathbb{F}^{q_j}$ where $\mathbb{F}$ is a finite field. The received signal in subcarrier j at Rx_i is given by:

$$\mathbf{y}_j^{(i)}[t] = \mathbf{G}_j^{q_j-n_j}\mathbf{x}_j^{(i)}[t] + S_j[t]\mathbf{G}_j^{q_j-k_j}\mathbf{x}_j^{(i')}[t], \qquad (1)$$

[1]Unlike its name suggests, the LD setup in our paper is not purely deterministic; there is a stochastic aspect related to the burstiness of interference. However, since the term deterministic in the interference channel literature is usually ascribed to the absence of receiver noise, we stick to this convention.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2697866, IEEE Transactions on Information Theory



(a) (b)

Fig. 1: Toy example with bursty interference in 2 subcarriers: (a) setup with subcarrier 1 ($n_1 = 1$, $k_1 = 1$) and subcarrier 2 ($n_2 = 1$, $k_2 = 3$) and marginal probability of interference in each subcarrier is $p = \frac{1}{2}$, (b) for achieving symmetric capacity, fresh symbols ($c_1[t]$ and $d_1[t]$) are sent in the top most level of both subcarriers, and the middle level is used to recover interfered symbols (in the previous block) of subcarrier 1.

where $\mathbf{G}_j$ is a $q_j \times q_j$ shift matrix in the terminology of deterministic channel models [7], $S_j[t]$ is a Bernoulli random variable (details in bursty interference model below) determining the presence of interference in subcarrier j at time index t, $\mathbf{x}_j^{(i')}[t]$ denotes the transmitted signal on subcarrier j of user $i' \neq i$, and parameters n_j and k_j represent the direct and interfering link strengths [7] in subcarrier j. Figure 2 shows the channel model for subcarrier j in the LD setup. Without loss of generality, we assume $q_j = \max(n_j, k_j)$ and let $\alpha_j = \frac{k_j}{n_j}$ denote the normalized strength of the interfering signal in subcarrier j. For every time instant, it is convenient to consider a subcarrier as indexed levels of bit pipes [7]; each bit pipe carries a symbol from $\mathbb{F}$. Note that, if we assume $S_j[t] = 1$ in (1), then the subcarrier channel model is precisely the usual 2-user LDIC [7], [8] where interference is assumed to be always present.

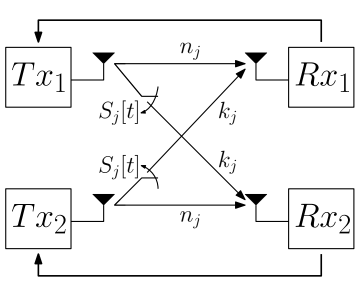


Fig. 2: Bursty interference channel (with feedback) for subcarrier j in LD setup: n_j and k_j represent direct and interfering link strengths. Presence of interference at time index t is determined by Bernoulli random variable $S_j[t]$.

In the GN setup, at discrete time index $t \in \{1, 2, \ldots N\}$, the transmitted signal in subcarrier j of Tx_i is $x_j^{(i)}[t] \in \mathbb{C}$, such that $\frac{1}{N}\sum_{t=1}^{N} |x_j^{(i)}[t]|^2 \leq 1$. The received signal in subcarrier j at Rx_i is given by:

$$y_j^{(i)}[t] = g_{D,j} x_j^{(i)}[t] + S_j[t] g_{I,j} x_j^{(i')}[t] + z_j^{(i)}[t], \tag{2}$$

where $g_{D,j}$, $g_{I,j} \in \mathbb{C}$ denote the direct and interfering channel gains, and $z_j^{(i)}[t] \sim \mathscr{CN}(0,1)$ is Gaussian noise. As in the LD setup, $S_j[t]$ is the interference state. In both LD and GN setups, Tx_i receives causal feedback from Rx_i (feedback consists of the received signal and the interference state).

Bursty interference model: We consider the same interference statistics for both LD and GN setups. As described above, the presence of interference in subcarrier j at time index t is given by a Bernoulli random variable $S_j[t]$ (takes values in $\{0,1\}$). The M Bernoulli random variables $\{S_1[t], S_2[t], \ldots S_M[t]\}$ have a joint probability distribution $\mathbb{P}(S_1[t] = s_1, S_2[t] = s_2, \ldots S_M[t] = s_M) = \mathbb{P}(S_1 = s_1, S_2 = s_2, \ldots S_M = s_M)$ instantiated i.i.d. over time. In this paper, we restrict the analysis to joint distributions with the same marginal probabilities for every $S_j[t]$, *i.e.*, $\forall j$, $\mathbb{E}(S_j[t]) = p$. The transmitters are assumed to know the above statistics, but are limited to causal information on the interference realizations in the subcarriers (through feedback).

Achievable rates: We consider the same rate requirements for both LD and GN setups. Base station Tx_i intends to send message $W^{(i)}$ to Rx_i over N time slots (time index $t \in \{1, 2, \ldots N\}$). Rate $R^{(i)}$ (corresponding to $W^{(i)}$) is considered achievable if the probability of decoding error is vanishingly small as $N \to \infty$.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2697866, IEEE Transactions on Information Theory

III. Main Results

Theorem 1 (LD setup capacity): The capacity region for $\left(R^{(1)}, R^{(2)}\right)$ in the LD setup is given by the following rate inequalities:

$$R^{(i)} \leq p\Delta + \sum_{j=1}^{M} n_j(1+p) - (n_j - k_j)^+ p, \quad (3)$$

$$R^{(i)} + pR^{(i')} \leq p\Delta + \sum_{j=1}^{M} n_j(1+p), \quad (4)$$

$$R^{(i)} + R^{(i')} \leq p\Delta + 2\sum_{j=1}^{M} n_j, \quad (5)$$

where i, $i' \in \{1,2\}$ and $i \neq i'$, and

$$\begin{aligned}\Delta &= \sum_{j=1}^{M} \max(n_j, k_j) + (n_j - k_j)^+ - 2n_j \\ &= \sum_{j:\alpha_j > 2} (k_j - 2n_j) - \sum_{j:\alpha_j \leq 1} k_j - \sum_{j:1<\alpha_j \leq 2} (2n_j - k_j). \quad (6)\end{aligned}$$

Remark 1: The value of Δ, as defined above, plays an important role in our achievability schemes. As we describe later in Section IV on inner bounds, $\Delta > 0$ implies that there are enough levels in subcarriers with $\alpha_j > 2$ (very strong interference) to recover the interfered signals for subcarriers with $\alpha_j \leq 1$ (weak interference) and $1 < \alpha_j \leq 2$ (strong interference). Also, as shown in Figure 3, the shape of the capacity region depends on the value of Δ. The details of the rate tuples $\left(R^{(1)}, R^{(2)}\right)$ marked in Figure 3 are listed below:

- $P_1: \left(p\Delta + \sum_{j=1}^{M} n_j(1+p) - (n_j - k_j)^+ p, 0\right)$
- $Q_1: \left(p\Delta + \sum_{j=1}^{M} n_j(1+p) - (n_j - k_j)^+ p, \sum_{j=1}^{M}(n_j - k_j)^+\right)$
- $D_1: \left(p\Delta + \sum_{j=1}^{M} n_j, \sum_{j=1}^{M} n_j\right)$
- $R_C \equiv (R_C, R_C): \left(\frac{p}{1+p}\Delta + \sum_{j=1}^{M} n_j, \frac{p}{1+p}\Delta + \sum_{j=1}^{M} n_j\right)$
- $R_{NC} \equiv (R_{NC}, R_{NC}): \left(\frac{p}{2}\Delta + \sum_{j=1}^{M} n_j, \frac{p}{2}\Delta + \sum_{j=1}^{M} n_j\right)$
- $D_2: \left(\sum_{j=1}^{M} n_j, p\Delta + \sum_{j=1}^{M} n_j\right)$
- $Q_2: \left(\sum_{j=1}^{M}(n_j - k_j)^+, p\Delta + \sum_{j=1}^{M} n_j(1+p) - (n_j - k_j)^+ p\right)$
- $P_2: \left(0, p\Delta + \sum_{j=1}^{M} n_j(1+p) - (n_j - k_j)^+ p\right)$.

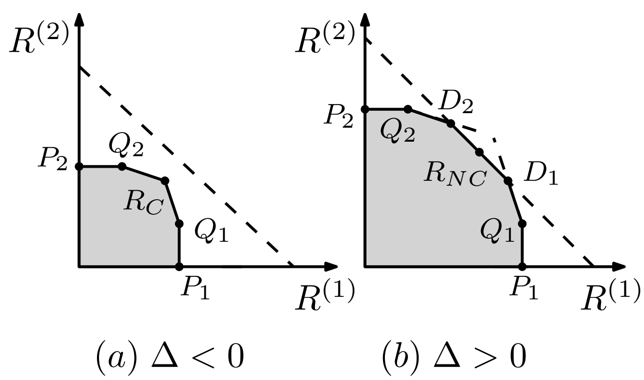


Fig. 3: Capacity region (LD setup) when $\Delta < 0$ and $\Delta > 0$. The dashed line representing inequality (5) is active only when $\Delta > 0$. Symmetric capacity (C_{sym}) for $\Delta < 0$ and $\Delta \geq 0$ is given by $R_C = \frac{p}{1+p}\Delta + \sum_{j=1}^{M} n_j$ and $R_{NC} = \frac{p}{2}\Delta + \sum_{j=1}^{M} n_j$ respectively.

Remark 2: For the case when $\Delta < 0$, the symmetric capacity is given by R_C as shown in Figure 3 (a). The subscript C in R_C stands for causal, and is related to the fact that R_C is derived from outer bounds which consider causal knowledge of bursty interference realizations. In a similar spirit, for symmetric capacity R_{NC} when $\Delta > 0$, the subscript NC stands for non-causal. We describe the proofs for these outer bounds in Section VI.

Corollary 1 (separability in LD setup): In the LD setup, for achieving the capacity region, treating subcarriers separately is optimal when:

1) $0 < p < 1$ and all $\alpha_j \leq 2$,
2) $0 < p < 1$ and all $\alpha_j \geq 2$,
3) $p \in \{1, 0\}$ (degenerate non-bursty case).

For the remaining cases, coding across subcarriers achieves the capacity region.

Theorem 2 (GN setup outer bounds): The following rate inequalities are outer bounds on achievable $(R^{(1)}, R^{(2)})$ in the GN setup:

$$\begin{aligned}R^{(i)} \leq \sum_{j=1}^{M} \Big(&(1-p)\log\left(1 + |g_{D,j}|^2\right) \\ &+ p\log\left(1 + |g_{D,j}|^2 + |g_{I,j}|^2\right)\Big), \quad (7)\end{aligned}$$

$$R^{(i)} + pR^{(i')} \leq p\Delta_G + (1+p)\sum_{j=1}^{M} \log\left(1 + |g_{D,j}|^2\right), \quad (8)$$

$$R^{(i)} + R^{(i')} \leq p\Delta_G + 2\sum_{j=1}^{M} \log\left(1 + |g_{D,j}|^2\right), \quad (9)$$

where i, $i' \in \{1,2\}$ and $i \neq i'$, and $\Delta_G = \sum_{j=1}^{M} \log\left(1 + (|g_{D,j}| + |g_{I,j}|)^2\right) + \log\left(1 + \frac{|g_{D,j}|^2}{1+|g_{I,j}|^2}\right) - 2\log\left(1 + |g_{D,j}|^2\right)$.

Theorem 3 (GN setup GDoF): In the GN setup, assuming $g_{D,j} = \sqrt{SNR}$, $g_{I,j} = \sqrt{INR_j}$ and $INR_j = SNR^{\beta_j}$ (rational β_j),

$$\begin{aligned}GDoF(\beta_1, \dots \beta_M) &= \limsup_{SNR \to \infty} \frac{C_{sym}(SNR, \beta_1, \dots \beta_M)}{M \log(SNR)} \\ &= 1 + \min\left(\frac{\frac{p}{2}\Delta_{GDoF}}{M}, \frac{\frac{p}{1+p}\Delta_{GDoF}}{M}\right), \quad (10)\end{aligned}$$

where C_{sym} denotes the symmetric capacity and $\Delta_{GDoF} = \sum_{j=1}^{M}(\max(1, \beta_j) + (1 - \beta_j)^+ - 2)$.

Corollary 2 (separability in GN setup): Similar to the separability in LD setup (Corollary 1), in the GN setup, treating subcarriers separately is GDoF optimal when all $\beta_j \leq 2$ or all $\beta_j \geq 2$.

IV. Inner bounds: LD setup

In this section, we focus on schemes for achieving the symmetric capacity in the LD setup (see Appendix F and G for achievability of remaining corner points in Figure 3). In Section IV-A, we briefly review the single carrier schemes in [11] and describe a *bursty relaying* technique (used in our multicarrier schemes). In Section IV-B, we mention the cases where treating subcarriers separately is optimal (*i.e.,* simply copying the optimal single carrier scheme [11] on each subcarrier leads to the symmetric capacity). For the remaining cases, we propose multicarrier schemes (covered in Sections IV-C and IV-D), which employ a helping mechanism where some *helper* levels in subcarriers with $\alpha_j > 2$ are used to recover interfered signals in subcarriers with $\alpha_j < 2$. For $\Delta \geq 0$ (Section IV-C), the helping mechanism is optimal; whereas for $\Delta < 0$ (Section IV-D) the helping mechanism is run in parallel with the single carrier schemes [11] to achieve symmetric capacity.

A. Single carrier symmetric capacity [11] and bursty relaying

The single carrier version of our setup (*i.e.,* $M = 1$) was studied in [11]. For notational consistency, we use $j = 1$ (subcarrier index) in stating the results from [11]. We simply restate below the schemes in [11] for the regimes $\alpha_1 \leq 1$ and $1 < \alpha_1 \leq 2$; but for the regime $\alpha_1 > 2$ we mention a slightly different scheme that makes describing our multicarrier schemes in Sections IV-C and IV-D more convenient.

Regime $\alpha_1 \leq 1$*:* For this regime, the symmetric capacity is $n_1 - \frac{p}{1+p}k_1$. To achieve this, a two phase scheme (same for Tx_1 and Tx_2) is used as briefly described below[2] (see [11] for details):

- Phase F: Transmitters in phase F at time index t send fresh symbols on all n_1 levels. If there is no interference at time index t (occurs w.p. $1-p$), all n_1 symbols can be decoded at the intended receiver and both transmitters stay in phase F for time index $t+1$. If there is interference (occurs w.p. p), only the bottom k_1 symbols get interfered at a receiver and the transmitters transition to phase R for time index $t+1$.
- Phase R: Transmitters send the past interference (obtained from receiver feedback) on the top k_1 levels and fresh symbols on the remaining $(n_1 - k_1)$ levels. Both transmitters transition to phase F for the next time index after phase R.

Figure 4 shows the underlying Markov chain for this scheme. Figure 5 shows the scheme for the setup where $n_1 = 3$ and $k_1 = 2$.

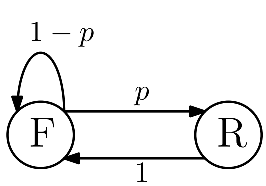


Fig. 4: Underlying Markov chain for the single carrier schemes in [11] for $\alpha_1 \leq 2$.

Regime $1 < \alpha_1 \leq 2$*:* For this regime, the symmetric capacity is $\frac{1-p}{1+p}n_1 + \frac{p}{1+p}k_1 = n_1 - \frac{p}{1+p}(2n_1 - k_1)$. To achieve this, a two phase scheme is used as briefly described below (see [11] for details):

- Phase F: Transmitters in phase F at time index t send fresh symbols on the top n_1 levels and the bottom $k_1 - n_1$ levels are not used. If there is no interference at time index t (occurs w.p. $1-p$), all n_1 symbols can be decoded at the intended receiver and both transmitters stay in phase F. If there is interference at time index t (occurs w.p. p), only $2n_1 - k_1$ symbols get interfered at a receiver and the transmitters transition to phase R for time index $t+1$.
- Phase R: Transmitters send fresh symbols in the top $k_1 - n_1$ levels. In the next $2n_1 - k_1$ levels (below the top $k_1 - n_1$ levels), the $2n_1 - k_1$ interfering symbols (obtained through receiver feedback) from the previous time index are sent. The remaining $k_1 - n_1$ levels in the bottom are not used. Both transmitters transition to phase F for the next time index after phase R.

[2]The scheme for $\alpha_1 = 1$ has slight variation from this scheme. For details, see [11].

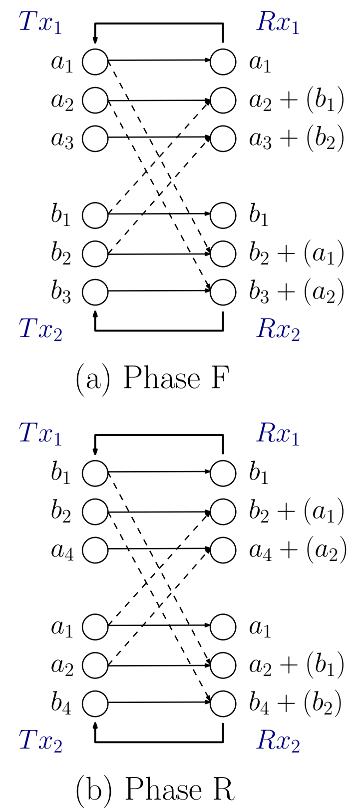


Fig. 5: Single carrier scheme [11] for $n_1 = 3$ and $k_1 = 2$. Symbols in $(\cdot)$ appear only when interference is present.

The underlying Markov chain in this scheme is same as the one in Figure 4. Figure 6 shows the scheme for the setup where $n_1 = 2$ and $k_1 = 3$.

Regime $\alpha_1 > 2$ *(bursty relaying):* For this regime, the symmetric capacity is $n_1 + \frac{p}{2}(k_1 - 2n_1)$. In [11], this is achieved using a Markov chain based scheme similar to the ones described above. To help describe the multicarrier schemes, we develop a block version of the scheme in [11] as follows. In each block of duration N_B, transmitters send fresh symbols on the top n_1 levels and never use the bottom n_1 levels. Since the bottom n_1 levels are never used, the fresh symbols from the top n_1 levels are always received interference free. This realizes rate n_1. From the $k_1 - 2n_1$ levels in the middle (below the top n_1 levels), we realize an additional rate $\frac{p}{2}(k_1 - 2n_1)$ over two blocks as follows. For the first block, Tx_i

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2697866, IEEE Transactions on Information Theory

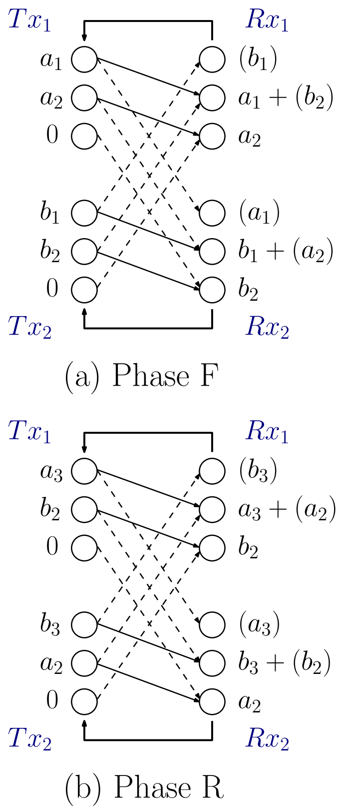


Fig. 6: Single carrier scheme [11] for $n_1 = 2$ and $k_1 = 3$. Symbols in $(\cdot)$ appear only when interference is present.

creates $N_B(k_1 - 2n_1)$ linear combinations from $pN_B(k_1 - 2n_1)$ fresh symbols and sends these linear combinations in the middle $k_1 - 2n_1$ levels. For large enough N_B, w.h.p. $Rx_{i'}$ receives $pN_B(k_1 - 2n_1)$ such linear combinations. $Rx_{i'}$ decodes the constituent fresh symbols from these linear combinations and sends them to $Tx_{i'}$ (through feedback). $Tx_{i'}$ now creates $N_B(k_1 - 2n_1)$ new linear combinations from these symbols and sends them in the $(k_1 - 2n_1)$ middle levels during the next block. W.h.p. Rx_i receives $pN_B(k_1 - 2n_1)$ such linear combinations and decodes all the constituent symbols. This leads to an additive rate of $\frac{pN_B(k_1-2n_1)}{2N_B} = \frac{p}{2}(k_1 - 2n_1)$ at Rx_i (and similarly at $Rx_{i'}$). In the remainder of this paper, we refer to this technique (for middle levels in subcarriers with $\alpha_j > 2$) as *bursty relaying* since Tx_i-Rx_i pair effectively acts as a relay for $Tx_{i'}$-$Rx_{i'}$ and vice versa. Figure 7 illustrates this technique of bursty relaying. Adding the rate from bursty relaying in $(k_1 - 2n_1)$ middle levels and rate n_1 from the top n_1 levels, we achieve rate $n_1 + \frac{p}{2}(k_1 - 2n_1)$.

B. Multicarrier separability

Using outer bounds (4) and (5) for LD setup and achievability rates for the single carrier schemes in [11], the following can be easily verified (see Appendix E for verification details):

- For $p \in \{0, 1\}$, *i.e.,* when interference is either never present or always present, the capacity region can be achieved by treating the subcarriers separately.

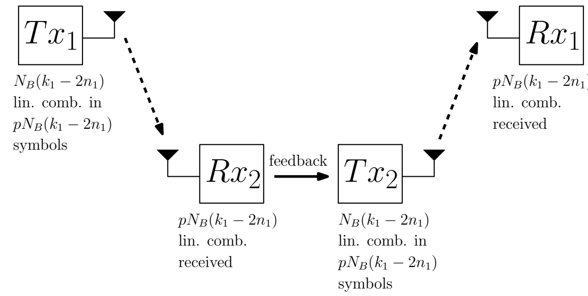


Fig. 7: Bursty relaying using $k_1 - 2n_1$ middle levels (below the top n_1 levels) when $\alpha_1 > 2$. As shown, Rx_2 receives $pN_B(k_1 - 2n_1)$ linear combinations in $pN_B(k_1 - 2n_1)$ symbols during a block of duration N_B. It decodes and sends the constituent symbols to Tx_2 which again creates $N_B(k_1 - 2n_1)$ linear combinations from these symbols. In the next block, Rx_1 receives $pN_B(k_1 - 2n_1)$ linear combinations from Tx_2 and decodes the constituent symbols.

- For $0 < p < 1$, when all subcarriers have $\alpha_j \leq 2$, the capacity region can be achieved by treating the subcarriers separately.
- For $0 < p < 1$, when all subcarriers have $\alpha_j \geq 2$, the capacity region can be achieved by treating the subcarriers separately.

Hence, the subcarriers are *separable* in the above cases. When we have subcarriers with $\alpha_j \leq 2$ as well as subcarriers with $\alpha_j > 2$ (and $0 < p < 1$), we employ coding across subcarriers (through a helping mechanism described in the next subsection) to achieve symmetric capacity; we assume such a *mixed* collection of subcarriers in describing our multicarrier schemes in Sections IV-C and IV-D.

C. Achieving symmetric capacity when $\Delta \geq 0$

As defined in (6),

$$\Delta = \sum_{j:\alpha_j > 2} (k_j - 2n_j) - \sum_{j:\alpha_j \leq 1} k_j - \sum_{j:1<\alpha_j \leq 2} (2n_j - k_j),$$

and when $\Delta \geq 0$, $C_{sym} = R_{NC} = \frac{p}{2}\Delta + \sum_{j=1}^{M} n_j$. We will now describe the achievability of R_{NC} using a block based scheme. In each block of duration N_B, fresh symbols are sent in the following levels (same for both transmitters by symmetry):

- All n_j levels for subcarriers with $\alpha_j \leq 1$.
- Top n_j levels for subcarriers with $\alpha_j > 1$.

In addition, the following levels are not used:

- Bottom $k_j - n_j$ levels of subcarriers with $1 < \alpha_j \leq 2$.
- Bottom n_j levels of subcarriers with $\alpha_j > 2$.

Because of the above choices, in every block (for large enough N_B):

- In subcarriers with $\alpha_j \leq 1$, w.h.p. $pN_B k_j$ fresh symbols get interfered.
- In subcarriers with $1 < \alpha_j \leq 2$, w.h.p. $pN_B(2n_j - k_j)$ fresh symbols get interfered.
- In subcarriers with $\alpha_j > 2$, the top n_j fresh symbols are always received interference free.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2697866, IEEE Transactions on Information Theory

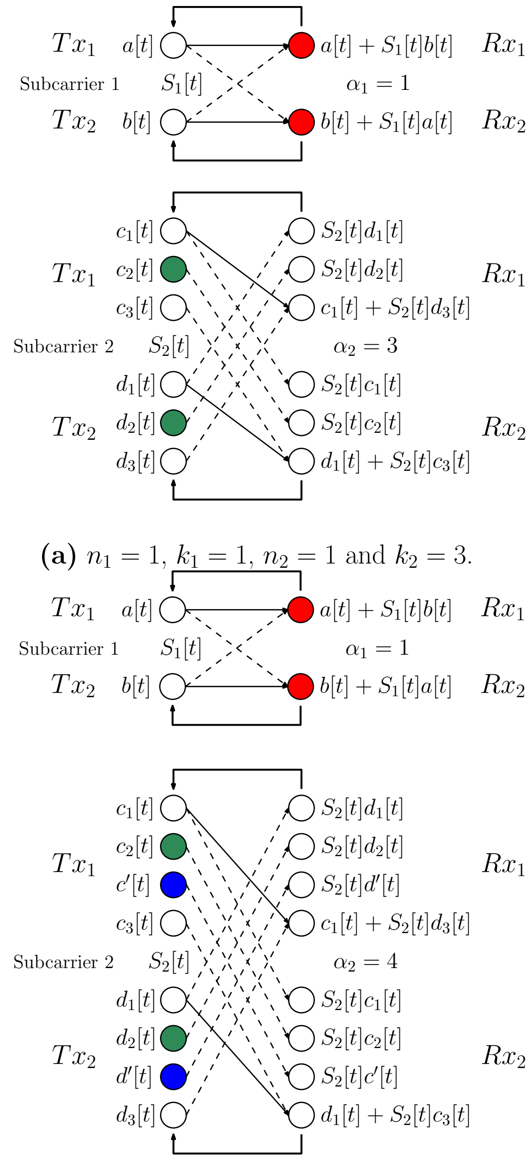
In total, each receiver needs to recover $pN_B(\sum_{j:\alpha_j\le 1} k_j + \sum_{j:1<\alpha_j\le 2} 2n_j - k_j)$ interfered symbols in each block. This recovery is done in a pipelined fashion in the next block using a *helping mechanism* described below.

Helping mechanism: We will use the term *helper levels* for the middle $k_j - 2n_j$ levels (below the top n_j levels) in subcarriers with $\alpha_j > 2$; hence $\sum_{j:\alpha_j>2} k_j - 2n_j$ helper levels in total. After each block, due to feedback from Rx_i, Tx_i knows exactly which of its transmitted symbols caused interference at $Rx_{i'}$. The number of such symbols, as described above, is w.h.p. equal to $pN_B(\sum_{j:\alpha_j\le 1} k_j + \sum_{j:1<\alpha_j\le 2} 2n_j - k_j)$. Tx_i now creates $N_B(\sum_{j:\alpha_j\le 1} k_j + \sum_{j:1<\alpha_j\le 2} 2n_j - k_j)$ linear combinations of these symbols and sends the linear combinations on any $(\sum_{j:\alpha_j\le 1} k_j + \sum_{j:1<\alpha_j\le 2} 2n_j - k_j)$ of the helper levels in the subsequent block. W.h.p. $pN_B(\sum_{j:\alpha_j\le 1} k_j + \sum_{j:1<\alpha_j\le 2} 2n_j - k_j)$ of such linear combinations are received at $Rx_{i'}$. This is sufficient to recover all the interfered symbols at $Rx_{i'}$ in the previous block.

As all the interfered symbols in a block are recovered using the above mechanism, we realize rate $\sum_{j=1}^{M} n_j$. If $\Delta > 0$, some of the helper levels are still available; precisely $(\sum_{j:\alpha_j>2} k_j - 2n_j) - (\sum_{j:\alpha_j\le 1} k_j + \sum_{j:1<\alpha_j\le 2} 2n_j - k_j) = \Delta$ of them. We realize an additional rate of $\frac{p}{2}\Delta$ from such leftover helper levels using the bursty relaying scheme described in Section IV-A. Adding the rate from the leftover helper levels to $\sum_{j=1}^{M} n_j$, we achieve the symmetric capacity $\frac{p}{2}\Delta + \sum_{j=1}^{M} n_j$.

Illustrative examples: The toy example in Section I considered two subcarriers with $n_1 = 1$, $k_1 = 1$, $n_2 = 1$ and $k_2 = 3$ (and $p = \frac{1}{2}$). As illustrated in the toy example, the middle level in the second subcarrier helped in recovering interfered symbols in the first subcarrier. With reference to our achievability scheme for $\Delta \ge 0$, the middle level in the second subcarrier is a helper level (green level in Figure 8 (a)) whereas the (only) level in the first subcarrier is a helped level (red level in Figure 8 (a)). Since there is only one helped level and one helper level, $\Delta = 1 - 1 = 0$ and $C_{sym} = 2$. To illustrate the idea behind our achievability scheme for $\Delta > 0$, we slightly modify the toy example as described below.

Example 1 ($n_1 = 1$, $k_1 = 1$, $n_2 = 1$ and $k_2 = 4$): Compared to the original toy example, we have modified only the second subcarrier such that it has one extra middle level (blue level in Figure 8 (b)). For this case, $\Delta = 2 - 1 = 1$ and $C_{sym} = 2 + \frac{1}{4}$. The helping mechanism is used as in the original toy example to achieve rate 2. Additional rate $\frac{1}{4}$ is achieved using the bursty relaying technique for the extra middle level in the second subcarrier (blue level in Figure 8 (b)).



(a) $n_1 = 1$, $k_1 = 1$, $n_2 = 1$ and $k_2 = 3$.

(b) $n_1 = 1$, $k_1 = 1$, $n_2 = 1$ and $k_2 = 4$.

Fig. 8: Toy example and its modification: (a) original toy example, and (b) Example 1.

D. Achieving symmetric capacity when $\Delta < 0$

When $\Delta < 0$, $C_{sym} = R_C = \frac{p}{1+p}\Delta + \sum_{j=1}^{M} n_j$. Before we proceed to the details, we give a high level idea of the scheme as follows. Simply copying the scheme for $\Delta \ge 0$ in Section IV-C does not work for this case since there are not enough helper levels $(\sum_{j:\alpha_j>2} k_j - 2n_j)$ compared to the number of levels facing interference $(\sum_{j:\alpha_j\le 1} k_j + \sum_{j:1<\alpha_j\le 2} 2n_j - k_j)$. The trick in this case is to *help as much as possible*. For each subcarrier with $\alpha_j < 2$, we select h_j *helped* levels; these levels face interference and the interfered symbols are recovered using the helping mechanism described in Section IV-C. The total number of helped levels $\sum_{j:\alpha_j<2} h_j$ equals the number of helper levels $(\sum_{j:\alpha_j>2} k_j - 2n_j)$. For the remaining interfered levels in subcarriers with $\alpha_j < 2$, we run the optimal single carrier scheme [11] (with slight modifications depending on the value of α_j as discussed later in Sections IV-D1, IV-D2 and IV-D3) in parallel with the helping mechanism. Adding the rates from the single carrier schemes and the helping mechanism, we achieve the symmetric capacity. This high level idea can also be illustrated by rewriting $R_C = \frac{p}{1+p}\Delta + \sum_{j=1}^{M} n_j$ as shown below:

$$\frac{p}{1+p}\Delta + \sum_{j=1}^{M} n_j = \left(\sum_{j:\alpha_j\ge 2} n_j\right) + \left(\sum_{j:\alpha_j<2} h_j\right)$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2697866, IEEE Transactions on Information Theory

$$
\begin{aligned}
&+\left(\sum_{j:\alpha_j\leq 1}(n_j-h_j)-\frac{p}{1+p}(k_j-h_j)\right)\\
&+\left(\sum_{j:1<\alpha_j<2}(n_j-h_j)-\frac{p}{1+p}(2(n_j-h_j)-(k_j-h_j))\right)\\
=&\left(\sum_{j:\alpha_j\geq 2}n_j\right)+\left(\sum_{j:\alpha_j<2}h_j\right)+\left(\sum_{j:\alpha_j\leq 1}\tilde{n}_j-\frac{p}{1+p}\tilde{k}_j\right)\\
&+\left(\sum_{j:1<\alpha_j<2}\tilde{n}_j-\frac{p}{1+p}(2\tilde{n}_j-\tilde{k}_j)\right), \qquad (11)
\end{aligned}
$$

where $\sum_{j:\alpha_j<2}h_j=\sum_{j:\alpha_j>2}k_j-2n_j$ is the total number of helped levels, and for subcarriers (with $\alpha_j<2$) being helped the effective direct and interfering link strengths are $\tilde{n}_j=n_j-h_j$ and $\tilde{k}_j=k_j-h_j$. The last two terms in (11) come from the optimal single carrier schemes for $\alpha_j<2$ (that run in parallel with the helping mechanism).

We now describe the achievability of R_C in detail. In subcarriers with $\alpha_j\geq 2$, the transmitters always send fresh symbols in the top n_j levels and never use the bottom n_j levels. This realizes rate $\sum_{j:\alpha_j\geq 2}n_j$. For each subcarrier with $\alpha_j<2$, we assign a non-negative integral value h_j with the following constraints: (*a*) $h_j\leq k_j$ for $\alpha_j\leq 1$ and $h_j\leq 2n_j-k_j$ for $1<\alpha_j<2$, (*b*) $\sum_{j:\alpha_j<2}h_j=\sum_{j:\alpha_j>2}k_j-2n_j$. Simply put, h_j denotes the number of helped levels in a subcarrier and the total number of such levels equals the number of helper levels available in subcarriers with $\alpha_j>2$. Having fixed h_j for each subcarrier with $\alpha_j<2$, we now describe the modifications needed in the optimal single carrier scheme [11] for parallel execution with the helping mechanism.

1) Modification for $\alpha_j<1$: The bottom h_j levels (of the direct link) are selected as helped levels as shown in Figure 9 (a) and interfered symbols in these levels are recovered using the helping mechanism described in Section IV-C. For the modified single carrier scheme, phase F remains the same as in [11] and the modification is only in Phase R. For illustration purposes consider that in phase F for a subcarrier with $\alpha_j<1$, Tx_1 sends fresh symbols $[a_1\ a_2\ \ldots a_{n_j}]$ (as shown in Figure 9 (a)) and Tx_2 sends fresh symbols $[b_1\ b_2\ \ldots b_{n_j}]$. If there is no interference, all the fresh symbols are received and the transmitters stay in phase F. If there is interference, the transmitters transition to phase R. In the scheme in [11], all k_j interfering symbols were sent on the top k_j levels in phase R; in the modified scheme the transmitters just send the top $\tilde{k}_j=k_j-h_j$ interfering symbols in the top $\tilde{k}_j$ levels as shown in Figure 9 (a). In the remaining levels, fresh symbols are sent (starred symbols in Figure 9 (a)). Ignoring the bottom h_j levels, the resulting system of linear equations at the receivers is exactly the same as in [11] with direct link strength $\tilde{n}_j$ and interfering link strength $\tilde{k}_j$. Thus at end of phase R, Rx_1 is able to decode $\{a_{\tilde{n}_j-\tilde{k}_j+1},a_{\tilde{n}_j-\tilde{k}_j+2},\ldots a_{\tilde{n}_j}\}$ (interfered symbols in phase F) and $\{a^*_{\tilde{n}_j+1},a^*_{\tilde{n}_j+2},\ldots a^*_{2\tilde{n}_j-\tilde{k}_j}\}$ (fresh symbols in phase R). To decode interfered symbols in the helped levels, the helping mechanism is used (which collects all interfered symbols in helped levels during a block of duration N_B and enables their recovery in the next block). So effectively, the rate obtained from a subcarrier with $\alpha_j\leq 1$ is $h_j+(\tilde{n}_j-\frac{p}{1+p}\tilde{k}_j)$.

2) Modification for $\alpha_j=1$: The case $k_j=n_j$ is just an aggregated version of the simple case $k_j=n_j=1$. For this simple case, either $h_j=0$ or $h_j=1$. If $h_j=1$, we use the helping mechanism to recover the interfered symbols. If $h_j=0$, there are no helped levels and we simply use the scheme for $\alpha_j=1$ in [11].

3) Modification for $1<\alpha_j<2$: The top h_j levels (of the direct link at the receiver) are selected as helped levels as shown in Figure 9 (b). Again, phase F remains the same as in [11] and the modification is only for phase R. For illustration purposes, consider that in phase F for a subcarrier with $1<\alpha_j<2$, Tx_1 sends fresh symbols $[a_{\tilde{n}_j+1}\ a_{\tilde{n}_j+2}\ \ldots a_{n_j}\ a_1\ a_2\ldots a_{\tilde{n}_j}]$ on the top n_j levels[3] (as shown in Figure 9 (b)). Similarly, Tx_2 sends fresh symbols $[b_{\tilde{n}_j+1}\ b_{\tilde{n}_j+2}\ \ldots b_{n_j}\ b_1\ b_2\ldots b_{\tilde{n}_j}]$ on the top n_j levels. The bottom k_j-n_j levels are not used. If there is no interference, all the fresh symbols are received and the transmitters stay in phase F. If there is interference, the transmitters transition to phase R. In phase R of the scheme in [11], the bottom k_j-n_j levels were not used and the $2n_j-k_j$ interfering symbols in phase F were sent on the $2n_j-k_j$ levels above the unused levels. In the modified scheme, the transmitters send only $2n_j-k_j-h_j=2\tilde{n}_j-\tilde{k}_j$ interfering symbols (from phase F) on the $2\tilde{n}_j-\tilde{k}_j$ levels above the k_j-n_j unused levels in the bottom. These interfering symbols correspond to the $2\tilde{n}_j-\tilde{k}_j$ levels below the top h_j levels in the direct link at the receiver as shown in Figure 9 (b). In the remaining levels, fresh symbols are sent (starred symbols in Figure 9 (b)). Ignoring the h_j helped levels, the resulting system of linear equations at the receivers is exactly the same as in [11] with direct link strength $\tilde{n}_j$ and interfering link strength $\tilde{k}_j$. Thus at end of phase R, Rx_1 is able to decode $\{a_1,a_2,\ldots a_{2\tilde{n}_j-\tilde{k}_j}\}$ (interfered symbols in phase F) and $\{a^*_{\tilde{n}_j+1},a^*_{\tilde{n}_j+2},\ldots a^*_{\tilde{k}_j}\}$ (fresh symbols in phase R). To decode interfered symbols in the helped levels, the helping mechanism is used (which collects all interfered symbols in helped levels during a block of duration N_B and enables their recovery in the next block). So effectively, the rate obtained from a subcarrier with $1<\alpha_j<2$ is $h_j+(\tilde{n}_j-\frac{p}{1+p}(2\tilde{n}_j-\tilde{k}_j))$.

Taking into account the above modifications and adding the rates across subcarriers we achieve rate R_C. To give an illustrative example of our achievability scheme for $\Delta<0$, we slightly modify the toy example in Section I as described below.

Example 2 ($n_1=2$, $k_1=2$, $n_2=1$ and $k_2=3$): Compared to the original toy example as shown in Figure 10 (a), we have modified only the first subcarrier. For this case, there are two levels in the first subcarrier which face may interference but there is only one helper level (green level in Figure 10 (b)) available in the second subcarrier. Hence $\Delta=1-2=-1$ and $C_{sym}=2+\frac{2}{3}$. We help the bottom level in the first subcarrier (as we did in the original toy example) and by simply copying the scheme in the original toy example we achieve rate 2. For the top level in the first subcarrier (gray level in Figure 10 (b)), we use the optimal single carrier scheme for $\alpha_1=1$ [11]

[3]This particular labeling of the symbols is just for convenience in describing the modification in phase R.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2697866, IEEE Transactions on Information Theory

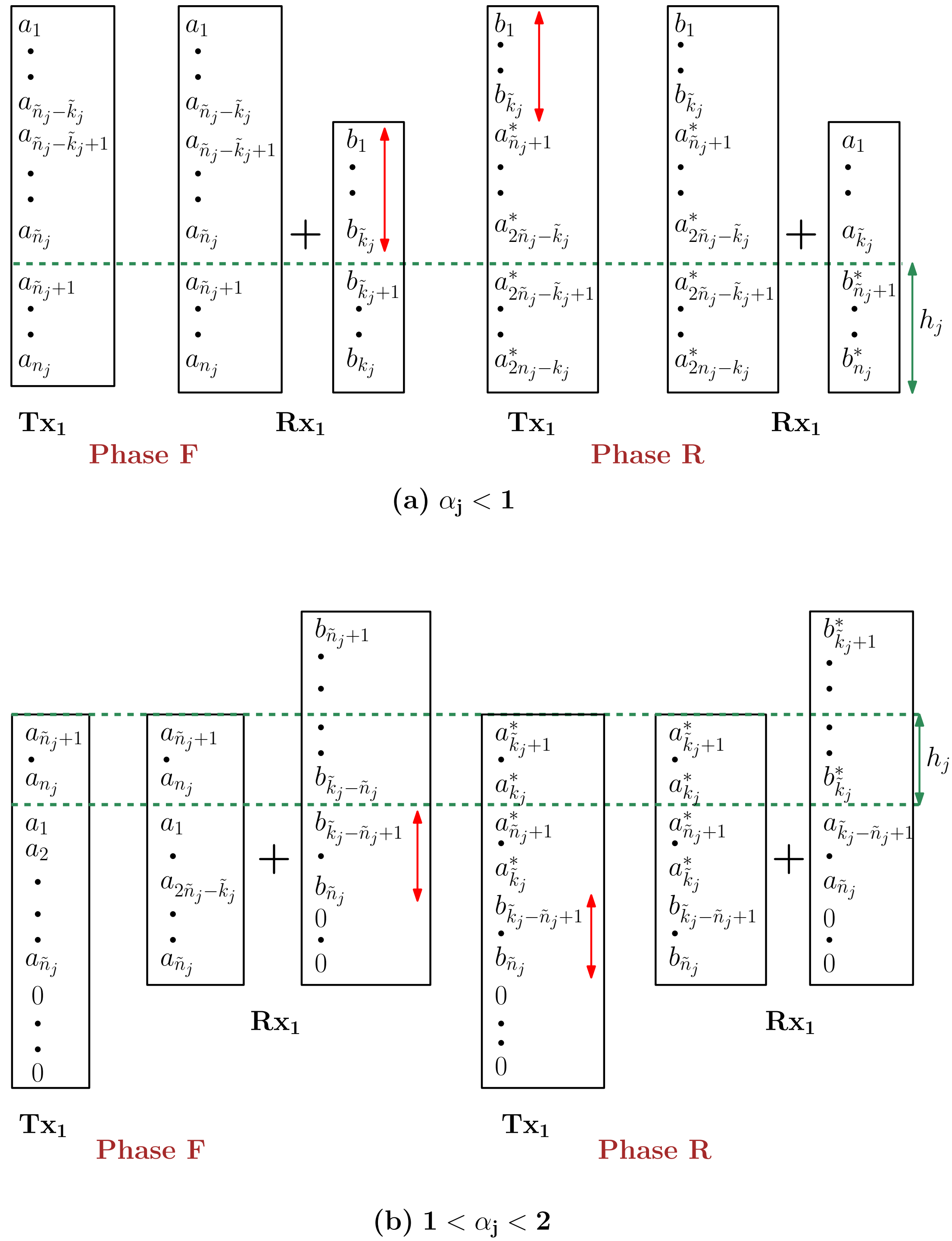


Fig. 9: Modified single carrier schemes for $\alpha_j < 1$ and $1 < \alpha_j < 2$ which run in parallel with the helping mechanism when $\Delta < 0$. Because of h_j helped levels, the effective direct and interfering link strengths are $\tilde{n}_j = n_j - h_j$ and $\tilde{k}_j = k_j - h_j$. The bidirectional red arrows indicate the interfering symbols (from phase F) sent in phase R of the modified scheme.

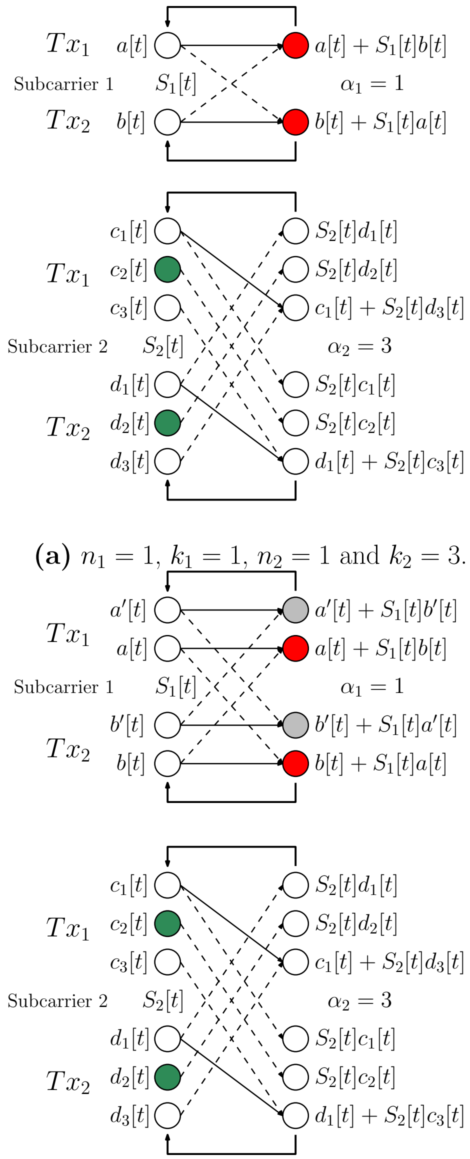
and achieve additional rate $\frac{2}{3}$. In this example, it is easy to see that the helping mechanism and the single carrier scheme can be executed in parallel.

V. GDoF: GN setup

In this section, we describe outer bounds followed by inner bounds (in Sections V-A and V-B) on the GDoF for GN setup. As mentioned in Section III, for the GDoF analysis we assume $g_{D,j} = \sqrt{SNR}$, $g_{I,j} = \sqrt{INR_j}$ and $INR_j = SNR^{\beta_j}$. We assume a rational β_j to simplify the achievability schemes (described in Sections V-A and V-B). With the above assumptions, the GDoF for GN setup is defined as follows:

$$GDoF\left(\beta_1, \beta_2, \ldots \beta_M\right) = \limsup_{SNR \to \infty} \frac{C_{sym}\left(SNR, \beta_1, \beta_2, \ldots \beta_M\right)}{M \log(SNR)},$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2697866, IEEE Transactions on Information Theory



(a) $n_1 = 1$, $k_1 = 1$, $n_2 = 1$ and $k_2 = 3$.

(b) $n_1 = 2$, $k_1 = 2$, $n_2 = 1$ and $k_2 = 3$.

Fig. 10: Toy example and its modification: (a) original toy example, and (b) Example 2.

where C_{sym} is the symmetric capacity. From outer bounds (8) and (9) for the GN setup (proved in Section VI), we have bounds on C_{sym} as follows:

$$
\begin{aligned}
C_{sym} &\le \min\Bigg(\frac{p}{2}\Delta_G + \sum_{j=1}^{M}\log\left(1+|g_{D,j}|^2\right), \\
&\qquad \frac{p}{1+p}\Delta_G + \sum_{j=1}^{M}\log\left(1+|g_{D,j}|^2\right)\Bigg) \\
&= \begin{cases} \frac{p}{2}\Delta_G + \sum_{j=1}^{M}\log\left(1+|g_{D,j}|^2\right) & \text{if } \Delta_G \ge 0, \\ \frac{p}{1+p}\Delta_G + \sum_{j=1}^{M}\log\left(1+|g_{D,j}|^2\right) & \text{if } \Delta_G < 0, \end{cases} \qquad (12)
\end{aligned}
$$

where

$$
\Delta_G = \sum_{j=1}^{M}\log\left(1+(|g_{D,j}|+|g_{I,j}|)^2\right) + \log\left(1+\frac{|g_{D,j}|^2}{1+|g_{I,j}|^2}\right) - 2\log\left(1+|g_{D,j}|^2\right).
$$

Using (12), the following outer bound on GDoF holds:

$$
\begin{aligned}
&GDoF(\beta_1,\ldots,\beta_M) \\
&\le \min\Bigg(\lim_{SNR\to\infty}\frac{\frac{p}{2}\Delta_G + \sum_{j=1}^{M}\log\left(1+|g_{D,j}|^2\right)}{M\log(SNR)}, \\
&\qquad \lim_{SNR\to\infty}\frac{\frac{p}{1+p}\Delta_G + \sum_{j=1}^{M}\log\left(1+|g_{D,j}|^2\right)}{M\log(SNR)}\Bigg) \\
&= \min\left(\frac{\frac{p}{2}\Delta_{GDoF}}{M}+1, \frac{\frac{p}{1+p}\Delta_{GDoF}}{M}+1\right), \qquad (13)
\end{aligned}
$$

where Δ_{GDoF} is derived from Δ_G as shown below:

$$
\begin{aligned}
&\Delta_{GDoF} \\
&= \lim_{SNR\to\infty}\frac{\Delta_G}{\log(SNR)} \\
&= \sum_{j=1}^{M}\left(\max(1,\beta_j) + (1-\beta_j)^+ - 2\right) \\
&= \left(\sum_{j:\beta_j>2}\beta_j - 2\right) - \left(\sum_{j:\beta_j\le 1}\beta_j\right) - \left(\sum_{j:1<\beta_j\le 2} 2-\beta_j\right). \qquad (14)
\end{aligned}
$$

In the remainder of this section, we describe achievability schemes (inner bounds) which achieve outer bound (13). The schemes for the GDoF setting mimic the achievability schemes for symmetric capacity in the LD setup by using techniques from [13]. Hence, the scheme for $\Delta_{GDoF} \ge 0$ (Section V-A) in the GDoF setting mimics the scheme for $\Delta \ge 0$ in LD setup and the scheme for $\Delta_{GDoF} < 0$ (Section V-B) mimics the scheme for $\Delta < 0$ in LD setup.

A. GDoF inner bound when $\Delta_{GDoF} \ge 0$

We use a block based scheme (block size N_B) which mimics the scheme for $\Delta \ge 0$ in Section IV-C for LD setup. For convenience in describing our scheme, we will work with the following *real* channel (the achievable rate for the complex channel in GN setup is just twice the achievable rate for this channel):

$$
y_j^{(i)}[t] = \sqrt{SNR}\, x_j^{(i)}[t] + (S_j[t])\sqrt{INR_j}\, x_j^{(i')}[t] + z_j^{(i)}[t], \qquad (15)
$$

where $x_j^{(i)}[t], x_j^{(i')}[t] \in \mathbb{R}$, $\frac{1}{N}\sum_{t=1}^{N}|x_j^{(i)}[t]|^2 \le 1$ and $z_j^{(i)}[t] \sim \mathcal{N}(0,1)$. Similar to the analysis in [13], we consider

$$
SNR = Q^{2m}, \qquad (16)
$$

where Q and m are positive integers. Furthermore, m is such that $\forall j \in \{1,2,\ldots M\}$, $m\beta_j$ is an integer (always possible since all β_j are rational). By letting m grow to infinity, we get a sequence of SNRs that approach infinity. Using (16), the received signal in (15) can be rewritten as follows:

$$
y_j^{(i)}[t] = Q^m x_j^{(i)}[t] + (S_j[t])\, Q^{m\beta_j} x_j^{(i')}[t] + z_j^{(i)}[t]. \qquad (17)
$$

Following [13], we will express positive real signals in Q-ary representation using Q-ary digits $0,1,\ldots Q-1$ (which we will

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2697866, IEEE Transactions on Information Theory

refer to as "qits", similar to [13]). To mimic the achievability scheme for $\Delta \geq 0$ in LD setup (Section IV-C), we use the following structure for the input signals (we drop the time index for convenience):

- For j with $\beta_j > 2$,
$$x_j^{(i)} = \left[0 \,.\, x_{j,m\beta_j}^{(i)} \, x_{j,m\beta_j-1}^{(i)} \dots x_{j,1}^{(i)}\right]_Q, \quad (18)$$
where $x_{j,1}^{(i)} = x_{j,2}^{(i)} = \ldots = x_{j,m}^{(i)} = 0$ and for the remaining $r \in \{1,2,\ldots m\beta_j\} - \{1,2,\ldots m\}$, $x_{j,r}^{(i)} \in \{1,2,\ldots Q-2\}$.
- For j with $\beta_j \leq 1$,
$$x_j^{(i)} = \left[0 \,.\, x_{j,m}^{(i)} \, x_{j,m-1}^{(i)} \dots x_{j,1}^{(i)}\right]_Q, \quad (19)$$
where $x_{j,r}^{(i)} \in \{1,2,\ldots \lfloor \frac{Q-1}{2} \rfloor - 1\}$ for $r \in \{1,2,\ldots m\}$.
- For j with $1 < \beta_j \leq 2$,
$$x_j^{(i)} = \left[0 \,.\, x_{j,m\beta_j}^{(i)} \, x_{j,m\beta_j-1}^{(i)} \dots x_{j,1}^{(i)}\right]_Q, \quad (20)$$
where $x_{j,1}^{(i)} = x_{j,2}^{(i)} = \ldots = x_{j,m(\beta_j-1)}^{(i)} = 0$ and for the remaining $r \in \{1,2,\ldots m\beta_j\} - \{1,2,\ldots m(\beta_j-1)\}$, $x_{j,r}^{(i)} \in \{1,2,\ldots \lfloor \frac{Q-1}{2} \rfloor - 1\}$.

The structure (*i.e.,* non-zero qits) used is same as in the scheme for LD setup (Section IV-C). The restrictions on the values taken by non-zero qits arises from techniques in [13] (these simplify the analysis by preventing carry overs when signals interfere, see [13] for details). In the absence of noise, it is easy to see the similarities between the LD setup and above setup; qits in a signals are similar to *levels* in the LD setup. The following example makes this similarity more precise for the case of subcarriers with $\beta_j > 2$.

Example 3: In a subcarrier with $\beta_j > 2$, the received signal at Rx_i after interference (in the absence of noise) is as follows:
$$\left[x_{j,m\beta_j}^{(i)} \, x_{j,m\beta_j-1}^{(i)} \dots x_{j,m\beta_j-m+1}^{(i)} \,.\, x_{j,m\beta_j-m}^{(i)} \dots x_{j,1}^{(i)}\right]_Q + \left[x_{j,m\beta_j}^{(i')} \, x_{j,m\beta_j-1}^{(i')} \dots x_{j,m+1}^{(i')} \, 0\,0 \ldots 0 \,.\, 0\,0\right]_Q. \quad (21)$$
Clearly, the top m qits of the direct signal (*i.e.,* $x_{j,m\beta_j}^{(i)} \, x_{j,m\beta_j-1}^{(i)} \dots x_{j,m\beta_j-m+1}^{(i)}$) are interference free in the above scenario and by doing a modulo Q^m operation at the receiver, one can completely recover the direct signal. Even in the presence of noise, due to bounded variance of the noise, the higher qits can be decoded with negligible probability of error (as $m \to \infty$).

Having shown the similarity between LD setup and the above setup in the absence of noise, we now describe the rates that we can achieve from the subcarriers in the GDoF setting.

a) $\beta_j \leq 1$: In this case, over a block only $(pN_B)m\beta_j$ qits in the direct signal are interfered. Assuming we are able to recover all (except $o(m)$) interfering qits (using the helping mechanism described for $\beta_j > 2$ below), we can achieve the following rate:
$$m \log_Q \left(\lfloor \frac{Q-1}{2} \rfloor - 1 \right) + o(m).$$
The above rate follows directly from the analysis in [13].

b) $1 < \beta_j \leq 2$: In this case, over a block only $(pN_B)m(2-\beta_j)$ qits in the direct signal are interfered. Assuming we are able to recover all (except $o(m)$) interfering qits (using the helping mechanism described for $\beta_j > 2$ below), we can achieve the following rate:
$$m \log_Q \left(\lfloor \frac{Q-1}{2} \rfloor - 1 \right) + o(m).$$

c) $\beta_j > 2$: The top m qits in the subcarriers with $\beta_j > 2$ are always received interference free. So from them we can achieve rate:
$$m \log_Q (Q-2) + o(m).$$

We now describe the helping mechanism for the GDoF setting. For removing the interfering qits for subcarriers with $\beta_j < 2$ in the previous block, we need to use $\sum_{j:1<\beta_j\leq 2} m(2-\beta_j) + \sum_{j:\beta_j\leq 1} m\beta_j$ *helper* qits in subcarriers with $\beta_j > 2$; these are the middle $m(\beta_j - 2)$ qits below the top m qits. Since $\Delta_{GDoF} \geq 0$, we have sufficient number of such helper qits to recover all interfering qits in subcarriers with $\beta_j < 2$. The helping mechanism is same as described for the LD setup (with minor changes for the Q-ary setup). From the leftover helper qits, we can achieve an additional rate using the bursty relaying technique. Summing the rates for all subcarriers we have the following inner bound (a factor of $\frac{1}{2}$ is included to account for the complex channel):
$$\begin{aligned} &\frac{1}{2} C_{sym}(SNR, \beta_1, \ldots \beta_M) \\ &\geq \left(m \sum_{j:\beta_j \leq 2} \log_Q \left(\lfloor \frac{Q-1}{2} \rfloor - 1 \right) + o(m) \right) \\ &\quad + \left(m \sum_{j:\beta_j > 2} \log_Q (Q-2) + o(m) \right) \\ &\quad + \left(\frac{p}{2} m \left(\sum_{j:\beta_j > 2} (\beta_j - 2) - \sum_{j:\beta_j \leq 1} \beta_j \right.\right. \\ &\qquad \left.\left. - \sum_{j:1<\beta_j\leq 2} (2-\beta_j) \right) \log_Q (Q-2) + o(m) \right). \end{aligned} \quad (22)$$

Hence,
$$\begin{aligned} &GDoF(\beta_1, \ldots \beta_M) \\ &= \limsup_{m\to\infty} \frac{C_{sym}(SNR, \beta_1, \ldots \beta_M)}{M \log_Q (Q^{2m})} \\ &\overset{(a)}{\geq} 1 + \\ &\frac{\frac{p}{2}\left(\left(\sum_{j:\beta_j>2} \beta_j - 2 \right) - \left(\sum_{j:\beta_j \leq 1} \beta_j \right) - \left(\sum_{j:1<\beta_j\leq 2} 2 - \beta_j \right) \right)}{M} \\ &= \frac{\frac{p}{2}\Delta_{GDoF}}{M} + 1, \end{aligned}$$
where (a) follows from large enough Q. Since the inner bound on GDoF matches the outer bound, we have a tight result when $\Delta_{GDoF} \geq 0$.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2697866, IEEE Transactions on Information Theory

B. GDoF inner bound when $\Delta_{GDoF} < 0$

As in the case of $\Delta_{GDoF} \geq 0$ in Section V-A, we focus on the real channel in (17) for our achievability scheme. The scheme for this case mimics the achievability of symmetric capacity in LD setup for $\Delta < 0$ by using the techniques from [13]. Since we have already illustrated the usage of techniques from [13] (for the case $\Delta_{GDoF} \geq 0$) in mimicking the LD setup schemes for the GDoF setting, we will briefly sketch the inner bound for $\Delta_{GDoF} < 0$.

Following the strategy of *helping as much possible* for the case $\Delta < 0$ in LD setup, we use the middle $m(\beta_j - 2)$ qits (below the top m qits) in subcarriers with $\beta_j > 2$ as helper qits. All the helper qits are used to recover interference in helped qits in subcarriers with $\beta_j < 2$ (each subcarrier with $\beta_j < 2$ has h_j helped qits and $\sum_{j:\beta_j<2} h_j = \sum_{j:\beta_j>2} m(\beta_j - 2)$). So we get the following rates from subcarriers:

- For j with $\beta_j \geq 2$: $m\log_Q(Q-2) + o(m)$.
- For j with $1 < \beta_j < 2$:
$$\left(\left(h_j + \frac{1-p}{1+p}(m-h_j) + \frac{p}{1+p}(m\beta_j - h_j)\right) \times \log_Q\left(\lfloor\frac{Q-1}{2}\rfloor - 1\right)\right) + o(m)$$
- For j with $\beta_j \leq 1$:
$$\left(\left(h_j + (m-h_j) - \frac{p}{1+p}(m\beta_j - h_j)\right) \times \log_Q\left(\lfloor\frac{Q-1}{2}\rfloor - 1\right)\right) + o(m).$$

It should be noted that due to noise, some of the interfering qits in phase F (of the single carrier scheme executed in parallel with the helping mechanism) may not be decoded correctly at Tx_i (after feedback) and this may affect the recovery of qits in phase R. However, it can be shown that such an *error propagation* leads to $o(m)$ reduction (compared to the case without noise) in the achievable rate for a subcarrier. Combining the rates from all subcarriers, we have the following bound (factor of 2 included for the complex channel):

$$\begin{aligned}
&C_{sym}(SNR, \beta_1, \ldots \beta_M) \\
&\geq 2\Bigg(\sum_{j:\beta_j<2} h_j + \sum_{j:\beta_j\geq 2} m + \sum_{j:\beta_j\leq 1} (m-h_j) - \frac{p}{1+p}(m\beta_j - h_j) \\
&\qquad + \sum_{j:1<\beta_j<2} \frac{1-p}{1+p}(m-h_j) + \frac{p}{1+p}(m\beta_j - h_j)\Bigg) \times \\
&\quad \log_Q\left(\lfloor\frac{Q-1}{2}\rfloor - 1\right) + o(m) \\
&\overset{(a)}{=} 2\left(\frac{p}{1+p} m\Delta_{GDoF} + \sum_{j=1}^{M} m\right) \log_Q\left(\lfloor\frac{Q-1}{2}\rfloor - 1\right) + o(m),
\end{aligned} \tag{23}$$

where (a) follows from $\sum_{j:\beta_j<2} h_j = \sum_{j:\beta_j>2} m(\beta_j - 2)$. Now, we have the following bound on the GDoF:

$$\begin{aligned}
&GDoF(\beta_1, \beta_2 \ldots \beta_M) \\
&= \limsup_{m\to\infty} \frac{C_{sym}(SNR, \beta_1, \ldots \beta_M)}{M\log_Q(Q^{2m})} \\
&\geq \lim_{m\to\infty} \frac{\left(\frac{p}{1+p} m\Delta_{GDoF} + \sum_{j=1}^{M} m\right) \log_Q\left(\lfloor\frac{Q-1}{2}\rfloor - 1\right) + o(m)}{mM} \\
&\overset{(a)}{=} \frac{\frac{p}{1+p}\Delta_{GDoF}}{M} + 1,
\end{aligned} \tag{24}$$

where (a) follows from large enough Q. The above inner bound matches outer bound (13) when $\Delta_{GDoF} < 0$ and this completes the GDoF characterization.

VI. Outer bounds: LD and GN setups

In this section, we focus on proofs of outer bounds in the LD and GN setups. We refer to outer bounds (4) and (8) as causal outer bounds as they account for the causal knowledge of subcarrier interference states at the transmitter. For proving these causal outer bounds, we use a subset entropy inequality by Madiman and Tetali which we describe in Section VI-A, prior to the proofs. Then we introduce some additional notation in Section VI-B followed by outer bound proofs for the LD setup (Section VI-C) and GN setup (Section VI-D).

A. Madiman-Tetali subset inequality

We now describe a subset entropy inequality by Madiman and Tetali (Theorem I' in [9]). Consider a hypergraph $(U, \mathscr{E})$ where U is a finite ground set and $\mathscr{E}$ is a collection of subsets of U. A function $\mathscr{G} : \mathscr{E} \to \mathbb{R}^+$ is called a fractional partition of $(U, \mathscr{E})$ if it satisfies the following condition $\forall j \in U$:

$$\sum_{E\in\mathscr{E}:j\in E} \mathscr{G}(E) = 1. \tag{25}$$

With the above definition, the subset entropy inequality can now be stated as follows:

$$\sum_{E\in\mathscr{E}} \mathscr{G}(E) H(X_E) \geq H(X_U), \tag{26}$$

where $\mathscr{G}$ is a fractional partition and the above inequality holds for any collection of jointly distributed random variables X_U. The differential entropy version of the above inequality has the same form [9]. To use these inequalities in our setups, we first choose a suitable fractional partition as explained below. For $\mathbf{s} \in \{0,1\}^M$, let $\mathbf{S}[t] = (S_1[t], S_2[t], \ldots S_M[t]) = \mathbf{s}$ denote the collection of interference states of all the M subcarriers at time index t. As specified in Section II, the occurrence of $\mathbf{S}[t] = \mathbf{s}$ is governed by the joint probability distribution $\mathbb{P}(\mathbf{S}[t] = \mathbf{s})$. To define a fractional partition, we consider the ground set $U = \{1, 2, \ldots M\}$ (*i.e.,* the index set of subcarriers) and view $\mathbf{s} \in \{0,1\}^M$ as a collection of M indicator functions for representing any subset of U. The power set of U (excluding subsets $\mathbf{s}$ such that $\mathbb{P}(\mathbf{S}[t] = \mathbf{s}) = 0$) is chosen as set $\mathscr{E}$. Now, we define a fractional partition $\mathscr{G} : \mathscr{E} \to \mathbb{R}^+$ as follows:

$$\mathscr{G}(E) = \frac{\mathbb{P}(\mathbf{S}[t] = \mathbf{s}_E)}{p}, \tag{27}$$

where $E \in \mathscr{E}$ and $\mathbf{s}_E$ denotes the joint state where only the subcarriers whose index is in set E face interference. The fractional partition condition holds as follows:

$$\sum_{E\in\mathscr{E}:j\in E} \frac{\mathbb{P}(\mathbf{S}[t] = \mathbf{s}_E)}{p} = \frac{\mathbb{E}(S_j[t])}{p} = 1. \tag{28}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2697866, IEEE Transactions on Information Theory

In Section VI-C1, we demonstrate the application of inequality (26), in conjunction with the fractional partition defined in (27), for proving outer bound (4). Similarly, in Section VI-D1, for proving outer bound (8) we use the differential entropy version [9] of inequality (26) with the same fractional partition.

B. Additional notation

For notational convenience, we use indicator functions $\mathbb{I}_{j\notin \mathbf{s}}$ and $\mathbb{I}_{j\in \mathbf{s}}$ to denote the absence and presence of interference in subcarrier j when the joint state realization across M subcarriers is $\mathbf{S}[t]=\mathbf{s}\in\{0,1\}^M$. Also, in the proofs we use $\sum_{\mathbf{s}}$ to denote $\sum_{\mathbf{s}\in\{0,1\}^M}$. The additional notation used for LD setup proofs is listed below in Table I.

The notation used for GN setup proofs is listed below in Table II. Some notation is common to both LD and GN setup proofs, and hence we have some repetitions from Table I in Table II for easier lookup.

C. Outer bounds: LD setup

For the LD setup, outer bounds (3) and (5) are straightforward (multicarrier) extensions of the outer bounds in the single carrier setup [11] (their proof is described in Appendix A and Appendix B respectively). Our main contribution in terms of outer bound techniques lies in proving outer bound (4) (and its corresponding version (8) for the GN setup). We focus on the proof of outer bound (4) in Section VI-C1.

1) Proof of outer bound (4): For proving outer bound (4), we first obtain a bound on $R^{(1)}$ followed by a bound on $R^{(2)}$. Finally, to obtain the bound on $R^{(1)}+pR^{(2)}$, we add the bounds on $R^{(1)}$ and $R^{(2)}$ accordingly, removing intermediate *interference* terms in the process. In particular, to facilitate the removal of such interference terms in the bound for $R^{(1)}+pR^{(2)}$, we use the subset entropy inequality (26) while bounding $R^{(1)}$. In addition, we also account for the causal knowledge of the state sequence $\mathbf{S}_{1:N}$ in the proof. We describe the proof details below.

Using Fano's inequality for Rx_1, for any $\varepsilon>0$, there exists a large enough N such that:

$$
\begin{aligned}
&NR^{(1)}-N\varepsilon\\
&\le I\left(W^{(1)};\mathbf{Y}^{(1)}_{1:N},\mathbf{S}_{1:N}\right)\\
&=\sum_{t=1}^{N} I\left(W^{(1)};\mathbf{Y}^{(1)}[t]\middle|\mathbf{Y}^{(1)}_{1:t-1},\mathbf{S}_{1:t-1},\mathbf{S}[t]\right)\\
&=\sum_{t=1}^{N}\sum_{\mathbf{s}}\mathbb{P}\left(\mathbf{S}[t]=\mathbf{s}\right)H\left(\mathbf{Y}^{(1)}_{\mathbf{s}}[t]\middle|\mathbf{Y}^{(1)}_{1:t-1},\mathbf{S}_{1:t-1},\mathbf{S}[t]=\mathbf{s}\right)\\
&\quad-\sum_{t=1}^{N}\sum_{\mathbf{s}}\mathbb{P}\left(\mathbf{S}[t]=\mathbf{s}\right)\times\\
&\qquad H\left(\mathbf{Y}^{(1)}_{\mathbf{s}}[t]\middle|W^{(1)},\mathbf{Y}^{(1)}_{1:t-1},\mathbf{S}_{1:t-1},\mathbf{S}[t]=\mathbf{s}\right)\\
&\overset{(a)}{\le}\sum_{t=1}^{N}\sum_{\mathbf{s}}\left[\mathbb{P}\left(\mathbf{S}[t]=\mathbf{s}\right)\sum_{j=1}^{M}\left\{n_j\mathbb{I}_{j\notin\mathbf{s}}+\max(n_j,k_j)\mathbb{I}_{j\in\mathbf{s}}\right\}\right]\\
&\quad-\sum_{t=1}^{N}\sum_{\mathbf{s}}\mathbb{P}\left(\mathbf{S}[t]=\mathbf{s}\right)\times
\end{aligned}
$$

TABLE I: Notation used in LD setup proofs.

$\mathbf{S}_{1:t}$	$\triangleq$	$(\mathbf{S}[1],\,\mathbf{S}[2],\,\ldots,\mathbf{S}[t])$
$\mathbf{Y}^{(i)}[t]$	$\triangleq$	$\left(\mathbf{y}^{(i)}_1[t],\,\mathbf{y}^{(i)}_2[t],\ldots,\mathbf{y}^{(i)}_M[t]\right)$, *i.e.*, received signal (across M subcarriers) for Rx_i at time index t
$\mathbf{Y}^{(i)}_{\mathbf{s}}[t]$	$\triangleq$	received signal (across M subcarriers) for Rx_i time t when $\mathbf{S}[t]=\mathbf{s}$; the difference between at $\mathbf{Y}^{(i)}[t]$ and $\mathbf{Y}^{(i)}_{\mathbf{s}}[t]$ is that the state at time t is specified as $\mathbf{s}$ in the latter
$\mathbf{Y}^{(i)}_{1:t}$	$\triangleq$	$\left(\mathbf{Y}^{(i)}[1],\,\mathbf{Y}^{(i)}[2],\,\ldots,\mathbf{Y}^{(i)}[t]\right)$
$\mathbf{V}^{(i)}_{\mathbf{s}}[t]$	$\triangleq$	interfering signals (across M subcarriers) for Rx_i when $\mathbf{S}[t]=\mathbf{s}$
$\mathbf{V}^{(i)}_{1:t}$	$\triangleq$	$\left(\mathbf{V}^{(i)}_{\mathbf{S}[1]}[1],\,\mathbf{V}^{(i)}_{\mathbf{S}[2]}[2],\ldots,\mathbf{V}^{(i)}_{\mathbf{S}[t]}[t]\right)$
$\tilde{\mathbf{V}}^{(i)}[t]$	$\triangleq$	interfering signals (across M subcarriers) at Rx_i when all its subcarriers face interference at time index t; equivalent to $\mathbf{V}^{(i)}_{\mathbf{s}}[t]$ with $\mathbf{s}=\{1,1,\ldots,1\}$
$\hat{\mathbf{X}}^{(i)}[t]$	$\triangleq$	received signal (across M subcarriers) at Rx_i when all its subcarriers are interference free at time index t; equivalent to $\mathbf{Y}^{(i)}_{\mathbf{s}}[t]$ with $\mathbf{s}=\{0,0,\ldots,0\}$

TABLE II: Notation used in GN setup proofs.

$\mathbf{S}_{1:t}$	$\triangleq$	$(\mathbf{S}[1],\,\mathbf{S}[2],\,\ldots,\mathbf{S}[t])$
$\mathbf{Y}^{(i)}[t]$	$\triangleq$	$\left(y^{(i)}_1[t],\,y^{(i)}_2[t],\ldots,y^{(i)}_M[t]\right)$, *i.e.*, received signal (across M subcarriers) for Rx_i at time index t
$\mathbf{Y}^{(i)}_{\mathbf{s}}[t]$	$\triangleq$	received signal (across M subcarriers) for Rx_i at time t when $\mathbf{S}[t]=\mathbf{s}$; the difference between $\mathbf{Y}^{(i)}[t]$ and $\mathbf{Y}^{(i)}_{\mathbf{s}}[t]$ is that the state at time t is specified as $\mathbf{s}$ in the latter
$\mathbf{Y}^{(i)}_{1:t}$	$\triangleq$	$\left(\mathbf{Y}^{(i)}[1],\,\mathbf{Y}^{(i)}[2],\,,\ldots,\mathbf{Y}^{(i)}[t]\right)$
$\mathbf{Z}^{(i)}[t]$	$\triangleq$	$\left(z^{(i)}_1[t],\,z^{(i)}_2[t],\ldots,z^{(i)}_M[t]\right)$, *i.e.*, receiver noise (across M subcarriers) for Rx_i at time index t
$\mathbf{Z}^{(i)}_{\mathbf{s}}[t]$	$\triangleq$	receiver noise in interfered subcarriers for Rx_i at time index t when $\mathbf{S}[t]=\mathbf{s}$
$\mathbf{Z}^{(i)}_{\mathbf{s}^c}[t]$	$\triangleq$	receiver noise in interference free subcarriers for Rx_i at time index t when $\mathbf{S}[t]=\mathbf{s}$
$\mathbf{V}^{(i)}_{\mathbf{s}}[t]\uplus\mathbf{Z}^{(i)}[t]$	$\triangleq$	$\left(S_1[t]g_{I,1}x^{(i')}_1[t]+z^{(i)}_1[t],\,S_2[t]g_{I,2}x^{(i')}_2[t]+z^{(i)}_2[t],\ldots,S_M[t]g_{I,M}x^{(i')}_M[t]+z^{(i)}_M[t]\right)$, *i.e.*, interfering signal (if present) plus noise, across M subcarriers, for Rx_i at

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2697866, IEEE Transactions on Information Theory

	time index t when $\mathbf{S}[t]=\mathbf{s}$
$\mathbf{V}_{\mathbf{s}}^{(i)}[t]\uplus\mathbf{Z}_{\mathbf{s}}^{(i)}[t]\triangleq$	interfering signal plus noise in interfered subcarriers for Rx_i at time index t when $\mathbf{S}[t]=\mathbf{s}$; this does not include the noise terms for subcarriers which do not face interference at time t (unlike $\mathbf{V}_{\mathbf{s}}^{(i)}[t]\uplus\mathbf{Z}^{(i)}[t]$)
$\mathbf{V}_{1:t}^{(i)}\uplus\mathbf{Z}_{1:t}^{(i)}\quad\triangleq$	$\left(\mathbf{V}_{\mathbf{S}[1]}^{(i)}[1]\uplus\mathbf{Z}^{(i)}[1],\ \mathbf{V}_{\mathbf{S}[2]}^{(i)}[2]\uplus\mathbf{Z}^{(i)}[2],\ \ldots,\mathbf{V}_{\mathbf{S}[t]}^{(i)}[t]\uplus\mathbf{Z}^{(i)}[t]\right)$
$\tilde{\mathbf{V}}^{(i)}[t]\uplus\mathbf{Z}^{(i)}[t]\triangleq$	interfering signal plus noise (across M subcarriers) at Rx_i when all its subcarriers face interference at time index t; equivalent to $\mathbf{V}_{\mathbf{s}}^{(i)}[t]\uplus\mathbf{Z}^{(i)}[t]$ with $\mathbf{s}=\{1,1,\ldots,1\}$
$\hat{\mathbf{X}}^{(i)}[t]\uplus\mathbf{Z}^{(i)}[t]\triangleq$	received signal (across M subcarriers) at Rx_i when all its subcarriers are interference free at time index t; equivalent to $\mathbf{Y}_{\mathbf{s}}^{(i)}[t]$ with $\mathbf{s}=\{0,0,\ldots,0\}$
$\hat{\mathbf{X}}_{1:t}^{(i)}\uplus\mathbf{Z}_{1:t}^{(i)}\quad\triangleq$	$\left(\hat{\mathbf{X}}^{(i)}[1]\uplus\mathbf{Z}^{(i)}[1],\ \hat{\mathbf{X}}^{(i)}[2]\uplus\mathbf{Z}^{(i)}[2],\ \ldots,\hat{\mathbf{X}}^{(i)}[t]\uplus\mathbf{Z}^{(i)}[t]\right)$

$$
\begin{aligned}
&H\left(\mathbf{V}_{\mathbf{s}}^{(1)}[t]\middle|W^{(1)},\mathbf{Y}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1},\mathbf{S}[t]=\mathbf{s}\right)\\
&=\sum_{t=1}^{N}\sum_{j=1}^{M}\sum_{\mathbf{s}}\left[n_j\mathbb{P}\left(\mathbf{S}[t]=\mathbf{s}\right)\mathbb{I}_{j\notin\mathbf{s}}+\max(n_j,k_j)\mathbb{P}\left(\mathbf{S}[t]=\mathbf{s}\right)\mathbb{I}_{j\in\mathbf{s}}\right]\\
&\quad-\sum_{t=1}^{N}\sum_{\mathbf{s}}\mathbb{P}\left(\mathbf{S}[t]=\mathbf{s}\right)\times\\
&\quad\quad H\left(\mathbf{V}_{\mathbf{s}}^{(1)}[t]\middle|W^{(1)},\mathbf{Y}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1},\mathbf{S}[t]=\mathbf{s}\right)\\
&\stackrel{(b)}{=}N\sum_{j=1}^{M}(1-p)n_j+p\max(n_j,k_j)\\
&\quad-\sum_{t=1}^{N}\sum_{\mathbf{s}}\mathbb{P}\left(\mathbf{S}[t]=\mathbf{s}\right)\times\\
&\quad\quad H\left(\mathbf{V}_{\mathbf{s}}^{(1)}[t]\middle|W^{(1)},\mathbf{Y}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1},\mathbf{S}[t]=\mathbf{s}\right)\\
&\stackrel{(c)}{=}N\sum_{j=1}^{M}(1-p)n_j+p\max(n_j,k_j)\\
&\quad-\sum_{t=1}^{N}\sum_{\mathbf{s}}\mathbb{P}\left(\mathbf{S}[t]=\mathbf{s}\right)H\left(\mathbf{V}_{\mathbf{s}}^{(1)}[t]\middle|W^{(1)},\mathbf{V}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1}\right)\\
&\stackrel{(d)}{\leq}N\sum_{j=1}^{M}(1-p)n_j+p\max(n_j,k_j)\\
&\quad-p\sum_{t=1}^{N}H\left(\tilde{\mathbf{V}}^{(1)}[t]\middle|W^{(1)},\mathbf{V}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1}\right)\\
&\stackrel{(e)}{=}N\sum_{j=1}^{M}n_j+(\max(n_j,k_j)-n_j)p\\
&\quad-p\sum_{t=1}^{N}H\left(\tilde{\mathbf{V}}^{(1)}[t]\middle|W^{(1)},\mathbf{V}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1}\right),
\end{aligned}\tag{29}
$$

where (a) follows from upper bounding the entropy of output in an interference free subcarrier by n_j (for an interfered subcarrier it is upper bounded by $\max(n_j,k_j)$). In addition, for step (a), we use the fact that $H\left(\mathbf{Y}_{\mathbf{s}}^{(1)}[t]\middle|W^{(1)},\mathbf{Y}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1},\mathbf{S}[t]=\mathbf{s}\right)=H\left(\mathbf{V}_{\mathbf{s}}^{(1)}[t]\middle|W^{(1)},\mathbf{Y}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1},\mathbf{S}[t]=\mathbf{s}\right)$ (*i.e.*, with the given conditioning, the direct signal can be determined and only the interfering signal contributes to the output uncertainty). Step (b) above follows from the fact that for a fixed subcarrier j, $\sum_{\mathbf{s}}\mathbb{P}\left(\mathbf{S}[t]=\mathbf{s}\right)\mathbb{I}_{j\in\mathbf{s}}=p$ (*i.e.*, the marginal probability of interference for subcarrier j), and (c) follows from the observation that the conditioning term $\mathbf{S}[t]=\mathbf{s}$ can be dropped because it is implicit in $\mathbf{V}_{\mathbf{s}}^{(1)}[t]$. To be more precise for step (c), $H\left(\mathbf{V}_{\mathbf{s}}^{(1)}[t]\middle|W^{(1)},\mathbf{Y}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1},\mathbf{S}[t]=\mathbf{s}\right)=H\left(\mathbf{V}_{\mathbf{s}}^{(1)}[t]\middle|W^{(1)},\mathbf{Y}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1}\right)$ because $\mathbf{V}_{\mathbf{s}}^{(1)}[t]$ already assumes $\mathbf{S}[t]=\mathbf{s}$, and $W^{(1)},\mathbf{Y}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1}$ are independent of $\mathbf{S}[t]$. In addition, for step (c) we have used the observation that given $W^{(1)}$, $\mathbf{Y}_{1:t-1}^{(1)}$ and $\mathbf{S}_{1:t-1}$, we can determine the direct signal till time $t-1$ and hence the interfering signal $\mathbf{V}_{1:t-1}^{(1)}$; also, given $W^{(1)}$, $\mathbf{V}_{1:t-1}^{(1)}$ and $\mathbf{S}_{1:t-1}$, we can determine $\mathbf{Y}_{1:t-1}^{(1)}$ (implying $H\left(\mathbf{V}_{\mathbf{s}}^{(1)}[t]\middle|W^{(1)},\mathbf{Y}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1}\right)=H\left(\mathbf{V}_{\mathbf{s}}^{(1)}[t]\middle|W^{(1)},\mathbf{V}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1}\right)$). Step (d) above follows from using the Madiman-Tetali subset inequality (26) for the fractional partition defined in (27). More precisely, in the above context, we can rewrite the Madiman-Tetali subset inequality (26) as

$$
\begin{aligned}
\sum_{\mathbf{s}}\frac{\mathbb{P}\left(\mathbf{S}[t]=\mathbf{s}\right)}{p}&H\left(\mathbf{V}_{\mathbf{s}}^{(1)}[t]\middle|W^{(1)},\mathbf{V}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1}\right)\\
&\geq H\left(\tilde{\mathbf{V}}^{(1)}[t]\middle|W^{(1)},\mathbf{V}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1}\right),
\end{aligned}
$$

where $\frac{\mathbb{P}\left(\mathbf{S}[t]=\mathbf{s}\right)}{p}$ corresponds to the fractional partition defined in (27) and the term involving entropy of subsets, *i.e.*, $\sum_{\mathbf{s}}\frac{\mathbb{P}\left(\mathbf{S}[t]=\mathbf{s}\right)}{p}H\left(\mathbf{V}_{\mathbf{s}}^{(1)}[t]\middle|W^{(1)},\mathbf{V}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1}\right)$ is bounded by the joint entropy term, *i.e.*, $H\left(\tilde{\mathbf{V}}^{(1)}[t]\middle|W^{(1)},\mathbf{V}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1}\right)$. Finally, step (e) simply follows by rewriting $(1-p)n_j+p\max(n_j,k_j)$ as $n_j+(\max(n_j,k_j)-n_j)p$.

Note that step (d) above is a crucial step in the proof since the resulting *interference* term (*i.e.*, the joint entropy term $H\left(\tilde{\mathbf{V}}^{(1)}[t]\middle|W^{(1)},\mathbf{V}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1}\right)$) in the bound for $R^{(1)}$ can be removed from the final outer bound using a similar interference term in the bound for $R^{(2)}$ (described below).

Using Fano's inequality for Rx_2, for any $\varepsilon>0$, there exists a large enough N such that:

$$
\begin{aligned}
&NR^{(2)}-N\varepsilon\\
&\leq I\left(W^{(2)};\mathbf{Y}_{1:N}^{(1)},\mathbf{Y}_{1:N}^{(2)},\mathbf{S}_{1:N},W^{(1)}\right)\\
&=I\left(W^{(2)};\mathbf{Y}_{1:N}^{(1)},\mathbf{Y}_{1:N}^{(2)},\mathbf{S}_{1:N}\middle|W^{(1)}\right)\\
&=\sum_{t=1}^{N}I\left(W^{(2)};\mathbf{Y}^{(2)}[t],\mathbf{Y}^{(1)}[t]\middle|\mathbf{Y}_{1:t-1}^{(2)},\mathbf{Y}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1},W^{(1)},\mathbf{S}[t]\right)\\
&=\sum_{t=1}^{N}\sum_{\mathbf{s}}\mathbb{P}\left(\mathbf{S}[t]=\mathbf{s}\right)\times\\
&\quad\Bigg(H\left(\mathbf{Y}_{\mathbf{s}}^{(2)}[t],\mathbf{Y}_{\mathbf{s}}^{(1)}[t]\middle|\mathbf{Y}_{1:t-1}^{(2)},\mathbf{Y}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1},W^{(1)},\mathbf{S}[t]=\mathbf{s}\right)
\end{aligned}
$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2697866, IEEE Transactions on Information Theory

$$
\begin{aligned}
&-H\Big(\mathbf{Y}_{\mathbf{s}}^{(2)}[t],\mathbf{Y}_{\mathbf{s}}^{(1)}[t]\Big|\mathbf{Y}_{1:t-1}^{(2)},\mathbf{Y}_{1:t-1}^{(1)}, \\
&\qquad\qquad \mathbf{S}_{1:t-1},W^{(1)},W^{(2)},\mathbf{S}[t]=\mathbf{s}\Big)\Big) \\
&\overset{(a)}{=}\sum_{t=1}^{N}\sum_{\mathbf{s}}\mathbb{P}(\mathbf{S}[t]=\mathbf{s})\times \\
&\quad H\left(\mathbf{Y}_{\mathbf{s}}^{(2)}[t],\mathbf{Y}_{\mathbf{s}}^{(1)}[t]\middle|\mathbf{Y}_{1:t-1}^{(2)},\mathbf{Y}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1},W^{(1)},\mathbf{S}[t]=\mathbf{s}\right) \\
&\overset{(b)}{=}\sum_{t=1}^{N}\sum_{\mathbf{s}}\mathbb{P}(\mathbf{S}[t]=\mathbf{s})\times \\
&\quad H\left(\hat{\mathbf{X}}^{(2)}[t],\mathbf{V}_{\mathbf{s}}^{(1)}[t]\middle|\mathbf{Y}_{1:t-1}^{(2)},\mathbf{Y}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1},W^{(1)},\mathbf{S}[t]=\mathbf{s}\right) \\
&\overset{(c)}{\leq}\sum_{t=1}^{N}\sum_{\mathbf{s}}\mathbb{P}(\mathbf{S}[t]=\mathbf{s})\times \\
&\quad H\left(\hat{\mathbf{X}}^{(2)}[t],\tilde{\mathbf{V}}^{(1)}[t]\middle|\mathbf{Y}_{1:t-1}^{(2)},\mathbf{Y}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1},W^{(1)},\mathbf{S}[t]=\mathbf{s}\right) \\
&\overset{(d)}{=}\sum_{t=1}^{N}H\left(\hat{\mathbf{X}}^{(2)}[t],\tilde{\mathbf{V}}^{(1)}[t]\middle|\mathbf{Y}_{1:t-1}^{(2)},\mathbf{Y}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1},W^{(1)}\right)\times \\
&\qquad \sum_{\mathbf{s}}\mathbb{P}(\mathbf{S}[t]=\mathbf{s}) \\
&\overset{(e)}{=}\sum_{t=1}^{N}H\left(\hat{\mathbf{X}}^{(2)}[t],\tilde{\mathbf{V}}^{(1)}[t]\middle|\mathbf{Y}_{1:t-1}^{(2)},\mathbf{Y}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1},W^{(1)}\right) \\
&\overset{(f)}{\leq}\sum_{t=1}^{N}H\left(\hat{\mathbf{X}}^{(2)}[t],\tilde{\mathbf{V}}^{(1)}[t]\middle|\mathbf{Y}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1},W^{(1)}\right) \\
&=\sum_{t=1}^{N}H\left(\hat{\mathbf{X}}^{(2)}[t],\tilde{\mathbf{V}}^{(1)}[t]\middle|\mathbf{V}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1},W^{(1)}\right), \qquad (30)
\end{aligned}
$$

where (a) follows from the observation that $\mathbf{Y}_{\mathbf{s}}^{(2)}[t]$ and $\mathbf{Y}_{\mathbf{s}}^{(1)}[t]$ can be determined given $\mathbf{Y}_{1:t-1}^{(2)}$, $\mathbf{Y}_{1:t-1}^{(1)}$, $\mathbf{S}_{1:t-1}$, $W^{(1)}$, $W^{(2)}$ and $\mathbf{S}[t]=\mathbf{s}$; hence, $H\left(\mathbf{Y}_{\mathbf{s}}^{(2)}[t],\mathbf{Y}_{\mathbf{s}}^{(1)}[t]\middle|\mathbf{Y}_{1:t-1}^{(2)},\mathbf{Y}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1},W^{(1)},W^{(2)},\mathbf{S}[t]=\mathbf{s}\right)=0$. Step (b) follows from the observation that given $\mathbf{Y}_{1:t-1}^{(2)}$, $\mathbf{Y}_{1:t-1}^{(1)}$, $\mathbf{S}_{1:t-1}$, $W^{(1)}$, and $\mathbf{S}[t]=\mathbf{s}$, we can determine the direct signal for Rx_1 and the interfering signal for Rx_2. Hence, the remaining output uncertainty stems from the interfering signal in Rx_1 and the direct signal in Rx_2. Step (c) follows by introducing additional interfering signals for Rx_1, and upper bounding $H\left(\hat{\mathbf{X}}^{(2)}[t],\mathbf{V}_{\mathbf{s}}^{(1)}[t]\middle|\mathbf{Y}_{1:t-1}^{(2)},\mathbf{Y}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1},W^{(1)},\mathbf{S}[t]=\mathbf{s}\right)$ by $H\left(\hat{\mathbf{X}}^{(2)}[t],\tilde{\mathbf{V}}^{(1)}[t]\middle|\mathbf{Y}_{1:t-1}^{(2)},\mathbf{Y}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1},W^{(1)},\mathbf{S}[t]=\mathbf{s}\right)$. Step (d) follows from the observation that $\hat{\mathbf{X}}^{(2)}[t]$ and $\tilde{\mathbf{V}}^{(1)}[t]$ do not depend on $\mathbf{S}[t]$, and step (e) simply follows from $\sum_{\mathbf{s}}\mathbb{P}(\mathbf{S}[t]=\mathbf{s})=1$. Step (f) follows from removing the conditioning on $\mathbf{Y}_{1:t-1}^{(2)}$.

Using inequalities (29) and (30),

$$
\begin{aligned}
&NR^{(1)}-N\varepsilon+pNR^{(2)}-pN\varepsilon \\
&\leq N\sum_{j=1}^{M}n_j+(\max(n_j,k_j)-n_j)\,p \\
&\quad +p\sum_{t=1}^{N}H\left(\hat{\mathbf{X}}^{(2)}[t]\middle|\tilde{\mathbf{V}}^{(1)}[t],W^{(1)},\mathbf{V}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1}\right) \qquad (31) \\
&\leq N\sum_{j=1}^{M}n_j+(\max(n_j,k_j)-n_j)\,p+p\sum_{t=1}^{N}H\left(\hat{\mathbf{X}}^{(2)}[t]\middle|\tilde{\mathbf{V}}^{(1)}[t]\right) \\
&\leq N\sum_{j=1}^{M}n_j+(\max(n_j,k_j)-n_j)\,p+p\sum_{t=1}^{N}\sum_{j=1}^{M}(n_j-k_j)^+ \\
&=Np\Delta+N\sum_{j=1}^{M}(1+p)n_j, \qquad (32)
\end{aligned}
$$

where $\Delta=\sum_{j=1}^{M}\max(n_j,k_j)+(n_j-k_j)^+-2n_j$. As mentioned before, the joint entropy term in (29) is effectively removed using (30) in (31). The bound on $pR^{(1)}+R^{(2)}$ follows by symmetry, and this completes the proof of outer bound (4).

The above proof demonstrates a connection between subset entropy inequalities and bursty interference in multicarrier systems. In a related context, [12] demonstrated a similar connection by using a sliding window subset entropy inequality [10] to show tight outer bounds for the case without feedback (in multicarrier systems with bursty interference). Intuitively, the connection stems from the presence of subsets of interfered subcarriers in the outer bound; this leads to entropies of subsets of interference terms, which is bounded by the joint entropy as demonstrated in the proof of (4).

D. Outer bounds: GN setup

The outer bound proofs for the GN setup resemble the outer bound proofs for the LD setup, with a few modifications required for the presence of additive Gaussian noise. Like in the LD setup, our main contribution in terms of outer bound techniques lies in proving (8). We describe the proof of (8) in Section VI-D1. The proof of outer bounds (7) and (9) is described in Appendix C and Appendix D respectively.

1) Proof of outer bound (8): The structure of the proof for outer bound (8) is similar to that for proving (4). We first obtain a bound on $R^{(1)}$ followed by a bound on $R^{(2)}$. For the bound on $R^{(1)}+pR^{(2)}$, we add the bounds on $R^{(1)}$ and $R^{(2)}$ accordingly, removing intermediate interference terms in the process. To facilitate the removal of intermediate interference terms in the bound for $R^{(1)}+pR^{(2)}$, we use the differential entropy version of the subset entropy inequality (26) while bounding $R^{(1)}$. In addition, we also account for the causal knowledge of the state sequence $\mathbf{S}_{1:N}$ in the proof. The proof details are described below.

Using Fano's inequality for Rx_1, for any $\varepsilon>0$, there exists a large enough N such that:

$$
\begin{aligned}
&NR^{(1)}-N\varepsilon \\
&\leq I\left(W^{(1)};\mathbf{Y}_{1:N}^{(1)},\mathbf{S}_{1:N}\right) \\
&=\sum_{t=1}^{N}I\left(W^{(1)};\mathbf{Y}^{(1)}[t]\middle|\mathbf{Y}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1},\mathbf{S}[t]\right) \\
&=\sum_{t=1}^{N}\sum_{\mathbf{s}}\mathbb{P}(\mathbf{S}[t]=\mathbf{s})\,h\left(\mathbf{Y}^{(1)}[t]\middle|\mathbf{Y}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1},\mathbf{S}[t]=\mathbf{s}\right) \\
&\quad -\sum_{t=1}^{N}\sum_{\mathbf{s}}\mathbb{P}(\mathbf{S}[t]=\mathbf{s})\,h\left(\mathbf{Y}^{(1)}[t]\middle|W^{(1)},\mathbf{Y}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1},\mathbf{S}[t]=\mathbf{s}\right) \\
&\leq\sum_{t=1}^{N}\sum_{\mathbf{s}}\mathbb{P}(\mathbf{S}[t]=\mathbf{s})\,h\left(\mathbf{Y}^{(1)}[t]\middle|\mathbf{S}[t]=\mathbf{s}\right)
\end{aligned}
$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2697866, IEEE Transactions on Information Theory

$$
\begin{aligned}
&-\sum_{t=1}^{N}\sum_{\mathbf{s}}\mathbb{P}\left(\mathbf{S}[t]=\mathbf{s}\right)h\left(\mathbf{Y}^{(1)}[t]\middle|W^{(1)},\mathbf{Y}^{(1)}_{1:t-1},\mathbf{S}_{1:t-1},\mathbf{S}[t]=\mathbf{s}\right)\\
&\overset{(a)}{\leq} N\sum_{j=1}^{M}(1-p)\log\left(1+|g_{D,j}|^2\right)+p\log\left(1+(|g_{D,j}|+|g_{I,j}|)^2\right)\\
&\quad+NM\log(\pi e)\\
&\quad-\sum_{t=1}^{N}\sum_{\mathbf{s}}\mathbb{P}\left(\mathbf{S}[t]=\mathbf{s}\right)h\left(\mathbf{Y}^{(1)}[t]\middle|W^{(1)},\mathbf{Y}^{(1)}_{1:t-1},\mathbf{S}_{1:t-1},\mathbf{S}[t]=\mathbf{s}\right)\\
&\overset{(b)}{=} N\sum_{j=1}^{M}(1-p)\log\left(1+|g_{D,j}|^2\right)+p\log\left(1+(|g_{D,j}|+|g_{I,j}|)^2\right)\\
&\quad+NM\log(\pi e)-\sum_{t=1}^{N}\sum_{\mathbf{s}}\mathbb{P}\left(\mathbf{S}[t]=\mathbf{s}\right)\times\\
&\quad h\left(\mathbf{V}^{(1)}_{\mathbf{s}}[t]\biguplus\mathbf{Z}^{(1)}[t]\middle|W^{(1)},\mathbf{Y}^{(1)}_{1:t-1},\mathbf{S}_{1:t-1},\mathbf{S}[t]=\mathbf{s}\right)\\
&\overset{(c)}{=} N\sum_{j=1}^{M}(1-p)\log\left(1+|g_{D,j}|^2\right)+p\log\left(1+(|g_{D,j}|+|g_{I,j}|)^2\right)\\
&\quad+NM\log(\pi e)-\sum_{t=1}^{N}\sum_{\mathbf{s}}\mathbb{P}\left(\mathbf{S}[t]=\mathbf{s}\right)\times\\
&\quad h\left(\mathbf{V}^{(1)}_{\mathbf{s}}[t]\biguplus\mathbf{Z}^{(1)}_{\mathbf{s}}[t]\middle|W^{(1)},\mathbf{Y}^{(1)}_{1:t-1},\mathbf{S}_{1:t-1},\mathbf{S}[t]=\mathbf{s}\right)\\
&\quad-\sum_{t=1}^{N}\sum_{\mathbf{s}}\mathbb{P}\left(\mathbf{S}[t]=\mathbf{s}\right)\times\\
&\quad h\left(\mathbf{Z}^{(1)}_{\mathbf{s}^c}[t]\middle|\mathbf{V}^{(1)}_{\mathbf{s}}[t]\biguplus\mathbf{Z}^{(1)}_{\mathbf{s}}[t],W^{(1)},\mathbf{Y}^{(1)}_{1:t-1},\mathbf{S}_{1:t-1},\mathbf{S}[t]=\mathbf{s}\right)\\
&\overset{(d)}{=} N\sum_{j=1}^{M}(1-p)\log\left(1+|g_{D,j}|^2\right)+p\log\left(1+(|g_{D,j}|+|g_{I,j}|)^2\right)\\
&\quad+NM\log(\pi e)-\sum_{t=1}^{N}\sum_{\mathbf{s}}\mathbb{P}\left(\mathbf{S}[t]=\mathbf{s}\right)\times\\
&\quad h\left(\mathbf{V}^{(1)}_{\mathbf{s}}[t]\biguplus\mathbf{Z}^{(1)}_{\mathbf{s}}[t]\middle|W^{(1)},\mathbf{Y}^{(1)}_{1:t-1},\mathbf{S}_{1:t-1},\mathbf{S}[t]=\mathbf{s}\right)\\
&\quad-\sum_{t=1}^{N}\sum_{\mathbf{s}}\mathbb{P}\left(\mathbf{S}[t]=\mathbf{s}\right)\sum_{j=1}^{M}\mathbb{I}_{j\notin\mathbf{s}}\log(\pi e)\\
&= N\sum_{j=1}^{M}(1-p)\log\left(1+|g_{D,j}|^2\right)+p\log\left(1+(|g_{D,j}|+|g_{I,j}|)^2\right)\\
&\quad+NM\log(\pi e)-\sum_{t=1}^{N}\sum_{\mathbf{s}}\mathbb{P}\left(\mathbf{S}[t]=\mathbf{s}\right)\times\\
&\quad h\left(\mathbf{V}^{(1)}_{\mathbf{s}}[t]\biguplus\mathbf{Z}^{(1)}_{\mathbf{s}}[t]\middle|W^{(1)},\mathbf{Y}^{(1)}_{1:t-1},\mathbf{S}_{1:t-1}\right)\\
&\quad-NM(1-p)\log(\pi e)\\
&\overset{(e)}{\leq} N\sum_{j=1}^{M}(1-p)\log\left(1+|g_{D,j}|^2\right)+p\log\left(1+(|g_{D,j}|+|g_{I,j}|)^2\right)\\
&\quad+NM\log(\pi e)\\
&\quad-p\sum_{t=1}^{N}h\left(\tilde{\mathbf{V}}^{(1)}[t]\biguplus\mathbf{Z}^{(1)}[t]\middle|W^{(1)},\mathbf{Y}^{(1)}_{1:t-1},\mathbf{S}_{1:t-1}\right)\\
&\quad-NM(1-p)\log(\pi e)\\
&\overset{(f)}{=} N\sum_{j=1}^{M}(1-p)\log\left(1+|g_{D,j}|^2\right)+p\log\left(1+(|g_{D,j}|+|g_{I,j}|)^2\right)\\
&\quad+NM\log(\pi e)\\
&\quad-p\sum_{t=1}^{N}h\left(\tilde{\mathbf{V}}^{(1)}[t]\biguplus\mathbf{Z}^{(1)}[t]\middle|W^{(1)},\mathbf{V}^{(1)}_{1:t-1}\biguplus\mathbf{Z}^{(1)}_{1:t-1},\mathbf{S}_{1:t-1}\right)\\
&\quad-NM(1-p)\log(\pi e),
\end{aligned}\tag{33}
$$

where (a) follows from the proof of (40) (see Appendix C). Step (b) follows from the observation that given $W^{(1)}$, $\mathbf{Y}^{(1)}_{1:t-1}$, $\mathbf{S}_{1:t-1}$ and $\mathbf{S}[t]=\mathbf{s}$, the direct signal in Rx_1 can be determined, and the remaining uncertainty in the output stems from the interfering signal and noise; as defined in Table II the term $\mathbf{V}^{(1)}_{\mathbf{s}}[t]\uplus\mathbf{Z}^{(1)}[t]$ in step (b) denotes the interfering signal (if present) plus the noise across the M subcarriers. Step (c) follows from splitting the term $\mathbf{V}^{(1)}_{\mathbf{s}}[t]\uplus\mathbf{Z}^{(1)}[t]$ on the basis of interfered subcarriers (leading to the term $\mathbf{V}^{(1)}_{\mathbf{s}}[t]\uplus\mathbf{Z}^{(1)}_{\mathbf{s}}[t]$) and interference free subcarriers (leading to the term $\mathbf{Z}^{(1)}_{\mathbf{s}^c}[t]$). Step (d) follows from the fact that the (differential) entropy corresponding to Gaussian noise in a subcarrier is $\log(\pi e)$. Step (e) follows by using the differential entropy version of the Madiman-Tetali subset inequality (26) for the fractional partition defined in (27). More precisely, in the above context, we can rewrite the Madiman-Tetali subset (differential) entropy inequality as

$$
\begin{aligned}
&\sum_{\mathbf{s}}\frac{\mathbb{P}\left(\mathbf{S}[t]=\mathbf{s}\right)}{p}h\left(\mathbf{V}^{(1)}_{\mathbf{s}}[t]\biguplus\mathbf{Z}^{(1)}_{\mathbf{s}}[t]\middle|W^{(1)},\mathbf{Y}^{(1)}_{1:t-1},\mathbf{S}_{1:t-1}\right)\\
&\quad\geq h\left(\tilde{\mathbf{V}}^{(1)}[t]\biguplus\mathbf{Z}^{(1)}[t]\middle|W^{(1)},\mathbf{Y}^{(1)}_{1:t-1},\mathbf{S}_{1:t-1}\right),
\end{aligned}
$$

where $\dfrac{\mathbb{P}\left(\mathbf{S}[t]=\mathbf{s}\right)}{p}$ corresponds to the fractional partition defined in (27) and $h\left(\tilde{\mathbf{V}}^{(1)}[t]\uplus\mathbf{Z}^{(1)}[t]\middle|W^{(1)},\mathbf{Y}^{(1)}_{1:t-1},\mathbf{S}_{1:t-1}\right)$ is the corresponding joint entropy term (we defined $\tilde{\mathbf{V}}^{(1)}[t]\uplus\mathbf{Z}^{(1)}[t]$ in Table II). As shown above, the term involving entropy of subsets, *i.e.*, $\sum_{\mathbf{s}}\frac{\mathbb{P}\left(\mathbf{S}[t]=\mathbf{s}\right)}{p}h\left(\mathbf{V}^{(1)}_{\mathbf{s}}[t]\biguplus\mathbf{Z}^{(1)}_{\mathbf{s}}[t]\middle|W^{(1)},\mathbf{Y}^{(1)}_{1:t-1},\mathbf{S}_{1:t-1}\right)$ is bounded by the corresponding joint entropy term. Finally, step (f) follows from the observation that using $W^{(1)}$, $\mathbf{Y}^{(1)}_{1:t-1}$ and $\mathbf{S}_{1:t-1}$ we can determine the direct signal at Rx_1 till time $t-1$ and determine $\mathbf{V}^{(1)}_{1:t-1}\uplus\mathbf{Z}^{(1)}_{1:t-1}$; at the same time, using $\mathbf{V}^{(1)}_{1:t-1}\uplus\mathbf{Z}^{(1)}_{1:t-1}$, $W^{(1)}$ and $\mathbf{S}_{1:t-1}$ we can determine $\mathbf{Y}^{(1)}_{1:t-1}$.

Note that step (e) above is a crucial step in the proof since the resulting *interference* term (*i.e.*, the joint entropy term) is effectively removed from the final outer bound using the bound on $R^{(2)}$ described below.

Using Fano's inequality for Rx_2, for any $\varepsilon>0$, there exists a large enough N such that:

$$
\begin{aligned}
&NR^{(2)}-N\varepsilon\\
&\leq I\left(W^{(2)};\mathbf{Y}^{(1)}_{1:N},\mathbf{Y}^{(2)}_{1:N},\mathbf{S}_{1:N},W^{(1)}\right)\\
&=I\left(W^{(2)};\mathbf{Y}^{(1)}_{1:N},\mathbf{Y}^{(2)}_{1:N},\mathbf{S}_{1:N}\middle|W^{(1)}\right)\\
&=\sum_{t=1}^{N}I\left(W^{(2)};\mathbf{Y}^{(2)}[t],\mathbf{Y}^{(1)}[t]\middle|\mathbf{Y}^{(2)}_{1:t-1},\mathbf{Y}^{(1)}_{1:t-1},\mathbf{S}_{1:t-1},W^{(1)},\mathbf{S}[t]\right)\\
&=\sum_{t=1}^{N}\sum_{\mathbf{s}}\mathbb{P}\left(\mathbf{S}[t]=\mathbf{s}\right)\times\\
&\quad I\left(W^{(2)};\mathbf{Y}^{(2)}[t],\mathbf{Y}^{(1)}[t]\middle|\mathbf{Y}^{(2)}_{1:t-1},\mathbf{Y}^{(1)}_{1:t-1},\mathbf{S}_{1:t-1},W^{(1)},\mathbf{S}[t]=\mathbf{s}\right)
\end{aligned}
$$

$$\begin{aligned}
&\stackrel{(a)}{=} \sum_{t=1}^{N}\sum_{\mathbf{s}} \mathbb{P}(\mathbf{S}[t]=\mathbf{s}) \times \\
&\quad I\Big(W^{(2)};\hat{\mathbf{X}}^{(2)}[t]\biguplus\mathbf{Z}^{(2)}[t],\mathbf{V}_{\mathbf{s}}^{(1)}[t]\biguplus\mathbf{Z}^{(1)}[t]\Big|\mathbf{Y}_{1:t-1}^{(2)}, \\
&\qquad\qquad \mathbf{Y}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1},W^{(1)},\mathbf{S}[t]=\mathbf{s}\Big) \\
&\stackrel{(b)}{=} \sum_{t=1}^{N}\sum_{\mathbf{s}} \mathbb{P}(\mathbf{S}[t]=\mathbf{s}) \times \\
&\quad I\Big(W^{(2)};\hat{\mathbf{X}}^{(2)}[t]\biguplus\mathbf{Z}^{(2)}[t],\mathbf{V}_{\mathbf{s}}^{(1)}[t]\biguplus\mathbf{Z}_{\mathbf{s}}^{(1)}[t]\Big|\mathbf{Y}_{1:t-1}^{(2)}, \\
&\qquad\qquad \mathbf{Y}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1},W^{(1)},\mathbf{S}[t]=\mathbf{s}\Big) \\
&\quad + \sum_{t=1}^{N}\sum_{\mathbf{s}} \mathbb{P}(\mathbf{S}[t]=\mathbf{s}) \times \\
&\quad I\Big(W^{(2)};\mathbf{Z}_{\mathbf{s}^c}^{(1)}[t]\Big|\hat{\mathbf{X}}^{(2)}[t]\biguplus\mathbf{Z}^{(2)}[t],\mathbf{V}_{\mathbf{s}}^{(1)}[t]\biguplus\mathbf{Z}_{\mathbf{s}}^{(1)}[t], \\
&\qquad\qquad \mathbf{Y}_{1:t-1}^{(2)},\mathbf{Y}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1},W^{(1)},\mathbf{S}[t]=\mathbf{s}\Big) \\
&\stackrel{(c)}{=} \sum_{t=1}^{N}\sum_{\mathbf{s}} \mathbb{P}(\mathbf{S}[t]=\mathbf{s}) \times \\
&\quad I\Big(W^{(2)};\hat{\mathbf{X}}^{(2)}[t]\biguplus\mathbf{Z}^{(2)}[t],\mathbf{V}_{\mathbf{s}}^{(1)}[t]\biguplus\mathbf{Z}_{\mathbf{s}}^{(1)}[t]\Big|\mathbf{Y}_{1:t-1}^{(2)}, \\
&\qquad\qquad \mathbf{Y}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1},W^{(1)},\mathbf{S}[t]=\mathbf{s}\Big) \\
&\stackrel{(d)}{\le} \sum_{t=1}^{N}\sum_{\mathbf{s}} \mathbb{P}(\mathbf{S}[t]=\mathbf{s}) \times \\
&\quad I\Big(W^{(2)};\hat{\mathbf{X}}^{(2)}[t]\biguplus\mathbf{Z}^{(2)}[t],\tilde{\mathbf{V}}^{(1)}[t]\biguplus\mathbf{Z}^{(1)}[t]\Big|\mathbf{Y}_{1:t-1}^{(2)},\mathbf{Y}_{1:t-1}^{(1)}, \\
&\qquad\qquad \mathbf{S}_{1:t-1},W^{(1)},\mathbf{S}[t]=\mathbf{s}\Big) \\
&= \sum_{t=1}^{N}\sum_{\mathbf{s}} \mathbb{P}(\mathbf{S}[t]=\mathbf{s}) \times \\
&\quad h\Big(\hat{\mathbf{X}}^{(2)}[t]\biguplus\mathbf{Z}^{(2)}[t],\tilde{\mathbf{V}}^{(1)}[t]\biguplus\mathbf{Z}^{(1)}[t]\Big|\mathbf{Y}_{1:t-1}^{(2)}, \\
&\qquad\qquad \mathbf{Y}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1},W^{(1)},\mathbf{S}[t]=\mathbf{s}\Big) \\
&\quad - \sum_{t=1}^{N}\sum_{\mathbf{s}} \mathbb{P}(\mathbf{S}[t]=\mathbf{s}) \times \\
&\quad h\Big(\hat{\mathbf{X}}^{(2)}[t]\biguplus\mathbf{Z}^{(2)}[t],\tilde{\mathbf{V}}^{(1)}[t]\biguplus\mathbf{Z}^{(1)}[t]\Big|\mathbf{Y}_{1:t-1}^{(2)},\mathbf{Y}_{1:t-1}^{(1)}, \\
&\qquad\qquad \mathbf{S}_{1:t-1},W^{(1)},W^{(2)},\mathbf{S}[t]=\mathbf{s}\Big) \\
&\stackrel{(e)}{=} \sum_{t=1}^{N}\sum_{\mathbf{s}} \mathbb{P}(\mathbf{S}[t]=\mathbf{s}) \times \\
&\quad h\Big(\hat{\mathbf{X}}^{(2)}[t]\biguplus\mathbf{Z}^{(2)}[t],\tilde{\mathbf{V}}^{(1)}[t]\biguplus\mathbf{Z}^{(1)}[t]\Big|\mathbf{Y}_{1:t-1}^{(2)}, \\
&\qquad\qquad \mathbf{Y}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1},W^{(1)},\mathbf{S}[t]=\mathbf{s}\Big) \\
&\quad - 2NM\log(\pi e) \\
&\stackrel{(f)}{\le} \sum_{t=1}^{N}\sum_{\mathbf{s}} \mathbb{P}(\mathbf{S}[t]=\mathbf{s}) \times \\
&\quad h\Big(\hat{\mathbf{X}}^{(2)}[t]\biguplus\mathbf{Z}^{(2)}[t],\tilde{\mathbf{V}}^{(1)}[t]\biguplus\mathbf{Z}^{(1)}[t]\Big|\mathbf{Y}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1},W^{(1)}\Big) \\
&\quad - 2NM\log(\pi e) \\
&\stackrel{(g)}{=} \sum_{t=1}^{N} h\Big(\hat{\mathbf{X}}^{(2)}[t]\biguplus\mathbf{Z}^{(2)}[t],\tilde{\mathbf{V}}^{(1)}[t]\biguplus\mathbf{Z}^{(1)}[t]\Big|\mathbf{Y}_{1:t-1}^{(1)}, \\
&\qquad\qquad \mathbf{S}_{1:t-1},W^{(1)}\Big) - 2NM\log(\pi e) \\
&= \sum_{t=1}^{N} h\Big(\hat{\mathbf{X}}^{(2)}[t]\biguplus\mathbf{Z}^{(2)}[t],\tilde{\mathbf{V}}^{(1)}[t]\biguplus\mathbf{Z}^{(1)}[t]\Big|\mathbf{V}_{1:t-1}^{(1)}\biguplus\mathbf{Z}_{1:t-1}^{(1)}, \\
&\qquad\qquad \mathbf{S}_{1:t-1},W^{(1)}\Big) - 2NM\log(\pi e), \qquad (34)
\end{aligned}$$

where (a) follows from the observation that given $\mathbf{Y}_{1:t-1}^{(2)}$, $\mathbf{Y}_{1:t-1}^{(1)}$, $\mathbf{S}_{1:t-1}$, $W^{(1)}$ and $\mathbf{S}[t]=\mathbf{s}$, we can determine the direct signal at Rx_1 and the interfering signal at Rx_2. Hence, the remaining uncertainty in the output at Rx_1 corresponds to the interfering signal (if present) plus noise (*i.e.*, the term $\mathbf{V}_{\mathbf{s}}^{(1)}[t]\uplus\mathbf{Z}^{(1)}[t]$), and the remaining uncertainty in the output at Rx_2 corresponds to the direct signal plus noise (*i.e.*, the term $\hat{\mathbf{X}}^{(2)}[t]\uplus\mathbf{Z}^{(2)}[t]$). Step (b) follows from splitting the term $\mathbf{V}_{\mathbf{s}}^{(1)}[t]\uplus\mathbf{Z}^{(1)}[t]$ on basis of interfered subcarriers (leading to the term $\mathbf{V}_{\mathbf{s}}^{(1)}[t]\uplus\mathbf{Z}_{\mathbf{s}}^{(1)}[t]$) and interference free subcarriers (leading to the term $\mathbf{Z}_{\mathbf{s}^c}^{(1)}[t]$). Step (c) follows from the independence of $\mathbf{Z}_{\mathbf{s}^c}^{(1)}[t]$ from $W^{(2)}$ despite the conditioning in step (b). Step (d) follows from introducing extra interference (plus noise) terms in $\mathbf{V}_{\mathbf{s}}^{(1)}[t]\uplus\mathbf{Z}_{\mathbf{s}}^{(1)}[t]$ leading to the term $\tilde{\mathbf{V}}^{(1)}[t]\uplus\mathbf{Z}^{(1)}[t]$. Step (e) follows from the observation that given $\mathbf{Y}_{1:t-1}^{(2)}$, $\mathbf{Y}_{1:t-1}^{(1)}$, $\mathbf{S}_{1:t-1}$, $W^{(1)}$, $W^{(2)}$ and $\mathbf{S}[t]=\mathbf{s}$, the remaining uncertainty in the outputs at Rx_1 and Rx_2 stems from the noise. Step (f) follows by removing the conditioning on $\mathbf{Y}_{1:t-1}^{(2)}$. Finally, step (g) follows from $\sum_{\mathbf{s}}\mathbb{P}(\mathbf{S}[t]=\mathbf{s})=1$.

Using inequalities (33) and (34),

$$\begin{aligned}
&NR^{(1)} - N\varepsilon + pNR^{(2)} - pN\varepsilon \\
&\le N\sum_{j=1}^{M}(1-p)\log\left(1+|g_{D,j}|^2\right) + p\log\Big(1+(|g_{D,j}|+|g_{I,j}|)^2\Big) \\
&\quad + NM\log(\pi e) \\
&\quad - p\sum_{t=1}^{N} h\Big(\tilde{\mathbf{V}}^{(1)}[t]\biguplus\mathbf{Z}^{(1)}[t]\Big|W^{(1)},\mathbf{V}_{1:t-1}^{(1)}\biguplus\mathbf{Z}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1}\Big) \\
&\quad - NM(1-p)\log(\pi e) \\
&\quad + p\sum_{t=1}^{N} h\Big(\hat{\mathbf{X}}^{(2)}[t]\biguplus\mathbf{Z}^{(2)}[t],\tilde{\mathbf{V}}^{(1)}[t]\biguplus\mathbf{Z}^{(1)}[t]\Big| \\
&\qquad\qquad \mathbf{V}_{1:t-1}^{(1)}\biguplus\mathbf{Z}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1},W^{(1)}\Big) - 2pNM\log(\pi e) \\
&= N\sum_{j=1}^{M}(1-p)\log\left(1+|g_{D,j}|^2\right) + p\log\Big(1+(|g_{D,j}|+|g_{I,j}|)^2\Big) \\
&\quad + p\sum_{t=1}^{N} h\Big(\hat{\mathbf{X}}^{(2)}[t]\biguplus\mathbf{Z}^{(2)}[t]\Big|\tilde{\mathbf{V}}^{(1)}[t]\biguplus\mathbf{Z}^{(1)}[t], \\
&\qquad\qquad \mathbf{V}_{1:t-1}^{(1)}\biguplus\mathbf{Z}_{1:t-1}^{(1)},\mathbf{S}_{1:t-1},W^{(1)}\Big) - pNM\log(\pi e) \qquad (35) \\
&\le N\sum_{j=1}^{M}(1-p)\log\left(1+|g_{D,j}|^2\right) + p\log\Big(1+(|g_{D,j}|+|g_{I,j}|)^2\Big) \\
&\quad + p\sum_{t=1}^{N}\sum_{j=1}^{M} h\Big(g_{D,j}x_j^{(2)}[t]+z_j^{(2)}[t]\Big|g_{I,j}x_j^{(2)}[t]+z_j^{(1)}[t]\Big) \\
&\quad - pNM\log(\pi e) \\
&\le N\sum_{j=1}^{M}(1-p)\log\left(1+|g_{D,j}|^2\right) + p\log\Big(1+(|g_{D,j}|+|g_{I,j}|)^2\Big)
\end{aligned}$$

$$
\begin{aligned}
&+p\sum_{t=1}^{N}\sum_{j=1}^{M}\log\left(\pi e\left(1+\frac{|g_{D,j}|^2}{1+|g_{I,j}|^2}\right)\right)-pNM\log(\pi e)\\
&=N\sum_{j=1}^{M}(1-p)\log\left(1+|g_{D,j}|^2\right)+p\log\left(1+(|g_{D,j}|+|g_{I,j}|)^2\right)\\
&\quad+p\log\left(1+\frac{|g_{D,j}|^2}{1+|g_{I,j}|^2}\right)\\
&\overset{(a)}{=}N\left((1+p)\sum_{j=1}^{M}\log\left(1+|g_{D,j}|^2\right)+p\Delta_G\right),
\end{aligned}
\tag{36}
$$

where (a) follows from $\Delta_G = \sum_{j=1}^{M}\log\left(1+(|g_{D,j}|+|g_{I,j}|)^2\right) + \log\left(1+\frac{|g_{D,j}|^2}{1+|g_{I,j}|^2}\right) - 2\log\left(1+|g_{D,j}|^2\right)$. As shown in (35), the joint entropy term in (33) is effectively removed using (34). The bound on $pR^{(1)}+R^{(2)}$ follows by symmetry, and this completes the proof of outer bound (8).

VII. Discussion

In this section, we discuss some aspects related to our system model, their role in the proofs, and directions for future research.

Burstiness model: In our setup, we consider an arbitrary joint distribution governing the interference states across subcarriers but constrain the marginal probabilities to be the same for each subcarrier. This constraint enables us to define a good fractional partition for the Madiman-Tetali subset inequality leading to tight outer bounds. From a practical perspective, our results with this constraint imply that even when the subcarriers experience a similar level of burstiness, some subcarriers can help others to achieve higher rates. In this context, considering the case of different marginal probabilities (of burstiness) for each subcarrier will be a natural extension of our results. Extending the results to non-i.i.d. bursts is another aspect which needs investigation. As an additional remark, the multicarrier system model in this paper can be interpreted as a single carrier setup with transmission blocks of M symbols, and feedback after every block. With such an interpretation, the interference can span multiple symbols in a block during the instantiation of a *burst*; this essentially provides a way to model interference bursts of longer duration in the single carrier setup.

Symmetric interference setup: In this paper, we focus on a symmetric interference channel for each subcarrier. This allows us to build on the capacity results from the single carrier setup [11]. For the case of asymmetric interference, even in the single carrier setup, the capacity characterization is not complete [11]. Hence, for extending the multicarrier results in this paper for asymmetric interference, a better understanding of asymmetric (bursty) interference in the single carrier setup is needed.

Output feedback: In this paper, we assume output feedback at the transmitter. From a practical view point, enabling channel state information (CSI) feedback is significantly cheaper than enabling output feedback. Hence, it will be of practical interest to extend the multicarrier schemes developed in this paper for the case when only CSI feedback is available at the transmitter.

Constant gap result and coding without structure: In this paper, for the setup with Gaussian noise (GN setup), we only give a tight generalized degrees of freedom characterization. Hence, the question of whether we can give a constant gap result for the GN setup remains to be explored. In this direction, using lattice codes along with ideas developed for GDoF achievability in this paper, may be a good candidate for the achievability scheme. A practical drawback of the GDoF achievability scheme in this paper is the use of structured coding for carefully aligning the interfered signal; finding alternative schemes which do not require such structured coding is another direction which remains to be explored.

Acknowledgment

The work of S. Mishra and S. Diggavi was supported in part by NSF grants 1514531 and 1314937. Additionally, they gratefully acknowledge support by Intel. The work of I-Hsiang Wang was supported by the Ministry of Science and Technology, Taiwan, under grants MOST 103-2221-E-002-089-MY2 and MOST 104-2622-8-002-002, and sponsored by MediaTek Inc., Hsin-chu, Taiwan.

Appendix

A. Proof of outer bound (3)

The proof of outer bound (3) is a straightforward (multicarrier) extension of the corresponding proof in the single carrier setup [11]. The proof details for (3) are described below.

Using Fano's inequality for Rx_1, for any $\varepsilon > 0$, there exists a large enough N such that:

$$
\begin{aligned}
&NR^{(1)}-N\varepsilon\\
&\le I\left(W^{(1)};\mathbf{Y}^{(1)}_{1:N},\mathbf{S}_{1:N}\right)\\
&=I\left(W^{(1)};\mathbf{Y}^{(1)}_{1:N}\middle|\mathbf{S}_{1:N}\right)\\
&\le H\left(\mathbf{Y}^{(1)}_{1:N}\middle|\mathbf{S}_{1:N}\right)\\
&\le\sum_{t=1}^{N}H\left(\mathbf{Y}^{(1)}[t]\middle|\mathbf{S}[t]\right)\\
&=\sum_{t=1}^{N}\sum_{\mathbf{s}}\mathbb{P}\left(\mathbf{S}[t]=\mathbf{s}\right)H\left(\mathbf{Y}^{(1)}[t]\middle|\mathbf{S}[t]=\mathbf{s}\right)\\
&\le\sum_{t=1}^{N}\sum_{\mathbf{s}}\mathbb{P}\left(\mathbf{S}[t]=\mathbf{s}\right)\sum_{j=1}^{M}n_j\mathbb{I}_{j\notin\mathbf{s}}+\max(n_j,k_j)\mathbb{I}_{j\in\mathbf{s}}\\
&=N\sum_{j=1}^{M}n_j+p\left(\max(n_j,k_j)-n_j\right)\\
&=Np\Delta+N\sum_{j=1}^{M}n_j(1+p)-(n_j-k_j)^{+}p,
\end{aligned}
\tag{37}
$$

where $\Delta=\sum_{j=1}^{M}\max(n_j,k_j)+(n_j-k_j)^{+}-2n_j$. The outer bound on $R^{(2)}$ follows by symmetry and this completes the proof of outer bound (3).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2697866, IEEE Transactions on Information Theory

B. Proof of outer bound (5)

Like the proof of (3), the proof of outer bound (5) is also a simple (multicarrier) extension of the corresponding sum rate bound in the single carrier setup [11]. The proof details are described below.

Using Fano's inequality for Rx_1 and Rx_2, for any $\varepsilon > 0$, there exists a large enough N such that:

$$
\begin{aligned}
& NR^{(1)} + NR^{(2)} - 2N\varepsilon \\
&\leq I\left(W^{(1)}; \mathbf{Y}^{(1)}_{1:N}, \mathbf{S}_{1:N}\right) + I\left(W^{(2)}; W^{(1)}, \mathbf{Y}^{(1)}_{1:N}, \mathbf{Y}^{(2)}_{1:N}, \mathbf{S}_{1:N}\right) \\
&= I\left(W^{(1)}; \mathbf{Y}^{(1)}_{1:N} \middle| \mathbf{S}_{1:N}\right) + I\left(W^{(2)}; \mathbf{Y}^{(1)}_{1:N}, \mathbf{Y}^{(2)}_{1:N} \middle| \mathbf{S}_{1:N}, W^{(1)}\right) \\
&= H\left(\mathbf{Y}^{(1)}_{1:N} \middle| \mathbf{S}_{1:N}\right) - H\left(\mathbf{Y}^{(1)}_{1:N} \middle| \mathbf{S}_{1:N}, W^{(1)}\right) \\
&\quad + H\left(\mathbf{Y}^{(1)}_{1:N}, \mathbf{Y}^{(2)}_{1:N} \middle| \mathbf{S}_{1:N}, W^{(1)}\right) \\
&= H\left(\mathbf{Y}^{(1)}_{1:N} \middle| \mathbf{S}_{1:N}\right) + H\left(\mathbf{Y}^{(2)}_{1:N} \middle| \mathbf{Y}^{(1)}_{1:N}, \mathbf{S}_{1:N}, W^{(1)}\right) \\
&= H\left(\mathbf{Y}^{(1)}_{1:N} \middle| \mathbf{S}_{1:N}\right) + H\left(\hat{\mathbf{X}}^{(2)}_{1:N} \middle| \mathbf{V}^{(1)}_{1:N}, \mathbf{S}_{1:N}, W^{(1)}\right) \\
&\leq \sum_{t=1}^{N} H\left(\mathbf{Y}^{(1)}[t] \middle| \mathbf{S}[t]\right) + \sum_{t=1}^{N} H\left(\hat{\mathbf{X}}^{(2)}[t] \middle| \mathbf{V}^{(1)}_{\mathbf{S}[t]}[t], \mathbf{S}[t]\right) \\
&= \sum_{t=1}^{N} \sum_{\mathbf{s}} \mathbb{P}(\mathbf{S}[t] = \mathbf{s}) H\left(\mathbf{Y}^{(1)}[t] \middle| \mathbf{S}[t] = \mathbf{s}\right) \\
&\quad + \sum_{t=1}^{N} \sum_{\mathbf{s}} \mathbb{P}(\mathbf{S}[t] = \mathbf{s}) H\left(\hat{\mathbf{X}}^{(2)}[t] \middle| \mathbf{V}^{(1)}_{\mathbf{S}[t]}[t], \mathbf{S}[t] = \mathbf{s}\right) \\
&\leq \sum_{t=1}^{N} \sum_{\mathbf{s}} \mathbb{P}(\mathbf{S}[t] = \mathbf{s}) \sum_{j=1}^{M} n_j \mathbb{I}_{j \notin \mathbf{s}} + \max(n_j, k_j) \mathbb{I}_{j \in \mathbf{s}} \\
&\quad + \sum_{t=1}^{N} \sum_{\mathbf{s}} \mathbb{P}(\mathbf{S}[t] = \mathbf{s}) \sum_{j=1}^{M} n_j \mathbb{I}_{j \notin \mathbf{s}} + (n_j - k_j)^+ \mathbb{I}_{j \in \mathbf{s}} \\
&= \sum_{t=1}^{N} \sum_{j=1}^{M} n_j (1-p) + \max(n_j, k_j) p \\
&\quad + \sum_{t=1}^{N} \sum_{j=1}^{M} n_j (1-p) + (n_j - k_j)^+ p \\
&= Np\Delta + 2N \sum_{j=1}^{M} n_j,
\end{aligned} \tag{38}
$$

where $\Delta = \sum_{j=1}^{M} \max(n_j, k_j) + (n_j - k_j)^+ - 2n_j$. This completes the proof of outer bound (5).

C. Proof of outer bound (7)

Using Fano's inequality for Rx_1, for any $\varepsilon > 0$, there exists a large enough N such that:

$$
\begin{aligned}
& NR^{(1)} - N\varepsilon \\
&\leq I\left(W^{(1)}; \mathbf{Y}^{(1)}_{1:N}, \mathbf{Y}^{(2)}_{1:N}, W^{(2)}, \mathbf{S}_{1:N}\right) \\
&= I\left(W^{(1)}; \mathbf{Y}^{(1)}_{1:N}, \mathbf{Y}^{(2)}_{1:N} \middle| W^{(2)}, \mathbf{S}_{1:N}\right) \\
&= h\left(\mathbf{Y}^{(1)}_{1:N}, \mathbf{Y}^{(2)}_{1:N} \middle| W^{(2)}, \mathbf{S}_{1:N}\right) \\
&\quad - h\left(\mathbf{Y}^{(1)}_{1:N}, \mathbf{Y}^{(2)}_{1:N} \middle| W^{(1)}, W^{(2)}, \mathbf{S}_{1:N}\right) \\
&= h\left(\mathbf{Y}^{(1)}_{1:N}, \mathbf{Y}^{(2)}_{1:N} \middle| W^{(2)}, \mathbf{S}_{1:N}\right) \\
&\quad - \sum_{t=1}^{N} h\left(\mathbf{Y}^{(1)}[t], \mathbf{Y}^{(2)}[t] \middle| \mathbf{Y}^{(2)}_{1:t-1}, \mathbf{Y}^{(1)}_{1:t-1}, W^{(1)}, W^{(2)}, \mathbf{S}_{1:N}\right) \\
&= h\left(\mathbf{Y}^{(1)}_{1:N}, \mathbf{Y}^{(2)}_{1:N} \middle| W^{(2)}, \mathbf{S}_{1:N}\right) \\
&\quad - \sum_{t=1}^{N} h\left(\mathbf{Z}^{(1)}[t], \mathbf{Z}^{(2)}[t] \middle| \mathbf{Y}^{(2)}_{1:t-1}, \mathbf{Y}^{(1)}_{1:t-1}, W^{(1)}, W^{(2)}, \mathbf{S}_{1:N}\right) \\
&= h\left(\mathbf{Y}^{(1)}_{1:N}, \mathbf{Y}^{(2)}_{1:N} \middle| W^{(2)}, \mathbf{S}_{1:N}\right) - \sum_{t=1}^{N} h\left(\mathbf{Z}^{(1)}[t], \mathbf{Z}^{(2)}[t]\right) \\
&= h\left(\mathbf{Y}^{(1)}_{1:N}, \mathbf{Y}^{(2)}_{1:N} \middle| W^{(2)}, \mathbf{S}_{1:N}\right) - 2NM \log(\pi e) \\
&= \sum_{t=1}^{N} h\left(\mathbf{Y}^{(1)}[t], \mathbf{Y}^{(2)}[t] \middle| \mathbf{Y}^{(1)}_{1:t-1}, \mathbf{Y}^{(2)}_{1:t-1}, W^{(2)}, \mathbf{S}_{1:N}\right) \\
&\quad - 2NM \log(\pi e) \\
&\leq \sum_{t=1}^{N} h\left(\mathbf{Y}^{(1)}[t], \mathbf{Y}^{(2)}[t] \middle| \mathbf{Y}^{(1)}_{1:t-1}, \mathbf{Y}^{(2)}_{1:t-1}, W^{(2)}, \mathbf{S}[t]\right) \\
&\quad - 2NM \log(\pi e) \\
&= \Bigg(\sum_{t=1}^{N} \sum_{\mathbf{s}} \mathbb{P}(\mathbf{S}[t] = \mathbf{s}) \times \\
&\quad h\left(\mathbf{Y}^{(1)}[t], \mathbf{Y}^{(2)}[t] \middle| \mathbf{Y}^{(1)}_{1:t-1}, \mathbf{Y}^{(2)}_{1:t-1}, W^{(2)}, \mathbf{S}[t] = \mathbf{s}\right) \Bigg) \\
&\quad - 2NM \log(\pi e) \\
&\leq \sum_{t=1}^{N} \sum_{\mathbf{s}} \mathbb{P}(\mathbf{S}[t] = \mathbf{s}) \times \\
&\quad \Bigg(\sum_{j=1}^{M} \mathbb{I}_{j \notin \mathbf{s}} \left(\log\left(\pi e \left(1 + |g_{D,j}|^2\right)\right) + \log(\pi e)\right) \\
&\quad + \mathbb{I}_{j \in \mathbf{s}} \log\left((\pi e)^2 \left(1 + |g_{D,j}|^2 + |g_{I,j}|^2\right)\right) \Bigg) - 2NM \log(\pi e) \\
&= N \sum_{j=1}^{M} (1-p) \log\left(1 + |g_{D,j}|^2\right) + p \log\left(1 + |g_{D,j}|^2 + |g_{I,j}|^2\right).
\end{aligned} \tag{39}
$$

The bound on $R^{(2)}$ follows by symmetry and this completes the proof of outer bound (7). We also prove a looser bound on $R^{(i)}$ as shown below (the proof for this looser bound is used in the proof of outer bounds (8) and (9)).

Using Fano's inequality for Rx_1, for any $\varepsilon > 0$, there exists a large enough N such that:

$$
\begin{aligned}
& NR^{(1)} - N\varepsilon \\
&\leq I\left(W^{(1)}; \mathbf{Y}^{(1)}_{1:N}, \mathbf{S}_{1:N}\right) \\
&= I\left(W^{(1)}; \mathbf{Y}^{(1)}_{1:N} \middle| \mathbf{S}_{1:N}\right) \\
&= h\left(\mathbf{Y}^{(1)}_{1:N} \middle| \mathbf{S}_{1:N}\right) - h\left(\mathbf{Y}^{(1)}_{1:N} \middle| W^{(1)}, \mathbf{S}_{1:N}\right) \\
&\leq h\left(\mathbf{Y}^{(1)}_{1:N} \middle| \mathbf{S}_{1:N}\right) - h\left(\mathbf{Y}^{(1)}_{1:N} \middle| W^{(2)}, W^{(1)}, \mathbf{S}_{1:N}\right) \\
&= h\left(\mathbf{Y}^{(1)}_{1:N} \middle| \mathbf{S}_{1:N}\right) - \sum_{t=1}^{N} h\left(\mathbf{Y}^{(1)}[t] \middle| \mathbf{Y}^{(1)}_{1:t-1}, W^{(2)}, W^{(1)}, \mathbf{S}_{1:N}\right) \\
&\leq h\left(\mathbf{Y}^{(1)}_{1:N} \middle| \mathbf{S}_{1:N}\right)
\end{aligned}
$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2697866, IEEE Transactions on Information Theory

$$
\begin{aligned}
&-\sum_{t=1}^{N} h(\mathbf{Y}^{(1)}[t]|\mathbf{Y}^{(2)}_{1:t-1},\mathbf{Y}^{(1)}_{1:t-1},W^{(2)},W^{(1)},\mathbf{S}_{1:N})\\
&= h\left(\mathbf{Y}^{(1)}_{1:N}\middle|\mathbf{S}_{1:N}\right)\\
&\quad -\sum_{t=1}^{N} h\left(\mathbf{Z}^{(1)}[t]\middle|\mathbf{Y}^{(2)}_{1:t-1},\mathbf{Y}^{(1)}_{1:t-1},W^{(2)},W^{(1)},\mathbf{S}_{1:N}\right)\\
&= h\left(\mathbf{Y}^{(1)}_{1:N}\middle|\mathbf{S}_{1:N}\right)-\sum_{t=1}^{N} h\left(\mathbf{Z}^{(1)}[t]\right)\\
&= h\left(\mathbf{Y}^{(1)}_{1:N}\middle|\mathbf{S}_{1:N}\right)-NM\log(\pi e)\\
&\leq \sum_{t=1}^{N} h\left(\mathbf{Y}^{(1)}[t]\middle|\mathbf{S}[t]\right)-NM\log(\pi e)\\
&= \sum_{t=1}^{N}\sum_{\mathbf{s}} \mathbb{P}\left(\mathbf{S}[t]=\mathbf{s}\right) h\left(\mathbf{Y}^{(1)}[t]\middle|\mathbf{S}[t]=\mathbf{s}\right)-NM\log(\pi e)\\
&\leq \Bigg(\sum_{t=1}^{N}\sum_{\mathbf{s}} \mathbb{P}\left(\mathbf{S}[t]=\mathbf{s}\right)\sum_{j=1}^{M}\log(\pi e(1+|g_{D,j}|^2))\mathbb{I}_{j\notin\mathbf{s}}\\
&\quad +\log\left(\pi e\left(1+(|g_{D,j}|+|g_{I,j}|)^2\right)\right)\mathbb{I}_{j\in\mathbf{s}}\Bigg)-NM\log(\pi e)\\
&= \Bigg(\sum_{t=1}^{N}\sum_{j=1}^{M}(1-p)\log\left(\pi e\left(1+|g_{D,j}|^2\right)\right)\\
&\quad +p\log\left(\pi e\left(1+(|g_{D,j}|+|g_{I,j}|)^2\right)\right)\Bigg)-NM\log(\pi e)\\
&= N\sum_{j=1}^{M}(1-p)\log\left(1+|g_{D,j}|^2\right)\\
&\quad +p\log\left(1+(|g_{D,j}|+|g_{I,j}|)^2\right). \qquad (40)
\end{aligned}
$$

As mentioned above, (40) is a looser bound compared to (7), but the above proof is used in proving outer bounds (8) and (9).

D. *Proof of outer bound (9)*

Using Fano's inequality for Rx_1 and Rx_2, for any $\varepsilon > 0$, there exists a large enough N such that:

$$
\begin{aligned}
&NR^{(1)}+NR^{(2)}-2N\varepsilon\\
&\leq I\left(W^{(1)};\mathbf{Y}^{(1)}_{1:N},\mathbf{S}_{1:N}\right)+I\left(W^{(2)};W^{(1)},\mathbf{Y}^{(1)}_{1:N},\mathbf{Y}^{(2)}_{1:N},\mathbf{S}_{1:N}\right)\\
&= I\left(W^{(1)};\mathbf{Y}^{(1)}_{1:N}\middle|\mathbf{S}_{1:N}\right)+I\left(W^{(2)};\mathbf{Y}^{(1)}_{1:N},\mathbf{Y}^{(2)}_{1:N}\middle|\mathbf{S}_{1:N},W^{(1)}\right)\\
&= h\left(\mathbf{Y}^{(1)}_{1:N}\middle|\mathbf{S}_{1:N}\right)-h\left(\mathbf{Y}^{(1)}_{1:N}\middle|\mathbf{S}_{1:N},W^{(1)}\right)\\
&\quad +h\left(\mathbf{Y}^{(1)}_{1:N},\mathbf{Y}^{(2)}_{1:N}\middle|\mathbf{S}_{1:N},W^{(1)}\right)\\
&\quad -h\left(\mathbf{Y}^{(1)}_{1:N},\mathbf{Y}^{(2)}_{1:N}\middle|\mathbf{S}_{1:N},W^{(1)},W^{(2)}\right)\\
&= h\left(\mathbf{Y}^{(1)}_{1:N}\middle|\mathbf{S}_{1:N}\right)+h\left(\mathbf{Y}^{(2)}_{1:N}\middle|\mathbf{Y}^{(1)}_{1:N},\mathbf{S}_{1:N},W^{(1)}\right)\\
&\quad -h\left(\mathbf{Y}^{(1)}_{1:N},\mathbf{Y}^{(2)}_{1:N}\middle|\mathbf{S}_{1:N},W^{(1)},W^{(2)}\right)\\
&= h\left(\mathbf{Y}^{(1)}_{1:N}\middle|\mathbf{S}_{1:N}\right)+h\left(\mathbf{Y}^{(2)}_{1:N}\middle|\mathbf{Y}^{(1)}_{1:N},\mathbf{S}_{1:N},W^{(1)}\right)\\
&\quad -\sum_{t=1}^{N} h\left(\mathbf{Y}^{(1)}[t],\mathbf{Y}^{(2)}[t]\middle|\mathbf{Y}^{(1)}_{1:t-1},\mathbf{Y}^{(2)}_{1:t-1},\mathbf{S}_{1:N},W^{(1)},W^{(2)}\right)\\
&= h\left(\mathbf{Y}^{(1)}_{1:N}\middle|\mathbf{S}_{1:N}\right)+h\left(\mathbf{Y}^{(2)}_{1:N}\middle|\mathbf{Y}^{(1)}_{1:N},\mathbf{S}_{1:N},W^{(1)}\right)\\
&\quad -\sum_{t=1}^{N} h\left(\mathbf{Z}^{(1)}[t],\mathbf{Z}^{(2)}[t]\middle|\mathbf{Y}^{(1)}_{1:t-1},\mathbf{Y}^{(2)}_{1:t-1},\mathbf{S}_{1:N},W^{(1)},W^{(2)}\right)\\
&= h\left(\mathbf{Y}^{(1)}_{1:N}\middle|\mathbf{S}_{1:N}\right)+h\left(\mathbf{Y}^{(2)}_{1:N}\middle|\mathbf{Y}^{(1)}_{1:N},\mathbf{S}_{1:N},W^{(1)}\right)\\
&\quad -2NM\log(\pi e)\\
&= h\left(\mathbf{Y}^{(1)}_{1:N}\middle|\mathbf{S}_{1:N}\right)+h\left(\hat{\mathbf{X}}^{(2)}_{1:N}\uplus\mathbf{Z}^{(2)}_{1:N}\middle|\mathbf{Y}^{(1)}_{1:N},\mathbf{S}_{1:N},W^{(1)}\right)\\
&\quad -2NM\log(\pi e)\\
&= h\left(\mathbf{Y}^{(1)}_{1:N}\middle|\mathbf{S}_{1:N}\right)\\
&\quad +h\left(\hat{\mathbf{X}}^{(2)}_{1:N}\uplus\mathbf{Z}^{(2)}_{1:N}\middle|\mathbf{V}^{(1)}_{1:N}\uplus\mathbf{Z}^{(1)}_{1:N},\mathbf{S}_{1:N},W^{(1)}\right)\\
&\quad -2NM\log(\pi e)\\
&\overset{(a)}{\leq} N\sum_{j=1}^{M}(1-p)\log\left(1+|g_{D,j}|^2\right)+p\log\left(1+(|g_{D,j}|+|g_{I,j}|)^2\right)\\
&\quad +h\left(\hat{\mathbf{X}}^{(2)}_{1:N}\uplus\mathbf{Z}^{(2)}_{1:N}\middle|\mathbf{V}^{(1)}_{1:N}\uplus\mathbf{Z}^{(1)}_{1:N},\mathbf{S}_{1:N},W^{(1)}\right)-NM\log(\pi e)\\
&\leq N\sum_{j=1}^{M}(1-p)\log\left(1+|g_{D,j}|^2\right)+p\log\left(1+(|g_{D,j}|+|g_{I,j}|)^2\right)\\
&\quad +\sum_{t=1}^{N}\sum_{\mathbf{s}}\mathbb{P}\left(\mathbf{S}[t]=\mathbf{s}\right)\times\\
&\quad h\left(\hat{\mathbf{X}}^{(2)}[t]\uplus\mathbf{Z}^{(2)}[t]\middle|\mathbf{V}^{(1)}_{\mathbf{s}}[t]\uplus\mathbf{Z}^{(1)}[t],W^{(1)},\mathbf{S}[t]=\mathbf{s}\right)\\
&\quad -NM\log(\pi e)\\
&\leq N\sum_{j=1}^{M}(1-p)\log\left(1+|g_{D,j}|^2\right)+p\log\left(1+(|g_{D,j}|+|g_{I,j}|)^2\right)\\
&\quad +\sum_{t=1}^{N}\sum_{\mathbf{s}}\mathbb{P}\left(\mathbf{S}[t]=\mathbf{s}\right)\sum_{j=1}^{M}\log(\pi e(1+|g_{D,j}|^2))\mathbb{I}_{j\notin\mathbf{s}}\\
&\quad +\log\left(\pi e\left(1+\frac{|g_{D,j}|^2}{1+|g_{I,j}|^2}\right)\right)\mathbb{I}_{j\in\mathbf{s}}-NM\log(\pi e)\\
&= N\sum_{j=1}^{M}(1-p)\log\left(1+|g_{D,j}|^2\right)+p\log\left(1+(|g_{D,j}|+|g_{I,j}|)^2\right)\\
&\quad +N\sum_{j=1}^{M}(1-p)\log\left(1+|g_{D,j}|^2\right)+p\log\left(1+\frac{|g_{D,j}|^2}{1+|g_{I,j}|^2}\right)\\
&= N\left(2\sum_{j=1}^{M}\log\left(1+|g_{D,j}|^2\right)+p\Delta_G\right) \qquad (41)
\end{aligned}
$$

where (a) follows from the proof of (40) (see Appendix C) and $\Delta_G = \sum_{j=1}^{M}\log\left(1+(|g_{D,j}|+|g_{I,j}|)^2\right)+\log\left(1+\frac{|g_{D,j}|^2}{1+|g_{I,j}|^2}\right)-2\log\left(1+|g_{D,j}|^2\right)$.

E. *Verification of multicarrier separability*

The verification for the non-bursty case (*i.e.*, when $p \in \{0,1\}$) is straightforward, and we focus on verifying the separability for the case when $0 < p < 1$ (*i.e.*, when the interfering link is bursty). We first verify the separability for symmetric capacity (*i.e.*, R_C and R_{NC}) in the case when all $\alpha_j \leq 2$, followed by the verification for the case when all $\alpha_j \geq 2$.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2697866, IEEE Transactions on Information Theory

All $\alpha_j \leq 2$*:* In this case, we can re-write Δ as shown below,

$$\begin{aligned}\Delta &= \sum_{j=1}^{M} \max(n_j, k_j) + (n_j - k_j)^+ - 2n_j \\ &= \sum_{j:\alpha_j>2} (k_j - 2n_j) - \sum_{j:\alpha_j\leq 1} k_j - \sum_{j:1<\alpha_j\leq 2} (2n_j - k_j) \\ &= -\sum_{j:\alpha_j\leq 1} k_j - \sum_{j:1<\alpha_j\leq 2} (2n_j - k_j). \end{aligned} \tag{42}$$

The sum of symmetric rates across subcarriers (using the optimal single carrier scheme [11] for each subcarrier) can be written as shown below:

$$\begin{aligned} &\left(\sum_{j:\alpha_j\leq 1} n_j - \frac{p}{1+p} k_j\right) + \left(\sum_{j:1<\alpha_j\leq 2} n_j - \frac{p}{1+p}(2n_j - k_j)\right) \\ &\stackrel{(a)}{=} \frac{p}{1+p}\Delta + \sum_{j=1}^{M} n_j = R_C, \end{aligned} \tag{43}$$

where (a) follows from (42). Hence, the symmetric capacity R_C (since $\Delta \leq 0$ in this case) can be achieved by simply using the optimal single carrier scheme [11] for each subcarrier.

All $\alpha_j \geq 2$*:* In this case,

$$\begin{aligned}\Delta &= \sum_{j=1}^{M} \max(n_j, k_j) + (n_j - k_j)^+ - 2n_j \\ &= \sum_{j:\alpha_j\geq 2} (k_j - 2n_j). \end{aligned} \tag{44}$$

The sum of symmetric rates across subcarriers (using the optimal single carrier scheme [11] for each subcarrier) can be written as shown below:

$$\sum_{j=1}^{M} n_j + \frac{p}{2}(k_j - 2n_j) \stackrel{(a)}{=} \frac{p}{2}\Delta + \sum_{j=1}^{M} n_j = R_{NC}, \tag{45}$$

where (a) follows from (44). Hence, the symmetric capacity R_{NC} (since $\Delta \geq 0$ in this case) can be achieved by simply using the optimal single carrier scheme [11] for each subcarrier.

Furthermore, in Appendix G we show that the corner points Q_1 and Q_2 can be achieved by treating the subcarriers separately (irrespective of α_j), and in Appendix F we show that separation is optimal for D_1 and D_2 when all $\alpha_j \leq 2$ or all $\alpha_j \geq 2$. In summary, the capacity region is achievable by treating the subcarriers separately when all $\alpha_j \leq 2$ or all $\alpha_j \geq 2$.

F. Achievability of corner points D_1 and D_2

As shown in Figure 3, these corner points appear when $\Delta > 0$. We will describe the achievability of D_1 and achievability of D_2 follows by symmetry. The achievability of D_1 is similar to achieving $R_{NC} = \frac{p}{2}\Delta + \sum_{j=1}^{M} n_j$ (described in Section IV-C); with a slight modification for subcarriers with $\alpha_j > 2$. The additive term $\frac{p}{2}\Delta$ appears in R_{NC} because of bursty relaying in the leftover helper levels (Δ in number). For D_1, to achieve $R^{(1)} = p\Delta + \sum_{j=1}^{M} n_j$, we use an asymmetric version of bursty relaying as follows: In every block Tx_1 sends $N_B\Delta$ linear combinations of $pN_B\Delta$ fresh symbols in the leftover helper levels. Rx_2 receives $pN_B\Delta$ such linear combinations in every block; it recovers the constituent symbols and forwards them to Tx_2. In the next block, Tx_2 creates $N_B\Delta$ linear combinations of the constituent symbols sent by Rx_1 and sends them on its leftover helper levels. Rx_1 receives $pN_B\Delta$ of these linear combinations and thus recovers the constituent symbols. So compared to R_{NC}, Rx_1 now gains an additional rate $\frac{p}{2}\Delta$ but Rx_2 loses[4] rate $\frac{p}{2}\Delta$. This completes the achievability of D_1.

In the special case when all $\alpha_j \leq 2$, or all $\alpha_j \geq 2$, D_1 and D_2 can be achieved by treating the subcarriers separately. Corner points D_1 and D_2 occur only when $\Delta > 0$. So between the conditions (i) all $\alpha_j \leq 2$, and (ii) all $\alpha_j \geq 2$, only condition (ii) (*i.e.*, all $\alpha_j \geq 2$) is relevant for D_1 and D_2 (for the other case $\Delta < 0$). Let us focus on D_1 when all $\alpha_j \geq 2$ (achievability for D_2 follows by symmetry). For this case, we can re-write $R^{(1)}$ for D_1 as shown below:

$$\begin{aligned} R^{(1)} &= p\Delta + \sum_{j=1}^{M} n_j \\ &= \sum_{j:\alpha_j\geq 2} p(k_j - 2n_j) + + \sum_{j=1}^{M} n_j \\ &\stackrel{(a)}{=} \sum_{j=1}^{M} p(k_j - 2n_j) + n_j \end{aligned}$$

where (a) follows from $\alpha_j \geq 2$ for all subcarriers. The rate tuple ($R_j^{(1)} = p(k_j - 2n_j) + n_j$, $R_j^{(2)} = n_j$) for subcarrier j is easily achievable by asymmetrically using the bursty relaying strategy for a single subcarrier (for $\alpha_j \geq 2$). To be more precise, in every block Tx_1 sends $N_B(k_j - 2n_j)$ linear combinations of $pN_B(k_j - 2n_j)$ fresh symbols in the $k_j - 2n_j$ helper levels. Rx_2 receives $pN_B(k_j - 2n_j)$ such linear combinations in every block; it recovers the constituent symbols and forwards them to Tx_2. In the next block, Tx_2 creates $N_B(k_j - 2n_j)$ linear combinations of the constituent symbols sent by Rx_1 and sends them on its helper levels. Rx_1 receives $pN_B(k_j - 2n_j)$ of these linear combinations and thus recovers the constituent symbols. Hence separation is optimal for D_1 when all $\alpha_j \geq 2$. Corner point D_2 can be achieved in a similar manner.

G. Achievability of corner points Q_1 and Q_2

Both Q_1 and Q_2 are achieved using a separation based scheme (*i.e.,* no coding across subcarriers). We first describe the achievability of Q_1; achievability of Q_2 follows by symmetry. In

$$\begin{aligned} Q_1 &= \left(R^{(1)}, R^{(2)}\right) \\ &= \left(p\Delta + \sum_{j=1}^{M} n_j(1+p) - (n_j - k_j)^+ p,\ \sum_{j=1}^{M} (n_j - k_j)^+\right), \end{aligned}$$

we can rewrite rate $R^{(1)}$ as follows:

$$\begin{aligned} &p\Delta + \sum_{j=1}^{M} n_j(1+p) - (n_j - k_j)^+ p \\ &= \sum_{j:\alpha_j\leq 1} n_j + \sum_{j:\alpha_j>1} n_j + (k_j - n_j)p. \end{aligned}$$

[4]The loss stems from Tx_2 not using its leftover helper levels for its own messages; it just uses them to relay messages for Rx_1.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2017.2697866, IEEE Transactions on Information Theory

Also, from the single carrier schemes in [11], the following rate tuples $\left(R^{(1)}, R^{(2)}\right)$ are achievable for a single carrier setup:

- $(n_j, n_j - k_j)$ for $\alpha_j \leq 1$.
- $(n_j + (k_j - n_j)p, 0)$ for $\alpha_j > 1$.

Clearly, achieving the above rate tuple for each subcarrier and summing rates across subcarriers leads to corner point Q_1. The achievability of Q_2 follows by symmetry.

References

[1] A. S. Avestimehr, S. N. Diggavi, C. Tian, and D. N. C. Tse, "An approximation approach to network information theory," Foundations and Trends in Communications and Information Theory, 2015.

[2] R. Etkin, D. Tse and H. Wang, "Gaussian interference channel capacity to within one bit," IEEE Transactions on Information Theory, vol. 54, pp. 5534-5562, Dec. 2008.

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

[4] N. Khude, V. Prabhakaran and P. Viswanath, "Opportunistic interference management," In Proc. IEEE Int. Symp. Inf. Theory (ISIT), 2009.

[5] N. Khude, V. Prabhakaran and P. Viswanath, "Harnessing bursty interference," In Proc. Information Theory Workshop (ITW), 2009.

[6] J. Chen, S. Yang, A. Ozgur and A. Goldsmith, "Achieving full DoF in heterogeneous parallel broadcast channels with outdated CSIT," IEEE Transactions on Information Theory, vol. 62, no. 7, pp. 4154-4171, July 2016.

[7] S. Avestimehr, S. Diggavi and D. Tse, "Wireless network information flow: a deterministic approach," IEEE Transactions on Information Theory, vol. 57, pp. 1872-1905, Apr. 2011.

[8] G. Bresler and D. Tse, "The two-user Gaussian interference channel: a deterministic view," European Transactions on Telecommunications, vol. 19, pp. 333-354, 2008.

[9] M. Madiman and P. Tetali, "Information inequalities for joint distributions, with interpretations and applications," IEEE Transactions on Information Theory, vol. 56, no. 6, pp. 2699-2713, June 2010.

[10] J. Jiang, N. Marukala and T. Liu, "Symmetrical multilevel diversity coding with an all-access encoder," In Proc. IEEE Int. Symp. Inf. Theory (ISIT), pp. 1662-1666, 2012.

[11] I.-H. Wang, C. Suh, S. Diggavi and P. Viswanath, "Bursty interference channel with feedback," In Proc. IEEE Int. Symp. Inf. Theory (ISIT) 2013. Extended version available: https://sites.google.com/site/ihsiangw/isit13preprintburstyic

[12] S. Mishra, I.-H. Wang and S. Diggavi, "Opportunistic interference management for multicarrier systems," In Proc. IEEE Int. Symp. Inf. Theory (ISIT) 2013. Extended version available: http://arxiv.org/abs/1305.2985

[13] S. A. Jafar and S. Vishwanath, "Generalized degrees of freedom of the symmetric Gaussian K user interference channel," IEEE Transactions on Information Theory, vol. 56, no. 7, pp. 3297-3303, Jul. 2010.

Shaunak Mishra received the B. Tech. degree in electronics and electrical communication engineering from the Indian Institute of Technology, Kharagpur, India in 2010, and the M.S. degree in electrical engineering from the University of California, Los Angeles in 2011. He is currently a Ph.D. student in the department of Electrical Engineering at University of California, Los Angeles. His research interests include statistics and information theory with applications in security, communication networks and machine learning. He is a recipient of the Henry Samueli fellowship at UCLA (2010).

I-Hsiang Wang received the B.Sc. degree in electrical engineering from National Taiwan University, Taiwan, in 2006. He received a Ph.D. degree in electrical engineering and computer sciences from the University of California at Berkeley, USA, in 2011. From 2011 to 2013, he was a postdoctoral researcher at Ècole Polytechnique Fèdèrale de Lausanne, Switzerland. Since 2013, he has been at the Department of Electrical Engineering in National Taiwan University, where he is now an assistant professor. His research interests include network information theory, wireless networks, and networked information and data processing. He was a finalist of the Best Student Paper Award of IEEE International Symposium on Information Theory, 2011.

Suhas N. Diggavi received the B. Tech. degree in electrical engineering from the Indian Institute of Technology, Delhi, India, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in 1998. After completing his Ph.D., he was a Principal Member Technical Staff in the Information Sciences Center, AT&T Shannon Laboratories, Florham Park, NJ. After that he was on the faculty of the School of Computer and Communication Sciences, EPFL, where he directed the Laboratory for Information and Communication Systems (LICOS). He is currently a Professor, in the Department of Electrical Engineering, at the University of California, Los Angeles, where he directs the Information Theory and Systems laboratory. His research interests include wireless network information theory, wireless networking systems, network data compression and network algorithms; more information can be found at http://licos.ee.ucla.edu. He has received several recognitions for his research including the 2013 IEEE Information Theory Society & Communications Society Joint Paper Award, the 2013 ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc) best paper award, the 2006 IEEE Donald Fink prize paper award. He is currently a Distinguished Lecturer and also serves on board of governors for the IEEE Information theory society. He is a Fellow of the IEEE. He has been an associate editor for IEEE Transactions on Information Theory, ACM/IEEE Transactions on Networking, IEEE Communication Letters, a guest editor for IEEE Selected Topics in Signal Processing and in the program committees of several IEEE conferences. He has also helped organize IEEE conferences including serving as the Technical Program Co-Chair for 2012 IEEE Information Theory Workshop (ITW) and the Technical Program Co-Chair for the 2015 IEEE International Symposium on Information Theory (ISIT). He has 8 issued patents.