An achievable rate region of broadcast relay channel with state feedback

Chao He, Sheng Yang, Pablo Piantanida

An Achievable Rate Region of Broadcast Relay Channel with State Feedback

Chao He, Sheng Yang, and Pablo Piantanida
Laboratoire des Signaux et Systèmes (L2S, UMR CNRS 8506)
CentraleSupélec - CNRS - Université Paris-Sud, 91192, Gif-sur-Yvette, FRANCE
Email: {chao.he, sheng.yang, pablo.piantanida}@centralesupelec.fr

Abstract—In this paper, we investigate the fast fading broadcast relay channel (BRC) with one source (macrocell BS), one relay (smallcell BS), and two destinations (mobile users). We assume that the instantaneous channel state information (CSI) is known only at the receivers' side, whereas the relay can obtain a noiseless CSI feedback from the destinations about the source-destination channels (with delay). An inner bound on the capacity region of this channel is derived based on a partial-decode-compress-forward (PDCF) scheme. The main idea is to let the relay decode a significant portion of the source messages and then generate and forward to the receivers some useful side information using both the decoded messages and the CSI feedback. Numerical results show that, thanks to the feedback, the proposed scheme provides a non-negligible gain over conventional decode-forward/compress-forward schemes in terms of sum-rate performance, especially in the high SNR regime. Quite remarkably, we show that overall transmit power of the macrocell network can be significantly reduced owing to the feedback from users to the smallcell BS.

I. INTRODUCTION

Future wireless communication systems will feature ultra dense deployment of base stations (BS) and heterogeneity. As the number of devices increases and the demanded traffic explodes, interference will become the bottleneck of overall network throughput. One way to accommodate the exponential growth of network data flow is to employ the spatial multiplexing gain provided by multi-antenna techniques (MIMO, or even massive MIMO). Another way is to exploit the spatial reuse by densifying the deployment of BSs while reducing the transmit power. Multiple tiers of mobile networks will coexist, as is already the case today, so that the mobiles can operate over multiple radios to achieve optimal performance.

In this paper, we consider a typical setup in which a macrocell BS serves multiple users who are in close proximity to a smallcell (e.g., femto/pico) BS. A natural question to ask is whether and how smallcell BS could improve downlink performance of the macrocell network. Essentially, the smallcell BS in this setting may be considered as a relay, and the overall channel becomes a broadcast relay channel (BRC).

The capacity of general BRC is unknown. In fact, even the capacity of relay channel is still open, except for a few special cases (e.g., degraded relay channel) [1]. The capacity region of static Gaussian noise broadcast channel without relay is known for a decade now [2]. This result can be easily extended to fading case when channel state information (CSI) is perfectly known at both transmitter (CSIT) and receivers. When no CSIT is ensured, however, the capacity is still unknown, with the exception of the spatially symmetric channel distribution [3]. It is well known that the degrees of freedom (DoF) gain in a downlink MIMO channel hinges on the quality of CSIT: full DoF can be achieved when channel estimation error vanishes with the signal-to-noise ratio (SNR) as $O(\mathsf{SNR}^{-1})$ [4], while no DoF gain at all can be obtained when the CSI error does not vanish with SNR [5].

In most practical scenarios, the CSIT quality may be inadequate for various reasons, namely, noisy CSIT caused by poor feedback quality (or insufficient uplink training, e.g., for TDD systems) or outdated feedback due to high user mobility. If the noisy CSIT could be avoided by investing more resource at the uplink, the timeliness of the CSI feedback cannot be guaranteed anyhow when the channel varies too rapidly. However, completely outdated CSIT is still very useful, for that almost linear DoF gain with the number of users and the number of transmit antennas at the BS can be obtained [6]. This conclusion holds as long as the transmitter can receive the CSI feedback (with arbitrary but finite delay) from all the receivers and that the CSI error vanishes as $O(\mathsf{SNR}^{-1})$. The proposed scheme, also referred to as the MAT scheme, can be interpreted as the space-time interference alignment which is possible with the delayed CSIT. Since [6], several extensions have been made in various configurations (cf. [7]–[11]). The main idea of these works consists in two phases: 1) the broadcast phase in which signal containing new messages is sent to the receivers, and 2) the multicast phase in which side information generated based on both the past channel and the sent messages is diffused. It is therefore intuitive to think of the transmitter as two separate virtual transmitters: one takes charge of the first phase without the need of any CSIT, whereas the other is responsible for the second phase and uses the delayed CSIT and the transmitted message. The second virtual transmitter is very much like a relay that "learns" the messages and the channel coefficients and "forwards" the side information to "help" the receivers decode their own messages.

In this work, we deploy a real relay that helps accomplish the space-time interference alignment *a la MAT*. The main motivation for this setting is that accurate (i.e. with vanishing error $O(\mathsf{SNR}^{-1})$) CSI feedback via uplink is costly in practice. The cost is in terms of both the uplink bandwidth and the uplink transmit power related to the feedback. In the proposed setup, the CSI is fed back from the receivers to the relay (e.g.,

smallcell BS) which is supposed to be at a closer distance than the BS. In other words, the feedback communication is localized and transmit power could be reduced.

The main contribution of the current work is therefore to investigate the potential gain of the use of relay in the presence of state feedback from the receivers to the relay. Specifically, we first derive an achievable rate region for the general two-user discrete-memoryless broadcast relay channel with state feedback to the relay. This channel includes the two-user Gaussian noise fading BRC as a special case. The proposed scheme is based on the aforementioned space-time interference alignment with the help of a relay. Then, with numerical simulation, we compare our scheme with the conventional decode-forward (DF) and compress-forward (CF) schemes directly adapted to the fading broadcast channel. Our results reveal that the proposed scheme provides a non-negligible gain over conventional schemes in terms of sum-rate performance, especially in the high SNR regime. Quite remarkably, we show that it is better for the macrocell BS not to transmit in its full power with our proposed scheme, whereas the conventional schemes should always use the full transmit power. The underlying implication is that the overall transmit power of the macrocell network can be significantly reduced thanks to the feedback from the users to the smallcell BS.

This paper is organized as follows. Section II provides achievable regions in general BRC while Section III characterizes the region by proposed scheme in Gaussian fading case. In Section IV, numerical results are shown followed by scheme illustration. The achievability is presented in Section V.

Throughout the paper, $(\cdot)^T$ and $(\cdot)^H$ denotes transpose and conjugated transpose of a matrix/vector, respectively. $\|\cdot\|$ is the Euclidean norm. $\mathrm{diag}(a_1, a_2, \cdots, a_k)$ is a diagonal matrix with diagonal entries $(a_1, a_2, \cdots, a_k)$. $\mathcal{T}_\delta^n(\cdot)$ is the jointly typical set. $\mathbb{E}(\cdot)$ is the expectation of a random variable. We use $f(x) = O(g(x))$ as the standard Landau notation. The logarithm is to base 2.

II. System Model and General Results

In this section, we provide the main results on the general two-user discrete memoryless state-dependent broadcast channel with a relay which always works in full-duplex mode. The channel is defined by the joint probability mass function (pmf) $p(y_1^n, y_2^n, y_r^n | x_s^n, x_r^n, s^n) p(s^n) = \prod_{i=1}^n p(y_{1i}, y_{2i}, y_{ri} | x_{si}, x_{ri}, s_i) p(s_i)$, where x_s and x_r are the transmitted signal at the source and relay, respectively; y_1, y_2, and y_r are the received signal at user 1, user 2, and the relay, respectively; s is the channel state; i is the time index. We assume that the source is not aware of the channel state; each user knows its instantaneous state perfectly; the relay can obtain a noiseless feedback from the destinations about the state S_i at time $i + 1$, i.e., with a delay of one symbol.

In the following, we start by providing the achievable regions by the conventional DF and CF schemes, adapted to our broadcast setting. Then, we present the main result of this work on the proposed partial decode and compress forward (PDCF) scheme. The main idea is to let the relay decode a significant

part of the source messages and then generate and forward to the receivers some useful side information employing both the decoded messages and the state feedback. The side information is exploited by each user to mitigate its own interference. It is worth noting that the proposed scheme contains the CF scheme as a special case. To make the presentation fluid, the proof of the result with the PDCF scheme is deferred to Section V. The proofs related to DF and CF schemes are rather standard and omitted due to the lack of space (cf. [12], [13]).

A. CF scheme

Theorem 1. (CF inner bound) *An achievable rate region for the memoryless state-dependent BRC with the CF scheme is given by the set of all rate pairs (R_1, R_2) satisfying:*

$$R_1 \leq \min \Big\{ I(X_r U_1; Y_1 | SQ) - I(Y_r; \hat{Y} | X_r U_1 Y_1 SQ),$$
$$I(U_1; \hat{Y} Y_1 | X_r SQ) \Big\}, \quad (1)$$

$$R_2 \leq \min \Big\{ I(X_r U_2; Y_2 | SQ) - I(Y_r; \hat{Y} | X_r U_2 Y_2 SQ),$$
$$I(U_2; \hat{Y} Y_2 | X_r SQ) \Big\}, \quad (2)$$

for $k \in \{1, 2\}$, and all pmf's that verify:

$$p(y_1, y_2, y_r, \hat{y}, x_r, u_1, u_2, x_s, s, q) = p(y_1, y_2, y_r | x_s, x_r, s)$$
$$\times p(\hat{y} | x_r, y_r) p(x_r) p(u_1) p(u_2) p(x_s | u_1, u_2, q) p(s) p(q),$$
$$0 \leq I(X_r; Y_k | U_k SQ) - I(Y_r; \hat{Y} | X_r U_k Y_k SQ).$$

This scheme uses a block-Markov coding combined with Wyner-Ziv compress of the relay output. The description of the compress is then broadcast to both users. At each user, backward decoding is performed [1].

B. DF scheme

Theorem 2. (DF inner bound) *An achievable rate region for the memoryless state-dependent BRC by employing DF scheme is given by the set of all rate pairs (R_1, R_2) satisfying:*

$$R_1 \leq \min\{I(U_1 X_1; Y_1 | SQ), I(X_1; Y_r | U_1 U_2 X_2 SQ)\}, \quad (3)$$
$$R_2 \leq \min\{I(U_2 X_2; Y_2 | SQ), I(X_2; Y_r | U_1 U_2 X_1 SQ)\}, \quad (4)$$
$$R_1 + R_2 \leq I(X_1 X_2; Y_r | U_1 U_2 SQ), \quad (5)$$

for all pmf's that verify:

$$p(y_1 y_2 y_r u_1 u_2 x_1 x_2 x_s x_r sq) = p(y_1 y_2 y_r | x_s x_r s) p(u_1) p(u_2)$$
$$p(x_1 | u_1) p(x_2 | u_2) p(x_s | x_1 x_2 q) p(x_r | u_1 u_2 q) p(s) p(q).$$

A similar block-Markov superposition coding scheme is performed. Instead of compression, the relay decodes the messages, re-encodes and forwards them to the destinations.

C. PDCF scheme

Theorem 3. (PDCF inner bound) *An inner bound on the capacity of memoryless state-dependent BRC with state feedback is given by the set of all rates pair (R_1, R_2) satisfying:*

$$R_1 \leq \min \Big\{ \{I_{11}\}^- + I_{31}, \{I_{21}\}^- + I_{41} + I_{51} \Big\}, \quad (6)$$

$$R_2 \leq \min \Big\{ \{I_{12}\}^- + I_{32}, \{I_{22}\}^- + I_{42} + I_{52} \Big\}, \quad (7)$$

$$R_1 + R_2 \leq I_6 + I_{51} + \{I_{21}\}^- + I_{52} + \{I_{22}\}^-, \quad (8)$$

for all pmf's that verify:

$$p(y_1, y_2, y_r, \hat{y}, y_r', x_s, x_r, u_1, u_2, x_1, x_2, s, q) =$$
$$p(y_1, y_2, y_r | x_s, x_r, s) p(\hat{y} | x_r, y_r') p(y_r' | y_r, u_1, u_2, s) p(x_r)$$
$$\times p(u_1) p(u_2) p(x_1 | u_1) p(x_2 | u_2) p(x_s | x_1, x_2, q) p(s) p(q)$$

$$0 \le I(X_r; Y_k | U_k X_k S Q) - I(\hat{Y}; Y_r' | X_r U_k X_k Y_k Q) \quad (9)$$

where $\{f\}^- = \min(0, f)$, *and* $k, j \in \{1, 2\}, (k \ne j)$,

$$I_{1k} = I(X_r; Y_k | SQ) - I(\hat{Y}; Y_r' | X_r Y_k Q), \quad (10)$$
$$I_{2k} = I(X_r; Y_k | U_k SQ) - I(\hat{Y}; Y_r' | X_r U_k Y_k Q), \quad (11)$$
$$I_{3k} = I(U_k X_k; Y_k \hat{Y} | X_r SQ), \quad (12)$$
$$I_{4k} = I(U_k; Y_r | X_r U_j SQ), \quad (13)$$
$$I_{5k} = I(X_k; Y_k \hat{Y} | X_r U_k SQ), \quad (14)$$
$$I_6 = I(U_1 U_2; Y_r | X_r SQ). \quad (15)$$

III. GAUSSIAN NOISE FADING CHANNEL

In this section, we investigate a fast fading BRC. For simplicity of description, we assume that the source is equipped with two antennas and the relay has two transmit and two receive antennas while each destination has a single antenna. The signal model is, for $i = 1, 2, \cdots, n$,

$$\mathbf{y}[i] = \mathbf{H}[i] \mathbf{x}_s[i] + \mathbf{G}[i] \mathbf{x}_r[i] + \mathbf{z}[i], \quad (16)$$
$$\mathbf{y}_r[i] = \mathbf{H}_r[i] \mathbf{x}_s[i] + \mathbf{z}_r[i], \quad (17)$$

where $\mathbf{x}_s[i], \mathbf{x}_r[i] \in \mathbb{C}^{2 \times 1}$ are the channel input at source and at relay subject to the power constraints $\frac{1}{n} \sum_{i=1}^{n} \mathbb{E}(\|\mathbf{x}_s[i]\|^2) \le P_s$ and $\frac{1}{n} \sum_{i=1}^{n} \mathbb{E}(\|\mathbf{x}_r[i]\|^2) \le P_r$, respectively; $\mathbf{y}[i] \triangleq [Y_1, Y_2]^T$ with Y_1, Y_2 being channel output at respective user while $\mathbf{y}_r[i]$ is the relay output; $\mathbf{z}$ and $\mathbf{z}_r$ are the additive Gaussian noises with temporarily and spatially i.i.d. entries $\mathcal{CN}(0, 1)$; $\mathbf{H}[i]$, $\mathbf{H}_r[i]$, and $\mathbf{G}[i]$ are temporarily and spatially i.i.d. channel matrix processes, where we define $\mathbf{H} \triangleq [\mathbf{h}_1, \mathbf{h}_2]^T, \mathbf{H}_r \triangleq [\mathbf{h}_{r_1}, \mathbf{h}_{r_2}]^T$ and $\mathbf{G} \triangleq [\mathbf{g}_1, \mathbf{g}_2]^T$. In PDCF scheme, we can identify $(\mathbf{H}[i], \mathbf{H}_r[i], \mathbf{G}[i])$ with S_i, $\mathbf{x}_s[i]$ with X_{si}, $\mathbf{x}_r[i]$ with X_{ri}, $\mathbf{y}_r[i]$ with Y_{ri}, and $\mathbf{y}[i]$ with (Y_{1i}, Y_{2i}) in the general memoryless channel. Then, we let X_r be $\mathbf{x}_r$ and X_s be $\mathbf{x}_s = \mathbf{x}_1 + \mathbf{x}_2$, where $\mathbf{x}_k = \mathbf{x}_k' + \mathbf{u}_k$ with $\mathbf{x}_k'$ being independent of $\mathbf{u}_k$. Moreover, we let Y_r' be $(\mathbf{y}_r^T, \mathbf{h}_1^T \mathbf{u}_2, \mathbf{h}_2^T \mathbf{u}_1)$, where $\mathbf{h}_1^T \mathbf{u}_2, \mathbf{h}_2^T \mathbf{u}_1$ are the interferences at user 1, 2, respectively. Finally, we let $\hat{Y}$ be compressed version of Y_r' with compression noise defined by $\hat{\mathbf{z}} \triangleq [\hat{z}_{r_1}, \hat{z}_{r_2}, \hat{z}_1, \hat{z}_2]^T$. Indeed, the scheme works as follows. Relay receives $\mathbf{h}_1, \mathbf{h}_2$ and recovers $\mathbf{u}_1, \mathbf{u}_2$ from $\mathbf{y}_r$ after source finishes sending $\mathbf{x}_s$, followed by generating $\mathbf{y}_r'$ as a function of interferences

$\mathbf{h}_1^T \mathbf{u}_2, \mathbf{h}_2^T \mathbf{u}_1$ and relay output $\mathbf{y}_r$. Then, $\mathbf{y}_r'$ is compressed to $\hat{\mathbf{y}}$ by adding compression noise $\hat{\mathbf{z}}$ and compress index is sent through $\mathbf{x}_r$. At each user, messages are decoded with the presence of these compressed side information. For simplicity, we choose $\mathbf{x}_r \sim \mathcal{CN}(\mathbf{0}, \frac{P_r}{2} \mathbf{I}_2)$, $\mathbf{x}_k' \sim \mathcal{CN}(\mathbf{0}, P_{x_k} \mathbf{I}_2)$, $\mathbf{u}_k \sim \mathcal{CN}(\mathbf{0}, P_{u_k} \mathbf{I}_2)$, with $P_1 + P_2 = \frac{P_s}{2}$ and $P_k = P_{x_k} + P_{u_k}$; and $\hat{\mathbf{z}} \sim \mathcal{CN}(\mathbf{0}, \text{diag}(N_{r_1}, N_{r_2}, N_1, N_2))$. To get concise expressions, we define $\mathbf{A}_{1k}, \mathbf{A}_{2k}, \mathbf{A}_{3k}$ in $(18), (19), (20)$ and $\mathcal{B}(\mathbf{T}) = \mathbf{T} \mathbf{T}^H$, $P_{k, x_j} = P_k + P_{x_j}$, $P_{u_1, u_2} = P_{u_1} + P_{u_2}$, $P_{x_1, x_2} = P_{x_1} + P_{x_2}$, $k, j \in \{1, 2\}, k \ne j$. Thus, we have:

$$I(X_r; Y_1 | SQ) = \mathbb{E} \log \left[1 + \frac{\|\mathbf{g}_1\|^2 (P_r/2)}{\|\mathbf{h}_1\|^2 (P_s/2) + 1} \right],$$

$$I(X_r; Y_1 | U_1 SQ) = \mathbb{E} \log \left[1 + \frac{\|\mathbf{g}_1\|^2 (P_r/2)}{\|\mathbf{h}_1\|^2 P_{2, x_1} + 1} \right],$$

$$I(X_r; Y_1 | U_1 X_1 SQ) = \mathbb{E} \log \left[1 + \frac{\|\mathbf{g}_1\|^2 (P_r/2)}{\|\mathbf{h}_1\|^2 P_2 + 1} \right],$$

$$I(U_1; Y_r | X_r U_2 SQ) = \mathbb{E} \log \left[\mathbf{I}_2 + \mathbf{H}_r \mathbf{H}_r^H P_{1, x_2} \right]$$
$$- \mathbb{E} \log \left[\mathbf{I}_2 + \mathbf{H}_r \mathbf{H}_r^H P_{x_1, x_2} \right],$$

$$I(U_1 U_2; Y_r | X_r SQ) = \mathbb{E} \log \left[\mathbf{I}_2 + \mathbf{H}_r \mathbf{H}_r^H (P_s/2) \right]$$
$$- \mathbb{E} \log \left[\mathbf{I}_2 + \mathbf{H}_r \mathbf{H}_r^H P_{x_1, x_2} \right],$$

$$I(X_1; Y_1 \hat{Y} | X_r U_1 SQ) = \mathbb{E} \log \left[\det(\mathbf{A}_{11} \mathbf{A}_{21}^{-1}) \right],$$

$$I(U_1 X_1; Y_1 \hat{Y} | X_r SQ) = \mathbb{E} \log \left[N_2^{-1} \det(\mathbf{A}_{31} \mathbf{A}_{21}^{-1}) \right],$$

$$I(\hat{Y}; Y_r' | X_r U_1 Y_1 SQ) = \mathbb{E} \log \left[\det(\mathbf{A}_{11}) \right]$$
$$- \mathbb{E} \log \left[N_{r_1} N_{r_2} N_1 (1 + \|\mathbf{h}_1\|^2 P_{2, x_1}) \right],$$

$$I(\hat{Y}; Y_r' | X_r Y_1 SQ) = \mathbb{E} \log \left[\det(\mathbf{A}_{31}) \right]$$
$$- \mathbb{E} \log \left[N_{r_1} N_{r_2} N_1 N_2 (1 + \|\mathbf{h}_1\|^2 (P_s/2)) \right],$$

$$I(\hat{Y}; Y_r' | X_r U_1 X_1 Y_1 SQ) = \mathbb{E} \log \left[\det(\mathbf{A}_{21}) \right]$$
$$- \mathbb{E} \log \left[N_{r_1} N_{r_2} N_1 (1 + \|\mathbf{h}_1\|^2 P_2) \right].$$

Corollary 1. (Gaussian PDCF inner bound) *In Gaussian BRC with source-destination (SD) CSI feedback at relay, a rate pair (R_1, R_2) is achievable if $(21), (22), (23)$ are satisfied for some $P_r, P_{x_1}, P_{x_2}, P_{u_1}, P_{u_2}, N_1, N_2, N_{r_1}, N_{r_2} \ge 0$, such that $P_{x_1} + P_{x_2} + P_{u_1} + P_{u_2} \le P_s/2$ and $0 \le \mathbb{E} \log \left[\left(N_{r_1} N_{r_2} N_k [\|\mathbf{g}_k\|^2 \frac{P_r}{2} + \|\mathbf{h}_k\|^2 P_j + 1] \right) \det(\mathbf{A}_{2k}^{-1}) \right]$, for $k, j = 1, 2$ and $k \ne j$.*

IV. SCHEME ILLUSTRATION WITH NUMERICAL RESULTS

In this section, a symmetric Gaussian BRC with i.i.d. Rayleigh fading is considered. We adopt a simplified propagation model with pathloss $L = d^{-n_p}$, where d defines the dis-

$$\mathbf{A}_{1k} = \text{diag}(1, 1 + N_{r_k}, 1 + N_{r_j}, N_k) + P_{u_j} \mathcal{B} \left([\mathbf{h}_k, \mathbf{h}_{r_k}, \mathbf{h}_{r_j}, \mathbf{h}_k]^T \right) + P_{x_1, x_2} \mathcal{B} \left([\mathbf{h}_k, \mathbf{h}_{r_k}, \mathbf{h}_{r_j}, \mathbf{0}]^T \right), \quad (18)$$

$$\mathbf{A}_{2k} = \text{diag}(1, 1 + N_{r_k}, 1 + N_{r_j}, N_k) + P_{u_j} \mathcal{B} \left([\mathbf{h}_k, \mathbf{h}_{r_k}, \mathbf{h}_{r_j}, \mathbf{h}_k]^T \right) + P_{x_j} \mathcal{B} \left([\mathbf{h}_k, \mathbf{h}_{r_k}, \mathbf{h}_{r_j}, \mathbf{0}]^T \right), \quad (19)$$

$$\mathbf{A}_{3k} = \text{diag}(1, 1 + N_{r_k}, 1 + N_{r_j}, N_k, N_j) + P_{u_j} \mathcal{B} \left([\mathbf{h}_k, \mathbf{h}_{r_k}, \mathbf{h}_{r_j}, \mathbf{h}_k, \mathbf{0}]^T \right) + P_{u_k} \mathcal{B} \left([\mathbf{h}_k, \mathbf{h}_{r_k}, \mathbf{h}_{r_j}, \mathbf{0}, \mathbf{h}_j]^T \right)$$
$$+ P_{x_1, x_2} \mathcal{B} \left([\mathbf{h}_k, \mathbf{h}_{r_k}, \mathbf{h}_{r_j}, \mathbf{0}, \mathbf{0}]^T \right). \quad (20)$$

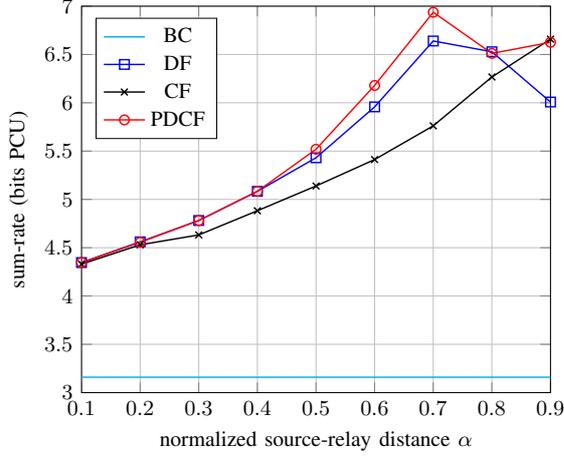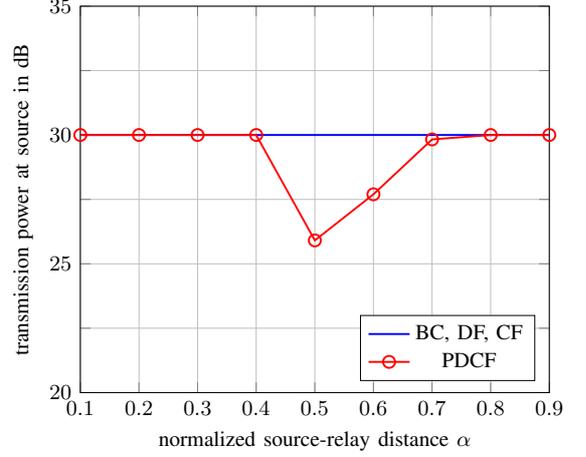Fig. 1. Sum-rate performance when $\mathrm{SNR} = 10\mathrm{dB}$.

Fig. 2. Optimal transmit power at source when $\mathrm{SNR} = 10\mathrm{dB}$.

tance between the communication nodes and n_p is the pathloss exponent. We take $n_p = 2$ as free space propagation. Hence, the fading channel coefficients are expressed as $\mathbf{H} \cdot \mathrm{L}_{sd}$, $\mathbf{H}_r \cdot \mathrm{L}_{sr}$, $\mathbf{G} \cdot \mathrm{L}_{rd}$, where $(\mathrm{L}_{sd}, \mathrm{L}_{sr}, \mathrm{L}_{rd})$ denote pathloss of (source-destination (SD), source-relay (SR), relay-destination (RD)) channels and (d_{sd}, d_{sr}, d_{rd}) are distance between (SD,SR,RD), respectively. For simplicity, the source, relay, and users lie on the same line where the two users are viewed as a single point in the deployment and yet have independent channels. Thus, source-relay-destination distance can be normalized as $(d_{sr} + d_{rd})/d_{sd} = 1$ and $d_{sd} = 10$ is chosen so that 9 integer-valued combinations of (d_{sr}, d_{rd}) are picked. We take the same power constraints ($P_s \leq P$ and $P_r \leq P$) over source and relay. Since the noises are normalized, we define $\mathrm{SNR} = P \cdot \mathrm{L}_{sd}$, which is the maximum received SNR at the users from the source. In the simulations, we only consider two

relaying modes in the proposed PDCF scheme. 1) The relay decodes completely the messages, reconstructs the interference components, and compresses them; the relay observation is not forwarded, i.e., we set (N_{r_1}, N_{r_2}) to infinity. 2) Instead of decoding message, the relay only compresses the observation. That is, it switches to the CF mode.

We compare our PDCF scheme with three reference schemes, namely, the broadcast channel (BC) without relay, the BRC with CF (Theorem 1), and the BRC with DF (Theorem 2). Note that only PDCF scheme exploits feedback while the other schemes do not. Moreover, all schemes are subject to no CSIT constraint. We consider two levels of received SNR, the medium to low level (10dB) and the high level (50dB).

The first observation from the simulation, although not explicit from the results, is that single-user transmission achieves the optimal sum-rate of each of the three reference schemes.

$$R_1 \leq \min\left\{\mathbb{E}\log\left(N_2^{-1}\det(\mathbf{A}_{31}\mathbf{A}_{21}^{-1})\right), \ \mathbb{E}\log\left[N_{r_1}N_{r_2}N_1\left(1 + \|\mathbf{h}_1\|^2\frac{P_s}{2} + \|\mathbf{g}_1\|^2\frac{P_r}{2}\right)\det(\mathbf{A}_{21}^{-1})\right],\right.$$
$$\mathbb{E}\log\det\left[\left(\mathbf{I} + \mathbf{H}_r\mathbf{H}_r^H P_{u_1}\left[\mathbf{I} + \mathbf{H}_r\mathbf{H}_r^H P_{x_1,x_2}\right]^{-1}\right)(\mathbf{A}_{11}\mathbf{A}_{21}^{-1})\right]$$
$$\left. + \min\left\{0, \mathbb{E}\log\left[N_{r_1}N_{r_2}N_1\left(1 + \|\mathbf{h}_1\|^2 P_{x_1,x_2} + \|\mathbf{g}_1\|^2\frac{P_r}{2}\right)\det(\mathbf{A}_{11}^{-1})\right]\right\}\right\}, \tag{21}$$

$$R_2 \leq \min\left\{\mathbb{E}\log\left(N_1^{-1}\det(\mathbf{A}_{32}\mathbf{A}_{22}^{-1})\right), \ \mathbb{E}\log\left[N_{r_1}N_{r_2}N_2\left(1 + \|\mathbf{h}_2\|^2\frac{P_s}{2} + \|\mathbf{g}_2\|^2\frac{P_r}{2}\right)\det(\mathbf{A}_{22}^{-1})\right],\right.$$
$$\mathbb{E}\log\det\left[\left(\mathbf{I} + \mathbf{H}_r\mathbf{H}_r^H P_{u_2}\left[\mathbf{I} + \mathbf{H}_r\mathbf{H}_r^H P_{x_1,x_2}\right]^{-1}\right)(\mathbf{A}_{12}\mathbf{A}_{22}^{-1})\right]$$
$$\left. + \min\left\{0, \mathbb{E}\log\left[N_{r_1}N_{r_2}N_2\left(1 + \|\mathbf{h}_2\|^2 P_{x_1,x_2} + \|\mathbf{g}_2\|^2\frac{P_r}{2}\right)\det(\mathbf{A}_{12}^{-1})\right]\right\}\right\}, \tag{22}$$

$$R_1 + R_2 \leq \mathbb{E}\log\det\left[\left(\mathbf{I} + \mathbf{H}_r\mathbf{H}_r^H P_{u_1,u_2}\left[\mathbf{I} + \mathbf{H}_r\mathbf{H}_r^H P_{x_1,x_2}\right]^{-1}\right)(\mathbf{A}_{11}\mathbf{A}_{21}^{-1})(\mathbf{A}_{12}\mathbf{A}_{22}^{-1})\right]$$
$$+ \min\left\{0, \mathbb{E}\log\left[N_{r_1}N_{r_2}N_1\left(1 + \|\mathbf{h}_1\|^2 P_{x_1,x_2} + \|\mathbf{g}_1\|^2\frac{P_r}{2}\right)\det(\mathbf{A}_{11}^{-1})\right]\right\}$$
$$+ \min\left\{0, \mathbb{E}\log\left[N_{r_1}N_{r_2}N_2\left(1 + \|\mathbf{h}_2\|^2 P_{x_1,x_2} + \|\mathbf{g}_2\|^2\frac{P_r}{2}\right)\det(\mathbf{A}_{12}^{-1})\right]\right\}. \tag{23}$$
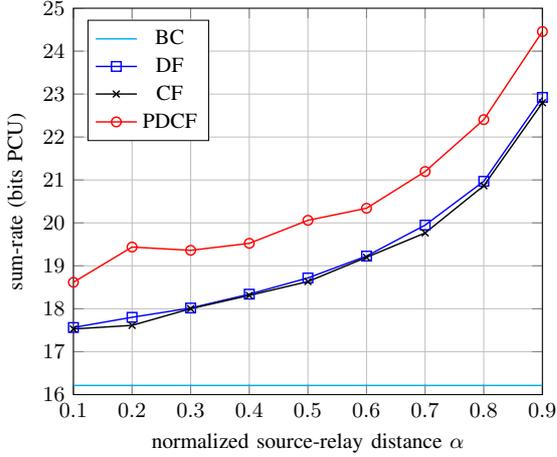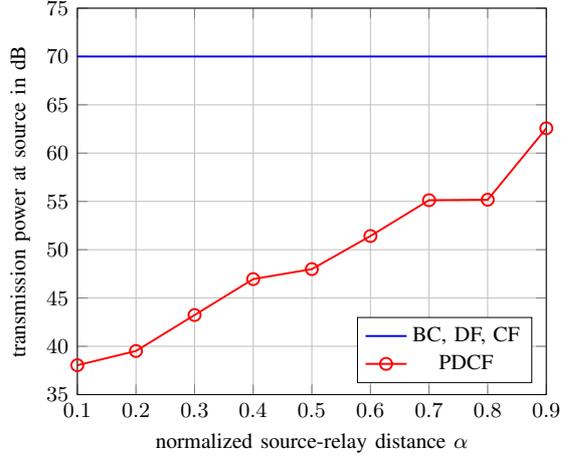
Fig. 3. Sum-rate performance when SNR = 50dB.

Fig. 4. Optimal transmit power at source when SNR = 50dB.

That is, no gain whatsoever could be obtained by serving more than one user at a time with these schemes. This is due to the fact that no CSIT is available and thus no downlink multiplexing gain can be exploited.

Then, let us look at the comparison in terms of sum-rate performance versus the relative relay position $\alpha = d_{sr}/d_{sd}$. At low SNR (Fig. 1), DF achieves peak rate at $\alpha = 0.7$ since the system consists of two multiple-input-single-output (MISO) SD channels, one multiple-input-multiple-output (MIMO) SR channel and two MISO RD channels. As α grows, throughput bottleneck switches from RD to SR channel due to the pathloss and the lack of CSI. Moreover, rate by CF increases monotonically with α. In fact, CF only compresses relay output so that the performance is less sensitive to SR channel. Interestingly, unlike in the static channel case, DF could not provide a performance gain when relay is close to the source. This is due to the lack of CSIT, which forbids the coherent transmission of the joint source-relay transmission once the message is decoded at the relay. The proposed scheme has strictly better performance for α between 0.4 and 0.8. At high SNR (Fig. 3), DF and CF have comparable performances and the sum-rate of all schemes increase with α. Indeed, numerical result shows clearly that the PDCF scheme, regardless of relay position, achieves non-negligible sum-rate gain over DF/CF schemes.

Finally, we examine the source power allocation of different schemes subject to the power constraints $P_s \leq P$. Fig. 2 and 4 are the power allocation counterparts of the sum-rate performances in Fig. 1 and 3, respectively. It is revealed that it is often better for the source not to transmit in its full power with our proposed scheme in contrast to conventional schemes. At low SNR (Fig. 2), we see that the source transmit power can be 4 dB lower than the maximum available power for the PDCF scheme for α around 0.5. At high SNR (Fig. 4), the same conclusion holds for all α and the gain can be up to 30 dB. Both Fig. 3 and Fig. 4 show that the PDCF scheme yields higher sum-rate performance with relatively less power consumed at source. The main reason is that, as

long as both users are served simultaneously, higher source power implies higher interference level at the receiver side. Therefore, the source must reduce its power to such a level that the corresponding interference could be resolved by the side-information sent from the relay. The amount of such side information is increased with the strength of RD channel, which explains the increasing source power with the decreasing relay-destination distance in Fig. 4. As a conclusion, we show that the transmit power of the macrocell network can be significantly reduced owing to the feedback from the users to the smallcell BS.

V. PROOF OF THEOREM 3

The original message pair (W_1, W_2) is split into $4 \times B$ mutual independent sub-messages $(m'_{1b}, m''_{1b}, m'_{2b}, m''_{2b})$ which are sent in $(B+L)$ n-length blocks, for $b = [1 : B]$, $k \in \{1, 2\}$, $L = O(B)$ and L is sufficiently large. At the last $(L-1)$ blocks, the last compress index is first decoded and then all compression indices and messages are decoded jointly and *backwardly*.

Code Generation: Given a pmf $p(\hat{y}|x_r, y'_r)p(y'_r|y_r, u_1, u_2, s)p(u_1)p(u_2)p(x_1|u_1)p(x_2|u_2)p(x_s|x_1, x_2)p(x_r)p(s)$ and the codebook is generated as follows.

1. Randomly generate $2^{nR'_k}$ independent sequences u_k^n, drawn i.i.d. from $P_{U_k^n}(u_k^n) = \prod_{i=1}^n P_{U_k}(u_{ki})$ and index them as $u_k^n(m'_k)$ with $m'_k \in \{1, 2, ..., 2^{nR'_k}\}$.
2. For every sequence u_k^n generated, randomly generate $2^{nR''_k}$ conditionally independent sequences x_k^n, drawn i.i.d. from $P_{X_k^n|U_k^n}(x_k^n|u_k^n) = \prod_{i=1}^n P_{X_k|U_k}(x_{ki}|u_{ki})$ and index them as $x_k^n(m''_k, m'_k)$ with $m''_k \in \{1, 2, ..., 2^{nR''_k}\}$.
3. Randomly and independently generate $2^{n\hat{R}}$ sequences x_r^n drawn i.i.d. from $P_{X_r^n}(x_r^n) = \prod_{i=1}^n P_{X_r}(x_{ri})$ and index them as $x_r^n(r)$ with $r \in \{1, 2, ..., 2^{n\hat{R}}\}$.
4. For each sequence x_r^n generated, randomly generate $2^{n\hat{R}}$ conditionally independent sequences $\hat{y}^n$ drawn i.i.d from $P_{\hat{Y}^n|X_r^n}(\hat{y}^n|x_r^n) = \prod_{i=1}^n P_{\hat{Y}|X_r}(\hat{y}_i|x_{ri})$ and index them as $\hat{y}^n(l, l')$ with $l, l' \in \{1, 2, ..., 2^{n\hat{R}}\}$.

5. Provide the codebook to the decoders, respectively.

Encoding

- For every block $b = [1 : B]$, the source conveys messages $(m''_{1b}, m''_{2b}, m'_{1b}, m'_{2b})$ based on $x_1^n(m''_{1b}, m'_{1b}), x_2^n(m''_{2b}, m'_{2b})$. For $b \in [B + 1 : B + L]$, dummy messages $(m'_{1b}, m'_{2b}, m''_{1b}, m''_{2b}) = (1, 1, 1, 1)$ are chosen.
- At block b, source encoder picks and sends $x_s(b)$ such that $x_s(b) \in \mathcal{T}_\delta^n \big(X_s | X_1^n(m''_{1b}, m'_{1b}), X_2^n(m''_{2b}, m'_{2b})\big)$.
- For each block $b = [1 : B]$, after receiving state sequence $s^n(b)$ and recovering the messages (m'_{1b}, m'_{2b}), relay encoder searches for at least one index l_b such that $\big(u_1^n(m'_{1b}), u_2^n(m'_{2b}), \hat{y}^n(l_b, l_{b-1}), x_r^n(l_{b-1}), y_r^n(b), s^n(b)\big)$ are jointly typical, also expressed as $(y_r'^n(b), \hat{y}^n(l_b, l_{b-1}), x_r^n(l_{b-1})) \in \mathcal{T}_\delta^n(\hat{Y} X_r Y_r')$, where $y_r' \triangleq (y_r, u_1, u_2, s)$ and $l_0 = 1$. The probability that such l_b exists goes to one as n tends to infinity if

$$\hat{R} \geq I(\hat{Y}; Y_r' | X_r S). \tag{24}$$

At block $b + 1$, relay sends $x_r^n(l_b)$. At last $L - 1$ blocks, relay repeats l_{B+1}.

Decoding and error events analysis

1) Before block $b + 1$, $(\hat{m}'_{1b}, \hat{m}'_{2b})$ are decoded at the relay such that $\big(u_1^n(\hat{m}'_{1b}), u_2^n(\hat{m}'_{2b}), x_r^n(l_b), y_r^n(b), s^n(b)\big) \in \mathcal{T}_\delta^n(U_1 U_2 X_r Y_r S)$. Thus, the error probability of decoding (m'_{1b}, m'_{2b}) goes to zero provided that,

$$R'_1 \leq I(U_1; Y_r | X_r U_2 S), \tag{25}$$
$$R'_2 \leq I(U_2; Y_r | X_r U_1 S), \tag{26}$$
$$R'_1 + R'_2 \leq I(U_1 U_2; Y_r | X_r S). \tag{27}$$

2) At destination, decoder awaits until the end of block $B + L$ and then receiver k searches for index $\hat{l}_{B+1}$ such that $\big(x_r^n(\hat{l}_{B+1}), x_1^n(1, 1), x_2^n(1, 1), y_k^n(b), s^n(b)\big) \in \mathcal{T}_\delta^n(X_r X_1 X_2 Y_k S)$ for all $b \in [B + 2 : B + L]$. $\hat{l}_{B+1}$ is obtained with error probability goes to zero provided,

$$\hat{R} \leq (L - 1) I(X_r; X_1 X_2 Y_k | S). \tag{28}$$

3) After the receivers get $\hat{l}_{B+1}$, destination k decodes the indices $(\hat{l}_b, \hat{m}'_{k(b+1)}, \hat{m}''_{k(b+1)})$ knowing $\hat{l}_{b+1}$, such that

$$\big(\hat{y}^n(\hat{l}_{b+1}, \hat{l}_b), u_k^n(\hat{m}'_{k(b+1)}), x_k^n(\hat{m}''_{k(b+1)}, \hat{m}'_{k(b+1)}),$$
$$x_r^n(\hat{l}_b), y_k^n(b+1), s^n(b+1)\big) \in \mathcal{T}_\delta^n(\hat{Y} U_k X_k X_r Y_k S).$$

Error events are analyzed as below to ensure the error probability for this decoding tends to zero, take $k = 1$,

a. The error event such that $(\hat{l}_b = l_b, \hat{m}''_{1(b+1)} \neq m''_{1(b+1)}, \hat{m}'_{1(b+1)} = m'_{1(b+1)})$ could be avoided if

$$R''_1 \leq I(X_1; Y_1 \hat{Y} | X_r U_1 S) \tag{29}$$

b. The error event such that $(\hat{l}_b = l_b, \hat{m}''_{1(b+1)} \neq m''_{1(b+1)}, \hat{m}'_{1(b+1)} \neq m'_{1(b+1)})$ could be avoided if

$$R'_1 + R''_1 \leq I(U_1 X_1; Y_1 \hat{Y} | X_r S) \tag{30}$$

c. The error event such that $(\hat{l}_b \neq l_b, \hat{m}''_{1(b+1)} = m''_{1(b+1)}, \hat{m}'_{1(b+1)} = m'_{1(b+1)})$ could be avoided if

$$\hat{R} \leq I(\hat{Y} X_r; Y_1 U_1 X_1 | S) \tag{31}$$

d. The error event such that $(\hat{l}_b \neq l_b, \hat{m}''_{1(b+1)} \neq m''_{1(b+1)}, \hat{m}'_{1(b+1)} = m'_{1(b+1)})$ could be avoided if

$$R''_1 + \hat{R} \leq I(X_1; Y_1 \hat{Y} | X_r U_1 S) + I(X_r; Y_1 | U_1 S)$$
$$+ I(\hat{Y}; U_1 Y_1 | X_r) \tag{32}$$

e. The error event such that $(\hat{l}_b \neq l_b, \hat{m}''_{1(b+1)} \neq m''_{1(b+1)}, \hat{m}'_{1(b+1)} \neq m'_{1(b+1)})$ could be avoided if

$$R'_1 + R''_1 + \hat{R} \leq I(U_1 X_1; Y_1 \hat{Y} | X_r S) + I(X_r; Y_1 | S)$$
$$+ I(\hat{Y}; Y_1 | X_r S) \tag{33}$$

4) We first remark that constraint (28) is neglectable as L goes to infinity. Moreover, we combine (24) and (31) that yields constraint (9). Indeed, we can obtain the constraints for user 2 by exchanging indices of (29)-(33). Then, we use some algebra to combine (24)-(27), (29)-(33) and the corresponding symmetric constraints for user 2. Finally, the proof for achievable rate region defined in Theorem 3 completes by adding a time-sharing auxiliary variable Q.

REFERENCES

[1] T. M. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 5, pp. 572–584, Sept. 1979.
[2] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1691–1706, July. 2003.
[3] 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.
[4] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, "Multiuser MIMO achievable rates with downlink training and channel state feedback," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2845–2866, June. 2010.
[5] A. G. Davoodi and S. A. Jafar, "Aligned image sets under channel uncertainty: Settling a conjecture by Lapidoth, Shamai and Wigger on the collapse of degrees of freedom under finite precision CSIT," *arXiv preprint arXiv:1403.1541*, 2014.
[6] M. A. Maddah-Ali and D. Tse, "Completely stale transmitter channel state information is still very useful," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4418–4431, July. 2012.
[7] 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, 2012.
[8] R. Tandon, S. A. Jafar, S. Shamai Shitz, and H. V. Poor, "On the synergistic benefits of alternating CSIT for the MISO BC," *IEEE Trans. Inf. Theory*, vol. 59, no. 7, pp. 4106–4128, July. 2013.
[9] X. Yi, S. Yang, D. Gesbert, and M. Kobayashi, "The degrees of freedom region of temporally-correlated MIMO networks with delayed CSIT," *IEEE Trans. Inf. Theory*, vol. 60, no. 1, pp. 494 – 514, Jan. 2014.
[10] J. Chen and P. Elia, "Toward the performance versus feedback tradeoff for the two-user MISO broadcast channel," *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 8336–8356, Dec. 2013.
[11] C. He, S. Yang, and P. Piantanida, "On the capacity of the fading broadcast channel with state feedback," in *IEEE Int. Symp. Communications, Control and Signal Processing (ISCCSP)*, May. 2014.
[12] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3037–3063, Sept. 2005.
[13] A. Behboodi and P. Piantanida, "Cooperative strategies for simultaneous and broadcast relay channels," *IEEE Trans. Inf. Theory*, vol. 59, no. 3, pp. 1417–1443, Mar. 2013.