



HAL
open science

Degrees of freedom of the MIMO interference channel with parallel multicasting

Jinyuan Chen, Andrea Goldsmith, Ayfer Ozgür, Sheng Yang

► **To cite this version:**

Jinyuan Chen, Andrea Goldsmith, Ayfer Ozgür, Sheng Yang. Degrees of freedom of the MIMO interference channel with parallel multicasting. IEEE International Symposium on Information Theory - (ISIT 2015), Jun 2015, Hong Kong, China. pp.1069-1073, 10.1109/ISIT.2015.7282619 . hal-01261221

HAL Id: hal-01261221

<https://centralesupelec.hal.science/hal-01261221>

Submitted on 10 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Degrees of Freedom of the MIMO Interference Channel with Parallel Multicasting

Jinyuan Chen
Stanford University, CA
jinyuanc@stanford.edu

Andrea Goldsmith
Stanford University, CA
andrea@ee.stanford.edu

Ayfer Özgür
Stanford University, CA
aozgur@stanford.edu

Sheng Yang
Centralesupélec, France
sheng.yang@centralesupelec.fr

Abstract—We investigate the degrees of freedom (DoF) for the two-user multiple-input multiple-output interference channel (MIMO IC) with parallel multicasting channels. Specifically, in addition to the MIMO IC, each transmitter is also connected to both receivers via an out-of-band multicast channel. Our main contribution lies in the characterization of the optimal sum DoF when the channel state information (CSI) on the MIMO IC is available to the transmitters with some delay (delayed CSIT). We show that jointly coding over the parallel multicast channels can achieve higher DoF than channel aggregation does. Furthermore, as long as the rate of the multicast channels is above a certain threshold, delayed CSIT is enough to achieve the same DoF performance as with instantaneous CSIT.

I. INTRODUCTION

Today’s wireless networks are heterogeneous, consisting of multiple radios with different capabilities, protocol stacks, and spectrum allocations. Such heterogeneous networks offer several advantages such as flexible resource allocation, better coverage, and higher capacity. In heterogeneous networks, transmitters can send information to users through parallel channels operating over different systems, such as a cellular system and a WiFi system.

While parallel channels have been studied extensively in the literature [1]–[8], these works typically considered parallel realizations of the same physical channel, e.g., over different time or frequency slots. Recently in [9], we studied heterogeneous parallel channels, where a transmitter sends information to receivers through two parallel channels: a multiple-input single-output broadcast channel (MISO BC) and a rate-limited multicast channel. We showed that jointly coding over the two channels can strictly outperform simple channel aggregation and can even achieve the same sum degrees of freedom (DoF) performance as with instantaneous CSIT when the CSIT on the MISO BC is completely stale, given that the multicast rate of the second channel is larger than a certain threshold. The main idea was to send information over the MISO BC at a rate above its capacity and use the second channel to send additional side information to allow for reliable decoding at both receivers (which we coined the overload-multicast strategy). Two key questions that remain open are whether the overload-multicast strategy can be optimally applied to other settings of heterogeneous parallel channels, and whether the insights obtained in [9] still hold, in particular that jointly

The work was partly supported by the NSF Center for Science of Information (CSoI) under grant NSF-CCF-0939370.

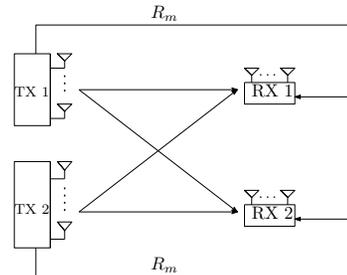


Fig. 1. Two-user MIMO IC with a rate-limited multicast channel.

coding over the parallel channels is optimal, and that outdated CSIT can achieve the same performance as with instantaneous CSIT.

With this motivation, in this work we study the following heterogeneous parallel channels. Two transmitters are connected to two receivers through two parallel channels: 1) the first channel is a multiple-input multiple-output interference channel (MIMO IC) with M antennas at each transmitter and N antennas at each receiver, and 2) the second channel is a rate-limited multicast channel through which each transmitter can multicast R_m bits/channel use of information to both receivers, as illustrated in Fig. 1.

Interestingly, we show that for this setting, jointly coding over the parallel channels still strictly outperforms the channel aggregation scheme when $M > N$ and only delayed CSIT is available. It is well known that the capacity of the Gaussian multi-antenna IC is very sensitive to the CSIT assumption. Let us consider an example with $M = 3$, $N = 2$. The sum DoF of the two-user MIMO IC are 3, $12/5$, and 2 for the cases with instantaneous CSIT [10], completely stale CSIT [11], [12], and no CSIT [13], respectively. Therefore, given that the first channel is a MIMO IC, it is obvious that channel aggregation also suffers from the same DoF loss due to the imperfect CSIT. In contrast, our proposed scheme can attain a total DoF of $3 + 2d_m$ with completely stale CSIT when the DoF of the parallel multicast channel d_m is larger than $3/2$. This DoF result is equivalent to the MIMO IC providing a sum DoF of 3 which, in isolation, is only possible when instantaneous CSIT is available. Analogous to [9], for this setting outdated CSIT can still achieve the same sum DoF as with instantaneous CSIT. This also suggests that the gain provided by feedback can be more significant when the multicast channel is present.

In this example, without the multicast channel, the DoF gain from feedback providing stale CSIT is $12/5 - 2 = 2/5$, while with the multicast channel, the DoF gain can be $(3 + 2d_m) - (2 + 2d_m) = 1$.

Our main contribution is the characterization of the optimal sum DoF of such channels for any rate of the multicast channel, any timeliness of the CSIT and any (M, N) antenna configuration. In particular, our result reveals an interesting tradeoff between the rate of the multicast channel and the CSIT timeliness when $M > N$: with timely CSIT a small value of d_m is enough to guarantee the optimal sum DoF, while with completely stale CSIT a large value of d_m is required to compensate for the DoF loss due to the CSIT staleness.

The main idea of our optimal scheme is based on the overload-multicast strategy, previously explored for parallel broadcast channels in [9]. However, the overload-multicast strategy proposed in this work brings some new perspectives. First of all, the overload-multicast strategy designed here has signals transmitted from two distributed transmitters. Since each transmitter does not know the message of the other, the signals transmitted from these two distributed transmitters have to be carefully designed in order to efficiently handle the interference. For example, each transmitter has to multicast side information that allows the unintended receiver to remove the interference caused by this transmitter. Second, this overload-multicast strategy achieves the optimal DoF performance for a larger range of (M, N) antenna configurations, i.e., for all $M > N$, by choosing the amount of overloading and multicasting as a function of the (M, N) antenna configuration. Note that the strategy in [9] was developed only for the MISO BC with a *single* antenna at each receiver.

II. SYSTEM MODEL

We focus on a two-user IC where transmitter k wishes to send information to receiver k , $k = 1, 2$, and the transmitters are connected to the two receivers through two parallel channels. The first channel is a MIMO IC with M antennas at each transmitter and N antennas at each receiver. The signals received over this MIMO IC by the two receivers at time t are

$$\mathbf{y}_1[t] = \mathbf{H}_{1,1}[t]\mathbf{x}_1[t] + \mathbf{H}_{1,2}[t]\mathbf{x}_2[t] + \mathbf{z}_1[t], \quad (1)$$

$$\mathbf{y}_2[t] = \mathbf{H}_{2,1}[t]\mathbf{x}_1[t] + \mathbf{H}_{2,2}[t]\mathbf{x}_2[t] + \mathbf{z}_2[t], \quad (2)$$

where $\mathbf{H}_{i,j}[t] \in \mathbb{C}^{N \times M}$ corresponds to the channel between transmitter j and receiver i at time t , and $\mathbf{z}_i[t]$ denotes the AWGN noise vector with distribution $\mathcal{CN}(\mathbf{0}, \mathbf{I})$, and $\mathbf{x}_i[t]$ denotes the transmitted signal vector of transmitter i at time t subject to an average power constraint P , where P is the signal-to-noise ratio (SNR) due to the noise power normalization, $i, j = 1, 2$. We assume a block fading model where all the channel coefficients remain constant during a coherence block of T_c channel uses and change independently from one block to the next. The channel coefficients in each block are i.i.d. complex Gaussian random variables with distribution $\mathcal{CN}(0, 1)$. We assume that the channel coefficients of $\mathbf{H}_{2,1}$ (cross-link) in each block are known to transmitter 1 only after γT_c channel uses, while the channel coefficients of $\mathbf{H}_{1,2}$

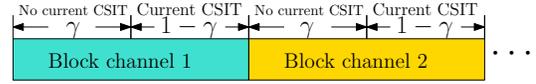


Fig. 2. Delayed CSIT and block channel model for the MIMO IC.

(cross-link) in each block are known to transmitter 2 only after γT_c channel uses, with $\gamma \in [0, 1]$. In other words, during the first γT_c channel uses, transmitter i only knows the channel coefficients of $\mathbf{H}_{j,i}$ ($j \neq i$) corresponding to the past blocks but not the coefficients of the current block. In the remaining $(1 - \gamma)T_c$ channel uses, the coefficients of $\mathbf{H}_{j,i}$ ($j \neq i$) of the current block are known perfectly to transmitter i (see Fig. 2). For this case, our emphasis is placed on the CSIT of the cross links simply because, as we will see, the DoF will be achieved without any CSIT knowledge of the channel realizations of the direct links. In terms of channel state information at the receivers, we assume that all the receivers perfectly and instantaneously know all the channel realizations of the cross links and direct links. In addition to the connection through the MIMO IC, we assume that each transmitter is also connected to the receivers through a parallel noiseless multicast channel, over which *each* transmitter can multicast information to both receivers at rate R_m bits per channel use.

We focus on the high SNR regime and the degrees of freedom performance of the system. For a given rate R_m of the multicast channel, and for an achievable rate pair $(R_1, R_2 | R_m)$, where R_k is the rate for user k , the corresponding DoF pair (d_1, d_2, \dots, d_K) is given by $d_k = \lim_{P \rightarrow \infty} \frac{R_k}{\log P}$, $k = 1, 2$. The corresponding DoF region \mathcal{D} is then the set of all achievable DoF tuples (d_1, d_2) , and the sum DoF is $d_{\text{sum}} = \sup\{d_1 + d_2 : (d_1, d_2) \in \mathcal{D}\}$. For notational convenience, we assume $R_m = d_m \log P$ and refer to d_m as the DoF of the multicast channel. d_m measures the multicast channel capacity in $\log P$ units and allows us to relate the multicast channel capacity to the capacity of the MIMO IC in the high SNR limit. Although we assume a noiseless multicast channel for simplicity, the analysis can be extended to the case where the multicast channel is noisy.

III. ILLUSTRATIVE EXAMPLE

To illustrate the main idea behind the proposed scheme, we start with an example and assume $M = 3, N = 2, R_m = 3/2 \log P$, $\gamma = 1$, where CSIT is completely outdated. For the sake of simplicity, we let $T_c = 1$ and assume that the multicast channel and the MIMO IC are synchronized in time. The scheme operates in packets of 6 symbols per transmitter-receiver pair. Packet i is communicated over channel uses $2i, 2i + 1$ of the MIMO IC (phase 1, phase 2) and channel uses $2(i + 1), 2(i + 1) + 1$ of the multicast channel (phase 3) as illustrated in Fig. 3. At the end of these three phases, each receiver can recover its 6 symbols which yields the optimal 6 sum DoF for the system. Next, we describe the transmission in three phases for a given packet.

1) *Phase 1 - transmit overload the MIMO IC*: In this phase, each transmitter sends 3 symbols in the following form

(ignoring the time index for simplicity):

$$\mathbf{x}_1 = [a_1 \ a_2 \ a_3]^\top, \quad \mathbf{x}_2 = [b_1 \ b_2 \ b_3]^\top, \quad (3)$$

where symbols a_i, b_i are intended for receiver 1 and receiver 2 respectively, for $i = 1, 2, 3$, and the power of each symbol is $P/3$. In this phase, we just use three signals received from three receiving antennas only. As we will show later on, we can achieve the optimal sum DoF by using those three signals only in this phase. Without loss of generality, we use two signals from two receiving antennas of receiver 1 and one signal from the first receiving antenna of receiver 2. Those received signals are given as

$$\mathbf{y}_1 = \mathbf{H}_{1,1} [a_1 \ a_2 \ a_3]^\top + \underbrace{\mathbf{H}_{1,2} [b_1 \ b_2 \ b_3]^\top}_{\triangleq [s_{1,1} \ s_{1,2}]^\top} + \mathbf{z}_1, \quad (4)$$

$$y_2 = \mathbf{h}_{2,2}^\top [b_1 \ b_2 \ b_3]^\top + \underbrace{\mathbf{h}_{2,1}^\top [a_1 \ a_2 \ a_3]^\top}_{\triangleq s_{2,1}} + \mathbf{z}_2, \quad (5)$$

where y_2 denotes the signal received from the first antenna of receiver 2, $\mathbf{h}_{2,2}^\top$ and $\mathbf{h}_{2,1}^\top$ denote the first row of channel matrices $\mathbf{H}_{2,2}$ and $\mathbf{H}_{2,1}$ respectively, where $[s_{1,1} \ s_{1,2}]^\top \triangleq \mathbf{H}_{1,2} [b_1 \ b_2 \ b_3]^\top$, $s_{2,1} \triangleq \mathbf{h}_{2,1}^\top [a_1 \ a_2 \ a_3]^\top$. Note that the total rate of the 6 symbols *overloaded* as in (3) surpasses the MIMO IC capacity if each symbol carries one DoF. One can see that if receiver 1 is able to learn the variables $s_{1,1}, s_{1,2}$ and $s_{2,1}$, then receiver 1 can remove the interference $s_{1,1}, s_{1,2}$ from \mathbf{y}_1 , and can use $s_{2,1}$ as an extra observation for decoding a_1, a_2, a_3 . On the other hand, with the knowledge of $s_{1,1}, s_{1,2}$ and $s_{2,1}$, receiver 2 can remove the interference $s_{2,1}$ from y_2 , and can use $s_{1,1}, s_{1,2}$ as an extra observation for decoding b_1, b_2, b_3 . Therefore, transmitter 1 and transmitter 2 will multicast the information about $s_{2,1}$ and $(s_{1,1}, s_{1,2})$ respectively to all receivers using the multicast channel in Phase 3.

2) *Phase 2 - transmit overload the MIMO IC*: Phase 2 is similar to Phase 1, but now the roles of the two transmitters are interchanged and the roles of the two receivers are interchanged. During phase 2, each transmitter sends 3 new symbols in the following form (ignoring the time index): $\mathbf{x}_1 = [a'_1 \ a'_2 \ a'_3]^\top$, $\mathbf{x}_2 = [b'_1 \ b'_2 \ b'_3]^\top$, where new symbols a'_i, b'_i are intended for receiver 1 and receiver 2, respectively, for $i = 1, 2, 3$, and the power of each symbol is $P/3$. In this phase, we only use two signals from two receiving antennas of receiver 2 and one signal from the first receiving antenna of receiver 1. Those received signals are given as

$$y_1 = \mathbf{h}_{1,1}^\top [a'_1 \ a'_2 \ a'_3]^\top + \underbrace{\mathbf{h}_{1,2}^\top [b'_1 \ b'_2 \ b'_3]^\top}_{\triangleq s'_{1,1}} + \mathbf{z}_1, \quad (6)$$

$$\mathbf{y}_2 = \mathbf{H}_{2,2} [b'_1 \ b'_2 \ b'_3]^\top + \underbrace{\mathbf{H}_{2,1} [a'_1 \ a'_2 \ a'_3]^\top}_{\triangleq [s'_{2,1} \ s'_{2,2}]^\top} + \mathbf{z}_2, \quad (7)$$

where y_1 denotes the signal of the first antenna of receiver 1, and $\mathbf{h}_{1,1}^\top$ and $\mathbf{h}_{1,2}^\top$ denote the first row of channel matrices $\mathbf{H}_{1,1}$ and $\mathbf{H}_{1,2}$ respectively. Again, with the knowledge of $s'_{1,1}, s'_{2,1}$ and $s'_{2,2}$ (cf. (6), (7)), both receiver 1 and receiver 2

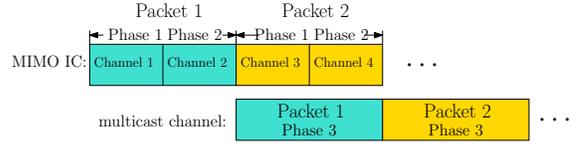


Fig. 3. Illustration of the three phases for the proposed scheme.

can decode their own symbols sent in Phase 2. Therefore, transmitter 1 and transmitter 2 will multicast the information about $(s'_{2,1}, s'_{2,2})$ and $s'_{1,1}$ respectively to all receivers in Phase 3.

3) *Phase 3 - multicast side information*: Phase 3 starts after the past CSI about $\mathbf{H}_{2,1}$ (resp. $\mathbf{H}_{1,2}$) is fed back to the transmitter 1 (resp. transmitter 2). Transmitter 1 (resp. transmitter 2) first *regenerates* $s_{2,1}, s'_{2,1}, s'_{2,2}$ (resp. $s_{1,1}, s_{1,2}, s'_{1,1}$) based on the past CSI, and then *quantizes* it into $\bar{s}_{2,1}, \bar{s}'_{2,1}, \bar{s}'_{2,2}$ (resp. $\bar{s}_{1,1}, \bar{s}_{1,2}, \bar{s}'_{1,1}$) by using $2R_m$ bits. Then, transmitter 1 (resp. transmitter 2) simply sends those $2R_m$ bits of the quantized information to both receivers through the multicast channel in two channel uses (since the multicast channel from each transmitter to both receivers has capacity R_m bits/channel use). After learning $\bar{s}_{1,1}, \bar{s}_{1,2}, \bar{s}_{2,1}$, receiver 1, 2 form their 3×3 MIMO observations of the form

$$\begin{bmatrix} \mathbf{y}_1 - [\bar{s}_{1,1} \ \bar{s}_{1,2}]^\top \\ \bar{s}_{2,1} \end{bmatrix} = \begin{bmatrix} \mathbf{H}_{1,1} \\ \mathbf{h}_{2,1}^\top \end{bmatrix} [a_1 \ a_2 \ a_3]^\top + \begin{bmatrix} \mathbf{z}_1 + [\tilde{s}_{1,1} \ \tilde{s}_{1,2}]^\top \\ -\bar{s}_{2,1} \end{bmatrix},$$

$$\begin{bmatrix} y_2 - \bar{s}_{2,1} \\ [\bar{s}_{1,1} \ \bar{s}_{1,2}]^\top \end{bmatrix} = \begin{bmatrix} \mathbf{h}_{2,2}^\top \\ \mathbf{H}_{1,2} \end{bmatrix} [b_1 \ b_2 \ b_3]^\top + \begin{bmatrix} z_2 + \bar{s}_{2,1} \\ -[\tilde{s}_{1,1} \ \tilde{s}_{1,2}]^\top \end{bmatrix},$$

respectively, corresponding to the transmission in phase 1, where $\tilde{s}_{i,j} \triangleq s_{i,j} - \bar{s}_{i,j}$, $i, j = 1, 2$, are the quantization errors. Since the power of $s_{i,j}$ is roughly P , it can be easily shown that the variance of the quantization error $\tilde{s}_{i,j}$ is at the noise level (cf. [14]). Therefore, with the help of the side information provided from the multicast channel, each receiver can recover its 3 symbols from the equivalent 3×3 MIMO channel, corresponding to the transmission of Phase 1. Similarly, each receiver can recover the other 3 symbols corresponding to the transmission of Phase 2. As a result, a sum DoF of 6 is achievable. Note that even if instantaneous perfect CSIT were available at the transmitters, the sum DoF performance could not scale better than 6. This example shows that completely outdated CSIT can be as good as instantaneous CSIT.

IV. MAIN RESULTS

We now return to the general system model of Section II, and provide the main results of this paper. The achievability proof is shown in Section V, while the converse proof can be found in the journal version of this work due to limited space here. First of all let us provide the maximum sum DoF achieved with instantaneous perfect CSIT ($\gamma = 0$):

$$d_{\text{sum}}^{(p)} \triangleq 2d_m + \min\{2M, 2N, \max\{M, N\}\}. \quad (8)$$

For this perfect CSIT case, channel aggregation is optimal in terms of sum DoF performance. The optimal sum DoF for the general case of γ is characterized in the following theorem.

Theorem 1 (Optimal DoF). For the channel model described in Section II, the optimal sum DoF is given by

$$d_{\text{sum}} = \begin{cases} d_{\text{sum}}^{(p)} & \text{if } M \leq N, \\ d_{\text{sum}}^{(p)} & \text{if } M > N, d_m \geq \frac{\min\{2N, M\} - 2(1-\gamma)N}{2}, \\ \Phi_o & \text{if } M > N, d_m < \frac{\min\{2N, M\} - 2(1-\gamma)N}{2}, \end{cases} \quad (9)$$

where $\Phi_o \triangleq \left(2 + \frac{4d_m}{N} + \frac{\min\{M, 2N\} - N}{\min\{M, 2N\}} 2(1-\gamma)\right) \frac{N \cdot \min\{M, 2N\}}{\min\{M, 2N\} + N}$.

The above result also characterizes the minimum DoF of the multicast channel, $d_m^* \triangleq \arg \min_{d_m} \{d_{\text{sum}}(d_m) = d_{\text{sum}}^{(p)}\}$, for achieving the instantaneous CSIT performance in terms of sum DoF, as given in the following corollary.

Corollary 1a (Minimum d_m). Given the CSIT delay fraction γ , the minimum value of d_m for achieving the instantaneous CSIT performance (in terms of sum DoF) is

$$d_m^* = \begin{cases} 0 & \text{if } M \leq N, \\ \max\{0, \frac{\min\{2N, M\} - 2(1-\gamma)N}{2}\} & \text{if } M > N. \end{cases} \quad (10)$$

When $M \leq N$, the above result shows that instantaneous CSIT performance is always achievable for any d_m . When $M > N$, the above result reveals that we need d_m^* DoF in the multicast channel to achieve the instantaneous CSIT performance. For example, when $M = 3, N = 2, \gamma = 1$, then $d_m^* = 3/2$.

Theorem 1 also characterizes the maximum CSIT delay fraction, $\gamma^* \triangleq \arg \max_{\gamma} \{d_{\text{sum}}(\gamma) = d_{\text{sum}}^{(p)}\}$, for achieving the maximum sum DoF, as given in the following corollary.

Corollary 1b (Maximum delay). The maximum CSIT delay fraction for achieving the maximum sum DoF is

$$\gamma^* = \begin{cases} 1 & \text{if } M \leq N, \\ \min\{1 - \frac{\min\{M, 2N\}}{2N} + \frac{d_m}{N}, 1\} & \text{if } M > N. \end{cases}$$

For a given multicast channel capacity, Corollary 1b gives the delay we can tolerate in feeding back the CSIT without sacrificing perfect CSIT performance, in the sum DoF sense. For example, when $M \geq 2N, d_m = N$, then $\gamma^* = 1$, i.e., completely stale CSIT is as good as instantaneous CSIT. On the other hand, when $M \geq 2N, d_m = N/2$, then $\gamma^* = 1/2$, i.e., we can tolerate a delay of a half coherence period and still achieve the instantaneous CSIT performance.

As we will show later on, for $M > N$ the optimal scheme of Theorem 1 is based on jointly coding over the parallel channels, while for $M \leq N$ channel aggregation is optimal in terms of the sum DoF. In the following proposition we provide the sum DoF for our system under channel aggregation, to use in comparison with the sum DoF in our overload-multicast scheme.

Proposition 1 (Channel aggregation). The sum DoF achieved with channel aggregation is given by

$$d_{\text{sum}}^{(ca)} = \begin{cases} d_{\text{sum}}^{(p)} & \text{if } M \leq N, \\ d_{\text{sum}}^{(p)} & \text{if } M > N, \gamma \leq 1 - \frac{\min\{M, 2N\}}{2N}, \\ \Phi_{ca} + 2d_m & \text{if } M > N, \gamma > 1 - \frac{\min\{M, 2N\}}{2N}, \end{cases} \quad (11)$$

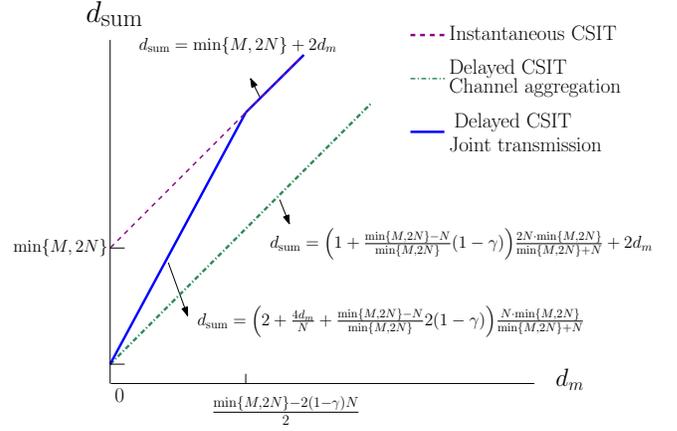


Fig. 4. Sum DoF d_{sum} vs d_m for optimal DoF with full CSIT, optimal DoF with delayed CSIT, and DoF achieved with channel aggregation under delayed CSIT: The case of $M > N$ and $\gamma > 1 - \frac{\min\{M, 2N\}}{2N}$.

which is strictly suboptimal for the case of $M > N$ and $\gamma > 1 - \frac{\min\{M, 2N\}}{2N}$, where $\Phi_{ca} \triangleq \left(2 \frac{\min\{2M, N\}}{\min\{M, N\}} + \frac{\min\{M, 2N\} - \min\{M, N\}}{\min\{M, 2N\}} 2(1-\gamma)\right) \frac{\min\{M, N\} \cdot \min\{M, 2N\}}{\min\{M, 2N\} + \min\{M, N\}}$.

The proof of (11) is simple. From Theorem 1 we note that the maximum sum DoF achieved over the MIMO IC only is $\min\{d_{\text{sum}}^{(p)} - 2d_m, \Phi_{ca}\}$. On the other hand, the maximum sum DoF of the multicast channel is $2d_m$. Hence channel aggregation gives the sum DoF as in (11). The results of Theorem 1 and Proposition 1 reveal that channel aggregation is strictly suboptimal for the case of $M > N$ and $\gamma > 1 - \frac{\min\{M, 2N\}}{2N}$, as illustrated in Fig. 4.

V. ACHIEVABILITY OF THEOREM 1

This section provides the achievability proof of Theorem 1. First of all, when $M \leq N$ the sum DoF in Theorem 1 can be easily achieved with no CSIT and with channel aggregation. When $M \geq 2N$, the sum DoF in Theorem 1 can be reduced to the one with $M = 2N$. Therefore, we just focus on the achievability proof for the case of $N < M \leq 2N$. The proposed scheme is an extension of the illustrative example in Section III. We next summarize the following basic strategies of this general scheme.

Whenever instantaneous CSIT is available over the MIMO IC, M new symbols are sent to two users with spatial zero-forcing (ZF) precoding, allowing each user to decode its $M - N$ private symbols and $(2N - M)$ common symbols in one channel use. Specifically those $(2N - M)$ common symbols are used for multicasting side information, as will be discussed.

When instantaneous CSIT is not available, the transmitters can do one of two things over different time fractions of each channel block: (i) they can overload the MIMO IC with independent symbols (over a δ fraction of the block), (ii) they can multicast side information (over a $\gamma - \delta$ fraction of the block). When each of the transmitters overloads symbols, it transmits M fresh symbols in one channel use by using the signaling technique in (3). In order to decode these M symbols

by the corresponding user, extra side information of $M \log P$ bits from the two transmitters will be sent to both users.

When the multicast channel capacity is very large, using it only a fraction η of the time may be sufficient to multicast the generated side information. During the remaining $(1 - \eta)$ fraction of time, the multicast channel can be used for sending new information to the corresponding user. Note that the multicasting rate and the single user transmission rate over the multicast channel are both R_m bits per channel use. The design of (δ, η) allows us to achieve the optimal sum DoF as given in the following.

a) When $d_m \geq \frac{M-2(1-\gamma)N}{2} \geq 0$: In this case we set $\delta = \gamma$, $\eta = \frac{M-2(1-\gamma)N}{2d_m}$, so that the amount of side information generated for one block matches the amount of side information multicast in next block, i.e.,

$$\underbrace{M\delta}_{\text{generated}} = \underbrace{(2N-M)(1-\gamma)}_{\text{multicast over MIMO IC, perfect CSIT}} + \underbrace{2\eta d_m}_{\text{over multicast channel}}. \quad (12)$$

Note that the first term in the RHS of (12) corresponds to the amount of side information multicast over the MIMO IC whenever instantaneous CSIT is available. Since the multicast channel capacity is large enough, the multicast channel is used in a time fraction η for multicasting the side information and in the rest of the time it is used for single user transmission, which yields the following optimal sum DoF

$$\begin{aligned} d_1 + d_2 &= \underbrace{2M\delta}_{\text{overloading}} + \underbrace{2(1-\eta)d_m}_{\text{single user, multicast channel}} + \underbrace{2(M-N)(1-\gamma)}_{\text{ZF}} \\ &= 2d_m + M. \end{aligned} \quad (13)$$

During overloading in the scheme we transmit M symbols per transmitter-receiver pair, hence achieve $2M$ DoF in total once the side information is multicast to both receivers (cf. (12)). By contrast, single user transmission over the multicast channel provides d_m DoF per transmitter-receiver pair. Finally the zero-forcing operation is performed in a $(1-\gamma)$ fraction of the time, which gives $2(M-N)(1-\gamma)$ DoF.

b) When $0 \leq d_m < \frac{M-2(1-\gamma)N}{2}$: In this case, the multicast channel capacity is insufficient. We set $\eta = 1$ and $\delta = \frac{2N-M+(M-N)\gamma+2d_m}{M+N}$, so that again the amount of side information generated for one block matches the amount of side information multicast in next block, i.e.,

$$M\delta = \underbrace{(2N-M)(1-\gamma)}_{\text{MIMO IC, perfect CSIT}} + \underbrace{N(\gamma-\delta)}_{\text{MIMO IC, no CSIT}} + \underbrace{2\eta d_m}_{\text{multicast channel}}.$$

In this case, the multicast channel is used fully for multicasting the side information, while the MIMO IC is used in time fractions δ and $\gamma - \delta$ for overloading as in strategy (i) and for side information multicasting as in strategy (ii) respectively, which allows us to achieve the following optimal sum DoF:

$$\begin{aligned} d_{\text{sum}} &= \underbrace{2M\delta}_{\text{overloading}} + \underbrace{2(M-N)(1-\gamma)}_{\text{ZF}} \\ &= (4Md_m + 2N(2M-N) - 2N(M-N)\gamma)/(M+N). \end{aligned}$$

c) When $\frac{M-2(1-\gamma)N}{2} < 0$: In this case the CSIT is not significantly delayed, i.e., $\gamma < 1 - \frac{M}{2N}$. Hence we just disregard the first $1 - \frac{M}{2N} - \gamma$ fraction of the current CSIT, which means that the CSIT delay fraction can be treated as $\tilde{\gamma} \triangleq 1 - \frac{M}{2N}$. Then one can easily repeat the same scheme as in Section V-a and replace γ with $\tilde{\gamma}$ to achieve the optimal sum DoF $d_{\text{sum}} = 2d_m + M$ as in (13).

VI. CONCLUSIONS

We have characterized the optimal sum DoF of the two-user MIMO IC with parallel rate-limited multicast channels, for any (M, N) antenna configuration of the MIMO IC, any DoF of the multicast channel and any timeliness of the CSIT. We showed that completely stale CSIT can achieve the same DoF performance as with instantaneous CSIT, if the multicast channel capacity is large enough. Our transmission scheme is based on a simple overload-multicast strategy, i.e., first transmit symbols over the MIMO IC at a rate above its capacity and then use the multicast channel to multicast additional information to enable reliable decoding.

REFERENCES

- [1] D. Hughes-Hartog, "The capacity of a degraded spectral Gaussian broadcast channel," Ph.D. dissertation, Stanford University, Jul. 1975.
- [2] A. El Gamal, "Capacity of the product and sum of two unmatched broadcast channels," *Probl. Peredachi Inf.*, vol. 6, no. 1, pp. 3 – 23, 1980.
- [3] A. J. Goldsmith and M. Effros, "The capacity region of broadcast channels with intersymbol interference and colored Gaussian noise," *IEEE Trans. Inf. Theory*, vol. 47, no. 1, pp. 219 – 240, Jan. 2001.
- [4] R. S. Cheng and S. Verdú, "Gaussian multiaccess channels with ISI: capacity region and multiuser water-filling," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 773 – 785, May 1993.
- [5] D. N. C. Tse and S. V. Hanly, "Multiaccess fading channels—part I: polyanatroid structure, optimal resource allocation and throughput capacities," *IEEE Trans. Inf. Theory*, vol. 44, no. 7, pp. 2796 – 2815, Nov. 1998.
- [6] L. Li and A. J. Goldsmith, "Capacity and optimal resource allocation for fading broadcast channels—Part I: Ergodic capacity," *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 1083 – 1102, Mar. 2001.
- [7] X. Shang, B. Chen, G. Kramer, and H. V. Poor, "Noisy-interference sum rate capacity of parallel Gaussian interference channels," *IEEE Trans. Inf. Theory*, vol. 57, no. 1, pp. 210 – 226, Jan. 2011.
- [8] V. R. Cadambe and S. A. Jafar, "Parallel Gaussian interference channels are not always separable," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 3983 – 3990, Sep. 2009.
- [9] J. Chen, S. Yang, A. Özgür, and A. Goldsmith, "Outdated CSIT can achieve full DoF in heterogeneous parallel channels," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2014.
- [10] S. A. Jafar and M. Fakhereddin, "Degrees of freedom for the MIMO interference channel," *IEEE Trans. Inf. Theory*, vol. 53, no. 7, pp. 2637 – 2642, Jul. 2007.
- [11] C. S. Vaze and M. K. Varanasi, "The degrees of freedom region and interference alignment for the MIMO interference channel with delayed CSIT," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4396 – 4417, Jul. 2012.
- [12] M. A. Maddah-Ali and D. N. C. Tse, "Completely stale transmitter channel state information is still very useful," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4418 – 4431, Jul. 2012.
- [13] C. Huang, S. A. Jafar, S. Shamai, and S. Vishwanath, "On degrees of freedom region of MIMO networks without channel state information at transmitters," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 849 – 857, Feb. 2012.
- [14] T. M. Cover and J. A. Thomas, *Elements of information theory*. New York: Wiley-Interscience, 2006.