Degrees-of-Freedom Region of the MISO Broadcast Channel with General Mixed-CSIT

Jinyuan Chen and Petros Elia
Mobile Communications Department
EURECOM
Sophia Antipolis, France
Email: {chenji,elia}@eurecom.fr

Abstract—In the setting of the two-user broadcast channel, recent work by Maddah-Ali and Tse has shown that knowledge of prior channel state information at the transmitter (CSIT) can be useful, even in the absence of any knowledge of current CSIT. Very recent work by Kobayashi et al., Yang et al., and Gou and Jafar, extended this to the case where, instead of no current CSIT knowledge, the transmitter has partial knowledge, and where under a symmetry assumption, the quality of this knowledge is identical for the different users' channels.

Motivated by the fact that in multiuser settings, the quality of CSIT feedback may vary across different links, we here generalize the above results to the natural setting where the current CSIT quality varies for different users' channels. For this setting we derive the optimal degrees-of-freedom (DoF) region, and provide novel multi-phase broadcast schemes that achieve this optimal region. Finally this generalization incorporates and generalizes the corresponding result in Maleki et al. which considered the broadcast channel with one user having perfect CSIT and the other only having prior CSIT.

I. INTRODUCTION

In many multiuser wireless communications scenarios, having sufficient CSIT is a crucial ingredient that facilitates improved performance. While being useful, perfect CSIT is also hard and time-consuming to obtain, hence the need for communication schemes that can utilize partial or delayed CSIT knowledge (see [1]–[6]). In this context of multiuser communications, we here consider the broadcast channel (BC), and specifically focus on the two-user multiple-input single-output (MISO) BC, where a two-antenna transmitter communicates to two single-antenna receivers. In this setting, the channel model takes the form

$$y_t^{(1)} = \boldsymbol{h}_t^\mathsf{T} \boldsymbol{x}_t + z_t^{(1)} \tag{1a}$$
$$y_t^{(2)} = \boldsymbol{g}_t^\mathsf{T} \boldsymbol{x}_t + z_t^{(2)}, \tag{1b}$$

where for any time instant t, $\boldsymbol{h}_t, \boldsymbol{g}_t \in \mathbb{C}^{2\times 1}$ represent the channel vectors for user 1 and 2 respectively, where $z_t^{(1)}, z_t^{(2)}$ represent unit power AWGN noise, where $\boldsymbol{x}_t$ is the input signal with power constraint $\mathbb{E}\left(\|\boldsymbol{x}_t\|^2\right) \leq P$, and where in this case, P also takes the role of the signal-to-noise ratio (SNR). It is well known that in this setting, the presence of full CSIT allows for the optimal 1 degree-of-freedom (DoF) per user, whereas the complete absence of CSIT causes a substantial

degradation to just $1/2$ DoF per user[1].

An interesting scheme that bridges this performance gap by utilizing partial CSIT knowledge, was recently presented in [1] which showed that delayed CSIT knowledge can still be useful in improving the DoF region of the broadcast channel. In the above described two-user MISO BC setting, and under the assumption that at time t, the transmitter knows the delayed channel states $(\boldsymbol{h}, \boldsymbol{g})$ up to time $t-1$, the work in [1] showed that each user can achieve $2/3$ DoF, providing a clear improvement over the case of no CSIT.

This result was later generalized in [7]–[9] which considered the natural extension where, in addition to the aforementioned perfect knowledge of prior CSIT, the transmitter also had imperfect knowledge of current CSIT; at time t the transmitter had estimates $\hat{\boldsymbol{h}}_t, \hat{\boldsymbol{g}}_t$ of $\boldsymbol{h}_t$ and $\boldsymbol{g}_t$, with estimation errors

$$\tilde{\boldsymbol{h}}_t = \boldsymbol{h}_t - \hat{\boldsymbol{h}}_t, \quad \tilde{\boldsymbol{g}}_t = \boldsymbol{g}_t - \hat{\boldsymbol{g}}_t \tag{2}$$

having i.i.d. Gaussian entries with power

$$\frac{1}{2}\mathbb{E}\left(\|\tilde{\boldsymbol{h}}_t\|^2\right) = \frac{1}{2}\mathbb{E}\left(\|\tilde{\boldsymbol{g}}_t\|^2\right) = P^{-\alpha},$$

for some non-negative parameter α that described the quality of the estimate of the current CSIT. In this setting of 'mixed' CSIT (perfect prior CSIT and imperfect current CSIT), and for d_1, d_2 denoting the DoF for the first and second user over the aforementioned two-user BC, the work in [7]–[9] showed the optimal DoF region to take the form,

$$\{d_1 \leq 1; \; d_2 \leq 1; \; 2d_1 + d_2 \leq 2 + \alpha; \; 2d_2 + d_1 \leq 2 + \alpha\} \tag{3}$$

corresponding to a polygon with corner points $\{(0,0), (1,0), (1,\alpha), (\frac{2+\alpha}{3}, \frac{2+\alpha}{3}), (\alpha,1), (0,1)\}$, nicely bridging the gap between the case of $\alpha = 0$ explored in [1], and the case of $\alpha = 1$ (and naturally $\alpha > 1$) corresponding to perfect CSIT.

A. Notation and conventions

Throughout this paper, $(\bullet)^{-1}$, $(\bullet)^\mathsf{T}$, $(\bullet)^\mathsf{H}$, respectively denote the inverse, transpose, and conjugate transpose of a matrix, while $(\bullet)^*$ denotes the complex conjugate, and $\|\bullet\|$ denotes the Euclidean norm. $|\bullet|$ denotes the magnitude of a scalar,

---

[1]We remind the reader that for an achievable rate pair (R_1, R_2), the corresponding DoF pair (d_1, d_2) is given by $d_i = \lim_{P\to\infty} \frac{R_i}{\log P}$, $i = 1, 2$. The corresponding DoF region is then the set of all achievable DoF pairs.

and diag($\bullet$) denotes a diagonal matrix. Logarithms are of base 2. $o(\bullet)$ comes from the standard Landau notation, where $f(x) = o(g(x))$ implies $\lim_{x\to\infty} f(x)/g(x) = 0$. We also use $\doteq$ to denote *exponential equality*, i.e., we write $f(P) \doteq P^B$ to denote $\lim_{P\to\infty} \frac{\log f(P)}{\log P} = B$. Finally, in the spirit of [7]–[9] we consider a unit coherence period, as well as perfect knowledge of channel state information at the receivers (perfect CSIR).

II. THE GENERALIZED MIXED-CSIT BROADCAST CHANNEL

Motivated by the fact that in multiuser settings, the quality of CSIT feedback may vary across different links, we extend the approach in [7]–[9] to consider unequal quality of current CSIT knowledge for h_t and g_t. Specifically under the same set of assumptions mentioned above, and in the presence of perfect prior CSIT, we now consider the case where at time t, the transmitter has estimates $\hat{h}_t, \hat{g}_t$ of the current h_t and g_t, with estimation errors

$$\tilde{h}_t = h_t - \hat{h}_t, \quad \tilde{g}_t = g_t - \hat{g}_t \qquad (4)$$

having i.i.d. Gaussian entries with power

$$\frac{1}{2}\mathbb{E}\left(\|\tilde{h}_t\|^2\right) = P^{-\alpha_1}, \quad \frac{1}{2}\mathbb{E}\left(\|\tilde{g}_t\|^2\right) = P^{-\alpha_2},$$

for some non-negative parameters α_1, α_2 that describe the generally unequal quality of the estimates of the current CSIT for the two users' links.

We proceed to describe the optimal DoF region of the general mixed-CSIT two-user MISO BC (two-antenna transmitter). The optimal schemes are presented in Section III, parts of the proof of the schemes' performance are presented in Appendix V, while the outer bound proof is placed in Appendix VI.

A. DoF region of the MISO BC with generalized mixed-CSIT

Without loss of generality, the rest of this work assumes that

$$1 \geq \alpha_1 \geq \alpha_2 \geq 0. \qquad (5)$$

Theorem 1: The DoF region of the two-user MISO BC with general mixed-CSIT, is given by

$$d_1 \leq 1, \qquad d_2 \leq 1 \qquad (6a)$$
$$2d_1 + d_2 \leq 2 + \alpha_1 \qquad (6b)$$
$$d_1 + 2d_2 \leq 2 + \alpha_2 \qquad (6c)$$

where the region is a polygon which, for $2\alpha_1 - \alpha_2 < 1$ has corner points

$$\{(0,0), (1,0), (1,\alpha_1), (\frac{2+2\alpha_1-\alpha_2}{3}, \frac{2+2\alpha_2-\alpha_1}{3}), (\alpha_2,1), (0,1)\},$$

and otherwise has corner points

$$\{(0,0), (1,0), (1, \frac{1+\alpha_2}{2}), (\alpha_2, 1), (0,1)\}.$$

The above corner points, and consequently the entire DoF inner bound, will be attained by the schemes to be described later on. The result generalizes the results in [7]–[9] as well as the result in [10] which considered the case of ($\alpha_1 = 1, \alpha_2 =$
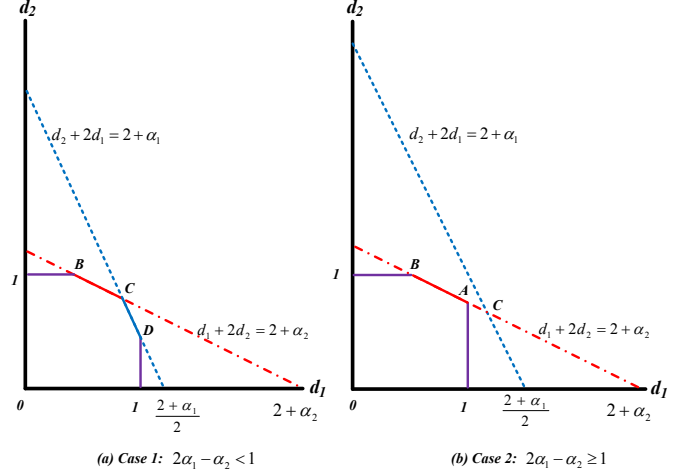

Fig. 1. DoF region when $2\alpha_1 - \alpha_2 < 1$ (case 1) and when $2\alpha_1 - \alpha_2 \geq 1$ (case 2). The corner points take the following values: $A = (1, \frac{1+\alpha_2}{2})$, $B = (\alpha_2, 1)$, $C = (\frac{2+2\alpha_1-\alpha_2}{3}, \frac{2+2\alpha_2-\alpha_1}{3})$ and $D = (1, \alpha_1)$.

0), where one user had perfect CSIT and the other only prior CSIT.

Figure 1 depicts the general DoF region for the case where $2\alpha_1 - \alpha_2 < 1$ (case 1) and the case where $2\alpha_1 - \alpha_2 \geq 1$ (case 2).

We proceed to describe the communication schemes.

III. DESIGN OF COMMUNICATION SCHEMES FOR THE TWO-USER GENERAL MIXED-CSIT MISO BC

As stated, without loss of generality, we assume that $1 \geq \alpha_1 \geq \alpha_2 \geq 0$. We describe the three schemes $\mathcal{X}_1, \mathcal{X}_2$ and $\mathcal{X}_3$ that achieve the optimal DoF region (in conjunction with time-division between these same schemes). Specifically scheme $\mathcal{X}_1$ achieves $C = (\frac{2+2\alpha_1-\alpha_2}{3}, \frac{2+2\alpha_2-\alpha_1}{3})$ (case 1), scheme $\mathcal{X}_2$ achieves DoF points $D = (1, \alpha_1)$ (case 1) and $A = (1, \frac{1+\alpha_2}{2})$ (case 2), and scheme $\mathcal{X}_3$ achieves $B = (\alpha_2, 1)$ (case 1 and case 2). The scheme description is done for $1 > \alpha_1 > \alpha_2 \geq 0$, and for rational α_1, α_2. The cases where $\alpha_1 = 1$, or $\alpha_1 = \alpha_2$, or where α_1, α_2 are not rational, can be readily handled with minor modifications. We proceed to describe the basic notation and conventions used in our schemes.

The schemes are designed with S phases (S varies from scheme to scheme), where the sth phase consists of T_s channel uses, $s = 1, 2, \cdots, S$. The vectors $h_{s,t}$ and $g_{s,t}$ will denote the channel vectors seen by the first and second user respectively during timeslot t of phase s, while $\hat{h}_{s,t}$ and $\hat{g}_{s,t}$ will denote the estimates of these channels at the transmitter during the same time, and $\tilde{h}_{s,t} = h_{s,t} - \hat{h}_{s,t}$, $\tilde{g}_{s,t} = g_{s,t} - \hat{g}_{s,t}$ will denote the estimation errors.

Furthermore $a_{s,t}$ and $a'_{s,t}$ will denote the independent information symbols that may be sent during phase-s, timeslot-t, and which are meant for user 1, while symbols $b_{s,t}$ and $b'_{s,t}$ are meant for user 2. Vectors $u_{s,t}$ and $v_{s,t}$ are the unit-norm beamformers for $a_{s,t}$ and $b_{s,t}$ respectively, chosen so that $u_{s,t}$ is orthogonal to $\hat{g}_{s,t}$, and so that $v_{s,t}$ is orthogonal to

$\hat{\boldsymbol{h}}_{s,t}$. Furthermore $\boldsymbol{u}'_{s,t}, \boldsymbol{v}'_{s,t}$ are the randomly chosen unit-norm beamformers for $a'_{s,t}$ and $b'_{s,t}$ respectively.

Another notation that will be shared between schemes includes

$$\bar{c}^{(b)}_{s,t} \triangleq \tilde{\boldsymbol{h}}^{\mathsf{T}}_{s,t}\boldsymbol{v}_{s,t}b_{s,t} + \boldsymbol{h}^{\mathsf{T}}_{s,t}\boldsymbol{v}'_{s,t}b'_{s,t},$$
$$\bar{c}^{(a)}_{s,t} \triangleq \tilde{\boldsymbol{g}}^{\mathsf{T}}_{s,t}\boldsymbol{u}_{s,t}a_{s,t} + \boldsymbol{g}^{\mathsf{T}}_{s,t}\boldsymbol{u}'_{s,t}a'_{s,t}, \quad t = 1, \cdots, T_s \quad (7)$$

that denotes the interference seen by user 1 and user 2 respectively, during timeslot t of phase s. For $\{\bar{c}^{(a)}_{s,t}, \bar{c}^{(b)}_{s,t}\}^{T_s}_{t=1}$ being the accumulated interference to both users during phase s, we will let $\{\hat{c}^{(a)}_{s,t}, \hat{c}^{(b)}_{s,t}\}^{T_s}_{t=1}$ be a quantized version of $\{\bar{c}^{(a)}_{s,t}, \bar{c}^{(b)}_{s,t}\}^{T_s}_{t=1}$, and we will consider the mapping where the total information in $\{\hat{c}^{(a)}_{s,t}, \hat{c}^{(b)}_{s,t}\}^{T_s}_{t=1}$ is split evenly across symbols $\{c_{s+1,t}\}^{T_{s+1}}_{t=1}$ transmitted during the next phase. In addition we use $\boldsymbol{w}_{s+1,t}$ to denote the randomly chosen unit-norm beamformer of $c_{s+1,t}$.

Furthermore, unless stated otherwise,

$$\boldsymbol{x}_{s,t} = \boldsymbol{w}_{s,t}\underbrace{c_{s,t}}_{P^{(c)}_s} + \boldsymbol{u}_{s,t}\underbrace{a_{s,t}}_{P^{(a)}_s} + \boldsymbol{u}'_{s,t}\underbrace{a'_{s,t}}_{P^{(a')}_s} + \boldsymbol{v}_{s,t}\underbrace{b_{s,t}}_{P^{(b)}_s} + \boldsymbol{v}'_{s,t}\underbrace{b'_{s,t}}_{P^{(b')}_s} \quad (8)$$

will be the general form of the transmitted vector at timeslot t of phase s. As noted above under each summand, the average power that is assigned to each symbol, throughout a specific phase, will be denoted as follows:

$$P^{(c)}_s \triangleq \mathbb{E}|c_{s,t}|^2, \quad P^{(a)}_s \triangleq \mathbb{E}|a_{s,t}|^2, \quad P^{(a')}_s \triangleq \mathbb{E}|a'_{s,t}|^2$$
$$P^{(b)}_s \triangleq \mathbb{E}|b_{s,t}|^2, \quad P^{(b')}_s \triangleq \mathbb{E}|b'_{s,t}|^2.$$

Furthermore each of the above symbols carries a certain amount of information, per timeslot, where this amount may vary across different phases. Specifically we use $r^{(a)}_s$ to mean that, during phase s, each symbol $a_{s,t}$, $t = 1, \cdots, T_s$, carries $r^{(a)}_s \log P + o(\log P)$ bits. Similarly we use $r^{(a')}_s, r^{(b)}_s, r^{(b')}_s, r^{(c)}_s$ to describe the prelog factor of the number of bits in $a'_{s,t}, b_{s,t}, b'_{s,t}, c_{s,t}$ respectively, again for phase s.

Finally the received signals during phase s for the first and second user, are respectively denoted as $y^{(1)}_{s,t}$ and $y^{(2)}_{s,t}$, where generally the signals take the following form

$$y^{(1)}_{s,t} = \boldsymbol{h}^{\mathsf{T}}_{s,t}x_{s,t} + z^{(1)}_{s,t},$$
$$y^{(2)}_{s,t} = \boldsymbol{g}^{\mathsf{T}}_{s,t}x_{s,t} + z^{(2)}_{s,t}, \quad t = 1, \cdots, T_s. \quad (9)$$

A. Scheme $\mathcal{X}_1$ achieving $C = \left(\frac{2+2\alpha_1-\alpha_2}{3}, \frac{2+2\alpha_2-\alpha_1}{3}\right)$ (case 1)

As stated, scheme $\mathcal{X}_1$ has S phases, where the phase durations $T_1, T_2, \cdots, T_S$ are chosen to be integers such that

$$T_2 = T_1\xi, \quad T_s = T_{s-1}\mu = T_1\xi\mu^{s-2}, \forall s \in \{3, 4, \cdots, S-1\},$$
$$T_S = T_{S-1}\gamma = T_1\xi\mu^{S-3}\gamma, \quad (10)$$

where $\xi = \frac{2-\alpha_1-\alpha_2}{1-\alpha_1-\Delta}$, $\mu = \frac{\alpha_1-\alpha_2+2\Delta}{1-\alpha_1-\Delta}$, $\gamma = \frac{\alpha_1-\alpha_2+2\Delta}{1-\alpha_2}$, and where Δ is any constant such that $0 < \Delta < \frac{1-2\alpha_1+\alpha_2}{3}$.

1) Phase 1: During phase 1 (T_1 channel uses), the transmit signal is

$$\boldsymbol{x}_{1,t} = \boldsymbol{u}_{1,t}a_{1,t} + \boldsymbol{u}'_{1,t}a'_{1,t} + \boldsymbol{v}_{1,t}b_{1,t} + \boldsymbol{v}'_{1,t}b'_{1,t}, \quad (11)$$

while the power and rate are set as

$$P^{(a)}_1 \doteq P, \quad P^{(a')}_1 \doteq P^{1-\alpha_2}, \quad P^{(b)}_1 \doteq P, \quad P^{(b')}_1 \doteq P^{1-\alpha_1}$$
$$r^{(a)}_1 = 1, \quad r^{(a')}_1 = 1-\alpha_2, \quad r^{(b)}_1 = 1, \quad r^{(b')}_1 = 1-\alpha_1. \quad (12)$$

The received signals at the two users then take the form

$$y^{(1)}_{1,t} = \underbrace{\boldsymbol{h}^{\mathsf{T}}_{1,t}\boldsymbol{u}_{1,t}a_{1,t}}_{P} + \underbrace{\boldsymbol{h}^{\mathsf{T}}_{1,t}\boldsymbol{u}'_{1,t}a'_{1,t}}_{P^{1-\alpha_2}} + \overbrace{\underbrace{\tilde{\boldsymbol{h}}^{\mathsf{T}}_{1,t}\boldsymbol{v}_{1,t}b_{1,t}}_{P^{1-\alpha_1}} + \underbrace{\boldsymbol{h}^{\mathsf{T}}_{1,t}\boldsymbol{v}'_{1,t}b'_{1,t}}_{P^{1-\alpha_1}}}^{\bar{c}^{(b)}_{1,t}} + \underbrace{z^{(1)}_{1,t}}_{P^0},$$

$$y^{(2)}_{1,t} = \overbrace{\underbrace{\tilde{\boldsymbol{g}}^{\mathsf{T}}_{1,t}\boldsymbol{u}_{1,t}a_{1,t}}_{P^{1-\alpha_2}} + \underbrace{\boldsymbol{g}^{\mathsf{T}}_{1,t}\boldsymbol{u}'_{1,t}a'_{1,t}}_{P^{1-\alpha_2}}}^{\bar{c}^{(a)}_{1,t}} + \underbrace{\boldsymbol{g}^{\mathsf{T}}_{1,t}\boldsymbol{v}_{1,t}b_{1,t}}_{P} + \underbrace{\boldsymbol{g}^{\mathsf{T}}_{1,t}\boldsymbol{v}'_{1,t}b'_{1,t}}_{P^{1-\alpha_1}} + \underbrace{z^{(2)}_{1,t}}_{P^0}, \quad (13)$$

where under each term we noted the order of the summand's average power.

At this point, and after the end of the first phase, the transmitter can use its knowledge of delayed CSIT to reconstruct $\{\bar{c}^{(a)}_{1,t}, \bar{c}^{(b)}_{1,t}\}^{T_1}_{t=1}$ (cf.(7)), and quantize each term as

$$\bar{c}^{(a)}_{1,t} = \hat{c}^{(a)}_{1,t} + \tilde{c}^{(a)}_{1,t}, \quad \bar{c}^{(b)}_{1,t} = \hat{c}^{(b)}_{1,t} + \tilde{c}^{(b)}_{1,t}, \quad t = 1, 2, \cdots, T_1,$$

where $\hat{c}^{(a)}_{1,t}, \hat{c}^{(b)}_{1,t}$ are the quantized values, and where $\tilde{c}^{(a)}_{1,t}, \tilde{c}^{(b)}_{1,t}$ are the quantization errors. Noting that $\mathbb{E}|\bar{c}^{(a)}_{1,t}|^2 \doteq P^{1-\alpha_2}$, $\mathbb{E}|\bar{c}^{(b)}_{1,t}|^2 \doteq P^{1-\alpha_1}$, we choose a quantization rate that assigns each $\hat{c}^{(a)}_{1,t}$ a total of $(1-\alpha_2)\log P + o(\log P)$ bits, and each $\hat{c}^{(b)}_{1,t}$ a total of $(1-\alpha_1)\log P + o(\log P)$ bits, thus allowing for $\mathbb{E}|\tilde{c}^{(a)}_{1,t}|^2 \doteq \mathbb{E}|\tilde{c}^{(b)}_{1,t}|^2 \doteq 1$ ([11]). At this point the $T_1(2-\alpha_1-\alpha_2)\log P + o(\log P)$ bits representing $\{\hat{c}^{(a)}_{1,t}, \hat{c}^{(b)}_{1,t}\}^{T_1}_{t=1}$, are distributed evenly across the set $\{c_{2,t}\}^{T_2}_{t=1}$ which will be sequentially transmitted during the next phase. This transmission of $\{c_{2,t}\}^{T_2}_{t=1}$ will help each of the users cancel the interference from the other user, and it will also serve as an extra observation that allows for decoding of all private information of that same user.

2) Phase 2: During phase 2 (T_2 channel uses), the transmit signal takes the exact form in (8)

$$\boldsymbol{x}_{2,t} = \boldsymbol{w}_{2,t}c_{2,t} + \boldsymbol{u}_{2,t}a_{2,t} + \boldsymbol{u}'_{2,t}a'_{2,t} + \boldsymbol{v}_{2,t}b_{2,t} + \boldsymbol{v}'_{2,t}b'_{2,t} \quad (14)$$

where we set power and rate as

$$\begin{aligned}
P^{(c)}_2 &\doteq P, & r^{(c)}_2 &= 1-\alpha_1-\Delta \\
P^{(a)}_2 &\doteq P^{\alpha_1+\Delta}, & r^{(a)}_2 &= \alpha_1+\Delta \\
P^{(a')}_2 &\doteq P^{\alpha_1-\alpha_2+\Delta}, & r^{(a')}_2 &= \alpha_1-\alpha_2+\Delta \quad (15) \\
P^{(b)}_2 &\doteq P^{\alpha_1+\Delta}, & r^{(b)}_2 &= \alpha_1+\Delta \\
P^{(b')}_2 &\doteq P^{\Delta}, & r^{(b')}_2 &= \Delta,
\end{aligned}$$

and where we note that $r^{(c)}_2$ satisfies $T_2 r^{(c)}_2 = T_1(2-\alpha_1-\alpha_2)$.
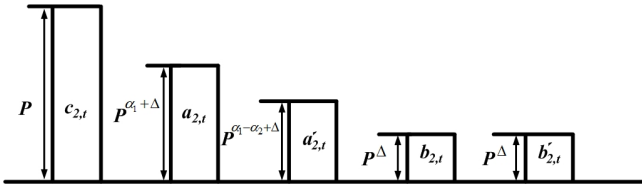
Fig. 2. Received power levels at user 1 (phase 2).

The received signals during this phase are given as

$$y_{2,t}^{(1)}=\underbrace{\boldsymbol{h}_{2,t}^{\mathsf{T}}\boldsymbol{w}_{2,t}c_{2,t}}_{P}+\underbrace{\boldsymbol{h}_{2,t}^{\mathsf{T}}\boldsymbol{u}_{2,t}a_{2,t}}_{P^{\alpha_1+\Delta}}+\underbrace{\boldsymbol{h}_{2,t}^{\mathsf{T}}\boldsymbol{u}_{2,t}^{'}a_{2,t}^{'}}_{P^{\alpha_1-\alpha_2+\Delta}}$$
$$+\underbrace{\tilde{\boldsymbol{h}}_{2,t}^{\mathsf{T}}\boldsymbol{v}_{2,t}b_{2,t}}_{P^{\Delta}}+\underbrace{\boldsymbol{h}_{2,t}^{\mathsf{T}}\boldsymbol{v}_{2,t}^{'}b_{2,t}^{'}}_{P^{\Delta}}+\underbrace{z_{2,t}^{(1)}}_{P^0}, \tag{16}$$

$$y_{2,t}^{(2)}=\underbrace{\boldsymbol{g}_{2,t}^{\mathsf{T}}\boldsymbol{w}_{2,t}c_{2,t}}_{P}+\underbrace{\tilde{\boldsymbol{g}}_{2,t}^{\mathsf{T}}\boldsymbol{u}_{2,t}a_{2,t}}_{P^{\alpha_1-\alpha_2+\Delta}}+\underbrace{\boldsymbol{g}_{2,t}^{\mathsf{T}}\boldsymbol{u}_{2,t}^{'}a_{2,t}^{'}}_{P^{\alpha_1-\alpha_2+\Delta}}$$
$$+\underbrace{\boldsymbol{g}_{2,t}^{\mathsf{T}}\boldsymbol{v}_{2,t}b_{2,t}}_{P^{\alpha_1+\Delta}}+\underbrace{\boldsymbol{g}_{2,t}^{\mathsf{T}}\boldsymbol{v}_{2,t}^{'}b_{2,t}^{'}}_{P^{\Delta}}+\underbrace{z_{2,t}^{(2)}}_{P^0}, \tag{17}$$

for $t=1,2,\cdots,T_2$, where under each term we noted the order of the summand's average power.

At this point, based on (16),(17), each user decodes $c_{2,t}$ by treating the other signals as noise. After decoding $\{c_{2,t}\}_{t=1}^{T_2}$ and fully reconstructing $\{\hat{c}_{1,t}^{(a)},\hat{c}_{1,t}^{(b)},\}_{t=1}^{T_1}$, user 1 goes back one phase and subtracts $\hat{c}_{1,t}^{(b)}$ from $y_{1,t}^{(1)}$ to remove (up to bounded noise) the interference corresponding to $\bar{c}_{1,t}^{(b)}$. The same user will also use the estimate $\hat{c}_{1,t}^{(a)}$ of $\bar{c}_{1,t}^{(a)}$ as an extra observation which, together with the observation $y_{1,t}^{(1)}$, present the user with a 2×2 MIMO channel that allows for decoding of both $a_{1,t}$ and $a_{1,t}^{'}$. Similarly user 2, after fully reconstructing $\{\hat{c}_{1,t}^{(a)},\hat{c}_{1,t}^{(b)},\}_{t=1}^{T_1}$, subtracts $\hat{c}_{1,t}^{(a)}$ from $y_{1,t}^{(2)}$, to remove (up to bounded noise) the interference corresponding to $\bar{c}_{1,t}^{(a)}$, and also uses the estimate $\hat{c}_{1,t}^{(b)}$ of $\bar{c}_{1,t}^{(b)}$ as an extra observation which, together with the observation $y_{1,t}^{(2)}$, allow for decoding of both $b_{1,t}$ and $b_{1,t}^{'}$. Further exposition to the details regarding the achievability of the mentioned rates, can be found in Appendix V.

Consequently after the end of the second phase, the transmitter can use its knowledge of delayed CSIT to reconstruct $\{\bar{c}_{2,t}^{(a)},\bar{c}_{2,t}^{(b)}\}_{t=1}^{T_2}$, and quantize each term to $\hat{c}_{2,t}^{(a)},\hat{c}_{2,t}^{(b)}$. With $\mathbb{E}|\bar{c}_{2,t}^{(a)}|^2\doteq P^{\alpha_1-\alpha_2+\Delta}$, $\mathbb{E}|\bar{c}_{2,t}^{(b)}|^2\doteq P^{\Delta}$, we choose a quantization rate that assigns each $\hat{c}_{2,t}^{(a)}$ a total of $(\alpha_1-\alpha_2+\Delta)\log P+o(\log P)$ bits, and each $\hat{c}_{2,t}^{(b)}$ a total of $\Delta\log P+o(\log P)$ bits, thus allowing for $\mathbb{E}|\tilde{c}_{2,t}^{(a)}|^2\doteq\mathbb{E}|\tilde{c}_{2,t}^{(b)}|^2\doteq1$. Then the $T_2(\alpha_1-\alpha_2+2\Delta)\log P+o(\log P)$ bits representing $\{\hat{c}_{2,t}^{(a)},\hat{c}_{2,t}^{(b)}\}_{t=1}^{T_2}$, are split evenly across the set $\{c_{3,t}\}_{t=1}^{T_3}$ which will be sequentially transmitted in the next phase so that user 1 can eventually decode $\{a_{2,t},a_{2,t}^{'}\}_{t=1}^{T_2}$, and user 2 can decode $\{b_{2,t},b_{2,t}^{'}\}_{t=1}^{T_2}$.

We now proceed with the general description of phase s.

3) Phase s, $3\le s\le S-1$: Phase s ($T_s=T_{s-1}\frac{\alpha_1-\alpha_2+2\Delta}{1-\alpha_1-\Delta}$ channel uses) is almost identical to phase 2, with one difference being the different relationship between T_s and T_{s-1}. The transmit signal takes the same form as in phase 2 (cf. (8),(14)), the rates and powers of the symbols are the same (cf. (15)) and the received signals $y_{s,t}^{(1)},y_{s,t}^{(2)}$ ($t=1,\cdots,T_s$) take the same form as in (16),(17).

Most of the actions are also the same, where based on (16),(17) (corresponding now to phase s), each user decodes $c_{s,t}$ by treating the other signals as noise, and then goes back one phase and reconstructs $\{\hat{c}_{s-1,t}^{(a)},\hat{c}_{s-1,t}^{(b)},\}_{t=1}^{T_{s-1}}$. As before, user 1 then subtracts $\hat{c}_{s-1,t}^{(b)}$ from $y_{s-1,t}^{(1)}$ to remove, up to bounded noise, the interference corresponding to $\bar{c}_{s-1,t}^{(b)}$. The same user also employs the estimate $\hat{c}_{s-1,t}^{(a)}$ of $\bar{c}_{s-1,t}^{(a)}$ as an extra observation which, together with the observation $y_{s-1,t}^{(1)}-\boldsymbol{h}_{s-1,t}^{\mathsf{T}}\boldsymbol{w}_{s-1,t}c_{s-1,t}-\hat{c}_{s-1,t}^{(b)}$ obtained after decoding $c_{s-1,t}$, allow for decoding of both $a_{s-1,t}$ and $a_{s-1,t}^{'}$. Similar actions are performed by user 2.

As before, after the end of phase s, the transmitter can use its knowledge of delayed CSIT to reconstruct $\{\bar{c}_{s,t}^{(a)},\bar{c}_{s,t}^{(b)}\}_{t=1}^{T_s}$, and quantize each term to $\hat{c}_{s,t}^{(a)},\hat{c}_{s,t}^{(b)}$ with the same rate as in phase 2 (($\alpha_1-\alpha_2+\Delta)\log P+o(\log P)$ bits for each $\hat{c}_{s,t}^{(a)}$, and $\Delta\log P+o(\log P)$ bits for each $\hat{c}_{s,t}^{(b)}$). Finally the accumulated $T_s(\alpha_1-\alpha_2+2\Delta)\log P+o(\log P)$ bits representing all the quantized values $\{\hat{c}_{s,t}^{(a)},\hat{c}_{s,t}^{(b)}\}_{t=1}^{T_s}$, are distributed evenly across the set $\{c_{s+1,t}\}_{t=1}^{T_{s+1}}$ which will be sequentially transmitted in the next phase. More details can be found in Appendix V.

4) Phase S: During the last phase ($T_S=T_{S-1}\frac{\alpha_1-\alpha_2+2\Delta}{1-\alpha_2}$ channel uses), the transmit signal is

$$\boldsymbol{x}_{S,t}=\boldsymbol{w}_{S,t}c_{S,t}+\boldsymbol{u}_{S,t}a_{S,t}+\boldsymbol{v}_{S,t}b_{S,t} \tag{18}$$

where we set power and rate as

$$\begin{aligned}P_S^{(c)}&\doteq P, & r_S^{(c)}&=1-\alpha_2\\ P_S^{(a)}&\doteq P^{\alpha_2}, & r_S^{(a)}&=\alpha_2\\ P_S^{(b)}&\doteq P^{\alpha_2}, & r_S^{(b)}&=\alpha_2.\end{aligned} \tag{19}$$

The received signals are

$$y_{S,t}^{(1)}=\underbrace{\boldsymbol{h}_{S,t}^{\mathsf{T}}\boldsymbol{w}_{S,t}c_{S,t}}_{P}+\underbrace{\boldsymbol{h}_{S,t}^{\mathsf{T}}\boldsymbol{u}_{S,t}a_{S,t}}_{P^{\alpha_2}}+\underbrace{\tilde{\boldsymbol{h}}_{S,t}^{\mathsf{T}}\boldsymbol{v}_{S,t}b_{S,t}}_{P^{\alpha_2-\alpha_1}}+\underbrace{z_{S,t}^{(1)}}_{P^0},$$
$$y_{S,t}^{(2)}=\underbrace{\boldsymbol{g}_{S,t}^{\mathsf{T}}\boldsymbol{w}_{S,t}c_{S,t}}_{P}+\underbrace{\tilde{\boldsymbol{g}}_{S,t}^{\mathsf{T}}\boldsymbol{u}_{S,t}a_{S,t}}_{P^0}+\underbrace{\boldsymbol{g}_{S,t}^{\mathsf{T}}\boldsymbol{v}_{S,t}b_{S,t}}_{P^{\alpha_2}}+\underbrace{z_{S,t}^{(2)}}_{P^0}, \tag{20}$$

for $t=1,2,\cdots,T_S$.

At this point, as before, the power and rate allocation of the different symbols allow both users to decode $c_{S,t}$ by treating the other signals as noise. Consequently user 1 can remove $\boldsymbol{h}_{S,t}^{\mathsf{T}}\boldsymbol{w}_{S,t}c_{S,t}$ from $y_{S,t}^{(1)}$ and decode $a_{S,t}$, and similarly user 2 can remove $\boldsymbol{g}_{S,t}^{\mathsf{T}}\boldsymbol{w}_{S,t}c_{S,t}$ from $y_{S,t}^{(2)}$ and decode $b_{S,t}$. Finally each user goes back one phase and reconstructs $\{\hat{c}_{S-1,t}^{(a)},\hat{c}_{S-1,t}^{(b)}\}_{t=1}^{T_{S-1}}$, which allows for decoding of $a_{S-1,t}$ and $a_{S-1,t}^{'}$ at user 1 and of $b_{S-1,t}$ and $b_{S-1,t}^{'}$ at user 2, all as described for the previous phases (see Appendix V for more details).

Table I summarizes the parameters of scheme $\mathcal{X}_1$. The use of symbol $\perp$ is meant to indicate precoding that is orthogonal to the channel estimate (rather than random). The table's last row indicates the prelog factor of the quantization rate.

TABLE I
SUMMARY OF SCHEME $\mathcal{X}_1$.

	Phase 1	Phase 2	Ph. s ($3 \leq s \leq S-1$)	Phase S
Duration	T_1	$T_1\xi$	$T_1\xi\mu^{s-2}$	$T_1\xi\mu^{S-3}\gamma$
$r^{(a)}$	1	$\alpha_1+\Delta$	$\alpha_1+\Delta$	α_2
$r^{(a')}$	$1-\alpha_2$	$\alpha_1-\alpha_2+\Delta$	$\alpha_1-\alpha_2+\Delta$	-
$r^{(b)}$	1	$\alpha_1+\Delta$	$\alpha_1+\Delta$	α_2
$r^{(b')}$	$1-\alpha_1$	Δ	Δ	-
$r^{(c)}$	-	$1-\alpha_1-\Delta$	$1-\alpha_1-\Delta$	$1-\alpha_2$
$P^{(a)}\perp$	P	$P^{\alpha_1+\Delta}$	$P^{\alpha_1+\Delta}$	P^{α_2}
$P^{(a')}$	$P^{1-\alpha_2}$	$P^{\alpha_1-\alpha_2+\Delta}$	$P^{\alpha_1-\alpha_2+\Delta}$	-
$P^{(b)}\perp$	P	$P^{\alpha_1+\Delta}$	$P^{\alpha_1+\Delta}$	P^{α_2}
$P^{(b')}$	$P^{1-\alpha_1}$	P^{Δ}	P^{Δ}	-
$P^{(c)}$	-	P	P	P
Quant.	$2-\alpha_1-\alpha_2$	$\alpha_1-\alpha_2+2\Delta$	$\alpha_1-\alpha_2+2\Delta$	0

a) DoF calculation for scheme $\mathcal{X}_1$: We proceed to add up the total amount of information transmitted during this scheme.

In accordance with the declared pre-log factors $r_s^{(a)}, r_s^{(a')}$ and phase durations (see Table I), we have that

$$d_1 = (T_1(2-\alpha_2) + \sum_{i=2}^{S-1} T_i(2\alpha_1-\alpha_2+2\Delta) + T_S\alpha_2)/(\sum_{i=1}^{S} T_i)$$

$$= (\sum_{i=2}^{S-1}(T_i(1-\alpha_1-\Delta)+T_i(\alpha_1+\Delta)) + T_S(1-\alpha_2)$$

$$+ T_S\alpha_2 + T_1\alpha_1 - \Delta\sum_{i=2}^{S}T_i)/(\sum_{i=1}^{S} T_i) \qquad (21)$$

$$= (1-\Delta) + \frac{T_1(\alpha_1+\Delta-1)+T_S\Delta}{\sum_{i=1}^{S}T_i}, \qquad (22)$$

where (21) considers the phase durations seen in (10). Considering that $0 < \mu < 1$ (see (10) for case 1), that $\sum_{i=0}^{S-3}\mu^i = \frac{1-\mu^{S-2}}{1-\mu}$, and given an asymptotically high S, we see that

$$d_1 = (1-\Delta) + \frac{\frac{T_2}{\xi}(\alpha_1+\Delta-1)+T_2\mu^{S-3}\gamma\Delta}{\frac{T_2}{\xi}+T_2(\frac{1}{1-\mu}+\mu^{S-3}(\gamma-\frac{\mu}{1-\mu}))} \qquad (23)$$

$$= (1-\Delta) + \frac{\frac{1}{\xi}(\alpha_1+\Delta-1)}{\frac{1}{\xi}+\frac{1}{1-\mu}}$$

$$= (1-\Delta) - \frac{1+\alpha_2-2\alpha_1-3\Delta}{3} = \frac{2+2\alpha_1-\alpha_2}{3}. \qquad (24)$$

Similarly, considering the values for $r_s^{(b)}, r_s^{(b')}$, we have that

$$d_2 = \frac{T_1(2-\alpha_1)+\sum_{i=2}^{S-1}T_i(\alpha_1+2\Delta)+T_S\alpha_2}{\sum_{i=1}^{S}T_i}$$

$$= \alpha_1+2\Delta + \frac{T_1(2-2\alpha_1-2\Delta)+T_S(\alpha_2-\alpha_1-2\Delta)}{\sum_{i=1}^{S}T_i}$$

$$= \alpha_1+2\Delta + \frac{\frac{T_2}{\xi}(2-2\alpha_1-2\Delta)+T_2\mu^{S-3}\gamma(\alpha_2-\alpha_1-2\Delta)}{\frac{T_2}{\xi}+T_2(\frac{1}{1-\mu}+\mu^{S-3}(\gamma-\frac{\mu}{1-\mu}))}$$

which, in the high S limit, gives

$$d_2 = \alpha_1+2\Delta + \frac{\frac{1}{\xi}(2-2\alpha_1-2\Delta)}{\frac{1}{\xi}+\frac{1}{1-\mu}}$$

$$= \alpha_1+2\Delta + \frac{2(1+\alpha_2-2\alpha_1-3\Delta)}{3} = \frac{2+2\alpha_2-\alpha_1}{3}. \qquad (25)$$

In conclusion, scheme $\mathcal{X}_1$ achieves DoF pair $C = (\frac{2+2\alpha_1-\alpha_2}{3}, \frac{2+2\alpha_2-\alpha_1}{3})$ (case 1).

B. Scheme $\mathcal{X}_2$ achieving $D = (1, \alpha_1)$ (case 1), and $A = (1, \frac{1+\alpha_2}{2})$ (case 2)

Scheme $\mathcal{X}_2$ is designed with S phases, with phase durations $T_1, T_2, \cdots, T_S$ chosen to be integers such that

$$T_2 = T_1\tau, \quad T_s = T_{s-1}\beta = T_1\tau\beta^{s-2}, \forall s \in \{3,4,\cdots,S-1\},$$
$$T_S = T_{S-1}\eta = T_1\tau\beta^{S-3}\eta, \qquad (26)$$

where $\tau = \frac{1-\alpha_2}{1-\alpha_1}$, $\beta = \frac{\alpha_1-\alpha_2}{1-\alpha_1}$, $\eta = \frac{\alpha_1-\alpha_2}{1-\alpha_2}$.

The scheme is similar to $\mathcal{X}_1$, but with a different power and rate allocation, and a different input structure since now user 2 only receives a single private information symbol.

1) Phase 1: During phase 1 (T_1 channel uses), the transmitter sends

$$\boldsymbol{x}_{1,t} = \boldsymbol{u}_{1,t}a_{1,t} + \boldsymbol{u}'_{1,t}a'_{1,t} + \boldsymbol{v}_{1,t}b_{1,t},$$

with power and rate set as

$$P_1^{(a)} \doteq P, \quad P_1^{(a')} \doteq P^{1-\alpha_2}, \quad P_1^{(b)} \doteq P^{\alpha_1}$$
$$r_1^{(a)} = 1, \quad r_1^{(a')} = 1-\alpha_2, \quad r_1^{(b)} = \alpha_1.$$

The received signals take the form

$$y_{1,t}^{(1)} = \underbrace{\boldsymbol{h}_{1,t}^\mathsf{T}\boldsymbol{u}_{1,t}a_{1,t}}_{P} + \underbrace{\boldsymbol{h}_{1,t}^\mathsf{T}\boldsymbol{u}'_{1,t}a'_{1,t}}_{P^{1-\alpha_2}} + \underbrace{\tilde{\boldsymbol{h}}_{1,t}^\mathsf{T}\boldsymbol{v}_{1,t}b_{1,t}}_{P^0} + \underbrace{z_{1,t}^{(1)}}_{P^0},$$

$$y_{1,t}^{(2)} = \underbrace{\overbrace{\tilde{\boldsymbol{g}}_{1,t}^\mathsf{T}\boldsymbol{u}_{1,t}a_{1,t}}^{\bar{c}_{1,t}^{(a)}}}_{P^{1-\alpha_2}} + \underbrace{\boldsymbol{g}_{1,t}^\mathsf{T}\boldsymbol{u}'_{1,t}a'_{1,t}}_{P^{1-\alpha_2}} + \underbrace{\boldsymbol{g}_{1,t}^\mathsf{T}\boldsymbol{v}_{1,t}b_{1,t}}_{P^{\alpha_1}} + \underbrace{z_{1,t}^{(2)}}_{P^0}.$$

After the end of the first phase, the transmitter reconstructs $\{\bar{c}_{1,t}^{(a)}\}_{t=1}^{T_1}$ (cf.(7)), and quantizes each term as

$$\bar{c}_{1,t}^{(a)} = \hat{c}_{1,t}^{(a)} + \tilde{c}_{1,t}^{(a)}, \quad t = 1,2,\cdots,T_1.$$

Noting that $\mathbb{E}|\bar{c}_{1,t}^{(a)}|^2 \doteq P^{1-\alpha_2}$, we choose a quantization rate that assigns each $\hat{c}_{1,t}^{(a)}$ a total of $(1-\alpha_2)\log P + o(\log P)$ bits, thus allowing for $\mathbb{E}|\tilde{c}_{1,t}^{(a)}|^2 \doteq 1$. Then the $T_1(1-\alpha_2)\log P + o(\log P)$ bits representing $\{\hat{c}_{1,t}^{(a)}\}_{t=1}^{T_1}$ are distributed evenly across the set $\{c_{2,t}\}_{t=1}^{T_2}$ which will be transmitted in the next phase. As before, transmission of $\{c_{2,t}\}_{t=1}^{T_2}$ aims to help user 2 cancel out interference, as well as aims to provide user 1 with an extra observation which will allow for decoding of the user's private information.

2) Phase 2: During phase 2 (T_2 channel uses), the transmitter sends

$$\boldsymbol{x}_{2,t} = \boldsymbol{w}_{2,t}c_{2,t} + \boldsymbol{u}_{2,t}a_{2,t} + \boldsymbol{u}'_{2,t}a'_{2,t} + \boldsymbol{v}_{2,t}b_{2,t}$$

with power and rate set as

$$
\begin{aligned}
P_2^{(c)} &\doteq P, & r_2^{(c)} &= 1 - \alpha_1 \\
P_2^{(a)} &\doteq P^{\alpha_1}, & r_2^{(a)} &= \alpha_1 \\
P_2^{(a')} &\doteq P^{\alpha_1-\alpha_2}, & r_2^{(a')} &= \alpha_1 - \alpha_2 \\
P_2^{(b)} &\doteq P^{\alpha_1}, & r_2^{(b)} &= \alpha_1,
\end{aligned}
\tag{28}
$$

where we note that $r_2^{(c)}$ satisfies $T_2 r_2^{(c)} = T_1(1 - \alpha_2)$.

The received signals in this phase are

$$y_{2,t}^{(1)} = \underbrace{\boldsymbol{h}_{2,t}^{\mathsf{T}}\boldsymbol{w}_{2,t}c_{2,t}}_{P} + \underbrace{\boldsymbol{h}_{2,t}^{\mathsf{T}}\boldsymbol{u}_{2,t}a_{2,t}}_{P^{\alpha_1}} + \underbrace{\boldsymbol{h}_{2,t}^{\mathsf{T}}\boldsymbol{u}'_{2,t}a'_{2,t}}_{P^{\alpha_1-\alpha_2}} + \underbrace{\tilde{\boldsymbol{h}}_{2,t}^{\mathsf{T}}\boldsymbol{v}_{2,t}b_{2,t}}_{P^0} + \underbrace{z_{2,t}^{(1)}}_{P^0}$$

$$\tag{29}$$

$$y_{2,t}^{(2)} = \underbrace{\boldsymbol{g}_{2,t}^{\mathsf{T}}\boldsymbol{w}_{2,t}c_{2,t}}_{P} + \underbrace{\tilde{\boldsymbol{g}}_{2,t}^{\mathsf{T}}\boldsymbol{u}_{2,t}a_{2,t}}_{P^{\alpha_1-\alpha_2}} + \underbrace{\boldsymbol{g}_{2,t}^{\mathsf{T}}\boldsymbol{u}'_{2,t}a'_{2,t}}_{P^{\alpha_1-\alpha_2}} + \underbrace{\boldsymbol{g}_{2,t}^{\mathsf{T}}\boldsymbol{v}_{2,t}b_{2,t}}_{P^{\alpha_1}} + \underbrace{z_{2,t}^{(2)}}_{P^0}$$

$$\tag{30}$$

for $t = 1, 2, \cdots, T_2$.

Then, based on (29),(30), each user decodes $c_{2,t}$ by treating the other signals as noise, and then proceeds to reconstruct $\{\hat{c}_{1,t}^{(a)}\}_{t=1}^{T_1}$. User 1 combines each $\hat{c}_{1,t}^{(a)}$ with its corresponding observation $y_{1,t}^{(1)}$, to introduce T_2 independent 2×2 MIMO channels that allow for decoding of all $a_{1,t}$ and $a'_{1,t}$. At the same time, user 2 subtracts $\hat{c}_{1,t}^{(a)}$ from $y_{1,t}^{(2)}$ to remove (up to bounded noise) the interference corresponding to $\bar{c}_{1,t}^{(a)}$, which in turn allows for decoding of $b_{1,t}$.

Consequently after the end of the second phase, the transmitter can use its knowledge of delayed CSIT to reconstruct $\{\bar{c}_{2,t}^{(a)}\}_{t=1}^{T_2}$, and quantize each term to $\hat{c}_{2,t}^{(a)}$. With $\mathbb{E}|\bar{c}_{2,t}^{(a)}|^2 \doteq P^{\alpha_1-\alpha_2}$, we choose a quantization rate that assigns each $\hat{c}_{2,t}^{(a)}$ a total of $(\alpha_1 - \alpha_2)\log P + o(\log P)$ bits, a choice that allows for $\mathbb{E}|\tilde{c}_{2,t}^{(a)}|^2 \doteq 1$. Then the $T_2(\alpha_1 - \alpha_2)\log P + o(\log P)$ bits representing $\{\hat{c}_{2,t}^{(a)}\}_{t=1}^{T_2}$, are distributed evenly across the set $\{c_{3,t}\}_{t=1}^{T_3}$ which will be transmitted in the next phase.

3) Phase s, $3 \leq s \leq S-1$: Phase s ($T_s = T_{s-1}\frac{\alpha_1-\alpha_2}{1-\alpha_1}$ channel uses) is almost identical to phase 2, except for the relationship between T_s and T_{s-1}. Specifically the transmit signal takes the same form as in phase 2

$$\boldsymbol{x}_{s,t} = \boldsymbol{w}_{s,t}\underbrace{c_{s,t}}_{P_s^{(c)}} + \boldsymbol{u}_{s,t}\underbrace{a_{s,t}}_{P_s^{(a)}} + \boldsymbol{u}'_{s,t}\underbrace{a'_{s,t}}_{P_s^{(a')}} + \boldsymbol{v}_{s,t}\underbrace{b_{s,t}}_{P_s^{(b)}},$$

the rates and powers of the symbols are the same (cf. (28)), and the received signals $y_{s,t}^{(1)}, y_{s,t}^{(2)}$ ($t = 1, \cdots, T_s$) take the same form as in (29),(30).

The actions are also the same, where based on (29),(30) (corresponding now to phase s), each user decodes $c_{s,t}$ by treating the other signals as noise, and then goes back one phase and reconstructs $\{\hat{c}_{s-1,t}^{(a)}\}_{t=1}^{T_{s-1}}$. As before, user 1 then employs the estimate $\hat{c}_{s-1,t}^{(a)}$ of $\bar{c}_{s-1,t}^{(a)}$ as an extra observation which, together with the observation $y_{s-1,t}^{(1)} - \boldsymbol{h}_{s-1,t}^{\mathsf{T}}\boldsymbol{w}_{s-1,t}c_{s-1,t}$

attained after decoding $c_{s-1,t}$, allow for decoding of both $a_{s-1,t}$ and $a'_{s-1,t}$. At the same time, user 2 subtracts $\hat{c}_{s-1,t}^{(a)}$ from $y_{s-1,t}^{(2)}$ to remove (up to bounded noise) the interference corresponding to $\bar{c}_{s-1,t}^{(a)}$, which allows for decoding of $b_{s-1,t}$.

Again as before, after the end of phase s, the transmitter can use delayed CSIT to reconstruct $\{\bar{c}_{s,t}^{(a)}\}_{t=1}^{T_s}$, and quantize each term to $\hat{c}_{s,t}^{(a)}$ with the same rate as in phase 2 (($\alpha_1 - \alpha_2)\log P + o(\log P)$ bits per channel use). Finally the total of the $T_s(\alpha_1 - \alpha_2)\log P + o(\log P)$ bits representing the quantized values $\{\hat{c}_{s,t}^{(a)}\}_{t=1}^{T_s}$ is split evenly to the set $\{c_{s+1,t}\}_{t=1}^{T_{s+1}}$ which will be transmitted in the next phase.

4) Phase S: During the last phase ($T_S = T_{S-1}\frac{\alpha_1-\alpha_2}{1-\alpha_2}$ channel uses), the transmitter sends

$$\boldsymbol{x}_{S,t} = \boldsymbol{w}_{S,t}c_{S,t} + \boldsymbol{u}_{S,t}a_{S,t} + \boldsymbol{v}_{S,t}b_{S,t} \tag{31}$$

with power and rates set as

$$
\begin{aligned}
P_S^{(c)} &\doteq P, & r_S^{(c)} &= 1 - \alpha_2 \\
P_S^{(a)} &\doteq P^{\alpha_2}, & r_S^{(a)} &= \alpha_2 \\
P_S^{(b)} &\doteq P^{\alpha_2}, & r_S^{(b)} &= \alpha_2.
\end{aligned}
\tag{32}
$$

resulting in received signals of the form

$$y_{S,t}^{(1)} = \underbrace{\boldsymbol{h}_{S,t}^{\mathsf{T}}\boldsymbol{w}_{S,t}c_{S,t}}_{P} + \underbrace{\boldsymbol{h}_{S,t}^{\mathsf{T}}\boldsymbol{u}_{S,t}a_{S,t}}_{P^{\alpha_2}} + \underbrace{\tilde{\boldsymbol{h}}_{S,t}^{\mathsf{T}}\boldsymbol{v}_{S,t}b_{S,t}}_{P^{\alpha_2-\alpha_1}} + \underbrace{z_{S,t}^{(1)}}_{P^0},$$

$$y_{S,t}^{(2)} = \underbrace{\boldsymbol{g}_{S,t}^{\mathsf{T}}\boldsymbol{w}_{S,t}c_{S,t}}_{P} + \underbrace{\tilde{\boldsymbol{g}}_{S,t}^{\mathsf{T}}\boldsymbol{u}_{S,t}a_{S,t}}_{P^0} + \underbrace{\boldsymbol{g}_{S,t}^{\mathsf{T}}\boldsymbol{v}_{S,t}b_{S,t}}_{P^{\alpha_2}} + \underbrace{z_{S,t}^{(2)}}_{P^0},$$

($t = 1, \cdots, T_S$).

As before, both receivers decode $c_{S,t}$ by treating all other signals as noise. Consequently user 1 removes $\boldsymbol{h}_{S,t}^{\mathsf{T}}\boldsymbol{w}_{S,t}c_{S,t}$ from $y_{S,t}^{(1)}$ and decodes $a_{S,t}$, and user 2 removes $\boldsymbol{g}_{S,t}^{\mathsf{T}}\boldsymbol{w}_{S,t}c_{S,t}$ from $y_{S,t}^{(2)}$ and decodes $b_{S,t}$. Finally each user goes back one phase and reconstructs $\{\hat{c}_{S-1,t}^{(a)}\}_{t=1}^{T_{S-1}}$, which in turn allows for decoding of $a_{S-1,t}$ and $a'_{S-1,t}$ at user 1 and of $b_{S-1,t}$ at user 2, all as described in the previous phases. The DoF achievability details follow those of scheme $\mathcal{X}_1$ (Appendix V).

Table II summarizes the parameters of scheme $\mathcal{X}_2$. The last row indicates the prelog factor of the quantization rate.

TABLE II
SUMMARY OF SCHEME $\mathcal{X}_2$.

	Phase 1	Phase 2	Ph.s ($3 \leq s \leq S-1$)	Phase S
Duration	T_1	$T_1\tau$	$T_1\tau\beta^{s-2}$	$T_1\tau\beta^{S-3}\eta$
$r^{(a)}$	1	α_1	α_1	α_2
$r^{(a')}$	$1-\alpha_2$	$\alpha_1-\alpha_2$	$\alpha_1-\alpha_2$	-
$r^{(b)}$	α_1	α_1	α_1	α_2
$r^{(c)}$	-	$1-\alpha_1$	$1-\alpha_1$	$1-\alpha_2$
$P^{(a)}\perp$	P	P^{α_1}	P^{α_1}	P^{α_2}
$P^{(a')}$	$P^{1-\alpha_2}$	$P^{\alpha_1-\alpha_2}$	$P^{\alpha_1-\alpha_2}$	-
$P^{(b)}\perp$	P^{α_1}	P^{α_1}	P^{α_1}	P^{α_2}
$P^{(c)}$	-	P	P	P
Quant.	$1-\alpha_2$	$\alpha_1-\alpha_2$	$\alpha_1-\alpha_2$	0

a) DoF calculation for scheme $\mathfrak{X}_2$: We proceed to add up the total amount of information transmitted during this scheme.

In accordance to the declared pre-log factors $r_s^{(a)}, r_s^{(a')}$ and phase durations (see Table II), and irrespective of whether α_1, α_2 fall under case 1 or case 2, we have that

$$
d_1 = (T_1(2-\alpha_2) + \sum_{i=2}^{S-1} T_i(2\alpha_1 - \alpha_2) + T_S\alpha_2)/(\sum_{i=1}^{S} T_i)
$$

$$
= (T_1 + T_1(1-\alpha_2) + \sum_{i=2}^{S-1}(T_i\alpha_1 + T_i(\alpha_1 - \alpha_2)) + T_S\alpha_2)/(\sum_{i=1}^{S} T_i)
$$

$$
= (T_1 + \sum_{i=2}^{S-1}(T_i(1-\alpha_1) + T_i\alpha_1) + T_S(1-\alpha_2) + T_S\alpha_2)/(\sum_{i=1}^{S} T_i) \tag{34}
$$

$$
= \frac{T_1 + T_2 + T_3 + \cdots + T_{S-1} + T_S}{T_1 + T_2 + \cdots + T_S} = 1 \tag{35}
$$

where (34) is due to (26).

Regarding the second user and the declared $r_s^{(b)}$, for case 1 ($2\alpha_1 - \alpha_2 < 1$) we see that

$$
d_2 = \frac{\sum_{i=1}^{S-1} T_i\alpha_1 + T_S\alpha_2}{\sum_{i=1}^{S} T_i} = \alpha_1 - \frac{T_S(\alpha_1 - \alpha_2)}{\sum_{i=1}^{S} T_i}
$$

$$
= \alpha_1 - \frac{T_1 \tau \beta^{S-3}\eta(\alpha_1 - \alpha_2)}{T_1 + T_1\tau\sum_{i=0}^{S-3}\beta^i + T_1\tau\beta^{S-3}\eta} \tag{36}
$$

$$
= \alpha_1 - \frac{\beta^{S-3}\eta(\alpha_1 - \alpha_2)}{\frac{1}{\tau} + \sum_{i=0}^{S-3}\beta^i + \beta^{S-3}\eta} \tag{37}
$$

$$
= \alpha_1 - \frac{\beta^{S-3}\eta(\alpha_1 - \alpha_2)}{\frac{1}{\tau} + \frac{1-\beta^{S-2}}{1-\beta} + \beta^{S-3}\eta}
$$

$$
= \alpha_1 - \frac{\beta^{S-3}\eta(\alpha_1 - \alpha_2)}{\frac{1}{\tau} + \frac{1}{1-\beta} + \beta^{S-3}(\eta - \frac{\beta}{1-\beta})} = \alpha_1, \tag{38}
$$

where we have used (26) to get (36), where we have used that $2\alpha_1 - \alpha_2 < 1$ implies $\beta < 1$, and where we have considered an asymptotically large S.

When $2\alpha_1 - \alpha_2 > 1$ ($\beta > 1$), then (37) gives that

$$
d_2 = \alpha_1 - \frac{\beta^{S-3}\eta(\alpha_1 - \alpha_2)}{\frac{1}{\tau} + \frac{1}{1-\beta} + \beta^{S-3}(\eta - \frac{\beta}{1-\beta})}
$$

$$
= \alpha_1 - \frac{\eta(\alpha_1 - \alpha_2)}{\frac{1-\beta+\tau}{\beta^{S-3}\tau(1-\beta)} + (\eta - \frac{\beta}{1-\beta})}
$$

which, in the high S regime, gives

$$
d_2 = \alpha_1 - \frac{\eta(\alpha_1 - \alpha_2)}{\eta - \frac{\beta}{1-\beta}} = \alpha_1 + \frac{1-2\alpha_1+\alpha_2}{2} = \frac{1+\alpha_2}{2}. \tag{39}
$$

When $2\alpha_1 - \alpha_2 = 1$ ($\beta = 1$), then (37) gives that $d_2 = \alpha_1 - \frac{\eta(\alpha_1 - \alpha_2)}{\frac{1}{\tau} + S - 2 + \eta}$ which, for large S, gives

$$
d_2 = \alpha_1 = \frac{1+\alpha_2}{2}. \tag{40}
$$

In conclusion, scheme $\mathfrak{X}_2$ achieves DoF pair $D = (1, \alpha_1)$ (case 1), else it achieves $A = (1, \frac{1+\alpha_2}{2})$.

C. Scheme $\mathfrak{X}_3$ achieving $B = (\alpha_2, 1)$

This is the simplest of all three schemes, and it consists of a single channel use[2] ($S = 1, T_1 = 1$) during which the transmitter sends

$$
\boldsymbol{x} = \boldsymbol{w}c + \boldsymbol{u}a + \boldsymbol{v}b,
$$

where $\boldsymbol{u}$ is orthogonal to $\hat{\boldsymbol{g}}$, $\boldsymbol{v}$ is orthogonal to $\hat{\boldsymbol{h}}$, and where the power and rates are set as

$$
\begin{aligned}
P^{(c)} &\doteq P, & r^{(c)} &= 1 - \alpha_1 \\
P^{(a)} &\doteq P^{\alpha_2}, & r^{(a)} &= \alpha_2 \\
P^{(b)} &\doteq P^{\alpha_1}, & r^{(b)} &= \alpha_1,
\end{aligned} \tag{41}
$$

resulting in received signals of the form

$$
y^{(1)} = \boldsymbol{h}^{\mathsf{T}}\boldsymbol{x} + z^{(1)} = \underbrace{\boldsymbol{h}^{\mathsf{T}}\boldsymbol{w}c}_{P} + \underbrace{\boldsymbol{h}^{\mathsf{T}}\boldsymbol{u}a}_{P^{\alpha_2}} + \underbrace{\tilde{\boldsymbol{h}}^{\mathsf{T}}\boldsymbol{v}b}_{P^0} + \underbrace{z^{(1)}}_{P^0},
$$

$$
y^{(2)} = \boldsymbol{g}^{\mathsf{T}}\boldsymbol{x} + z^{(2)} = \underbrace{\boldsymbol{g}^{\mathsf{T}}\boldsymbol{w}c}_{P} + \underbrace{\tilde{\boldsymbol{g}}^{\mathsf{T}}\boldsymbol{u}a}_{P^0} + \underbrace{\boldsymbol{g}^{\mathsf{T}}\boldsymbol{v}b}_{P^{\alpha_1}} + \underbrace{z^{(2)}}_{P^0}.
$$

After transmission, both receivers first decode c by treating the other signals as noise, and then user 1 utilizes its knowledge of $\{\boldsymbol{h}, \boldsymbol{g}, \hat{\boldsymbol{h}}, \hat{\boldsymbol{g}}\}$ to reconstruct $\boldsymbol{h}^{\mathsf{T}}\boldsymbol{w}c$ and remove it from $y^{(1)}$, thus being able to decode a, while after decoding c, user 2 removes $\boldsymbol{g}^{\mathsf{T}}\boldsymbol{w}c$ from $y^{(2)}$, and decodes b. The details for the achievability of $r^{(a)}, r^{(b)}, r^{(c)}$ follow closely the exposition in Appendix V. Consequently the DoF point ($d_1 = \alpha_2, \quad d_2 = 1$) can be achieved by associating c to information intended entirely for the second user.

IV. CONCLUSIONS

The work provided analysis and communication schemes for the setting of the two-user MISO BC with general mixed CSIT. The work can be seen as a natural extension of the result in [10] and of the recent results in [1], [7]–[9], to the case where the CSIT feedback quality varies across different links.

V. APPENDIX - DETAILS OF ACHIEVABILITY PROOF

We will here focus on achievability details for scheme $\mathfrak{X}_1$. The clarifications of the details carry over easily to the other two schemes.

Regarding $r_s^{(c)}$ ($2 \leq s \leq S - 1$ - see (15)), we recall that during phase s, both users decode $c_{s,t}$ (from $y_{s,t}^{(1)}, y_{s,t}^{(2)}, t = 1, \cdots, T_s$ - see (29),(30)) by treating all other signals as noise. Consequently for $\boldsymbol{H} \triangleq \{\boldsymbol{h}_{i,j}, \boldsymbol{g}_{i,j}, \hat{\boldsymbol{h}}_{i,j}, \hat{\boldsymbol{g}}_{i,j}, \forall i, j\}$, we note that

$$
\begin{aligned}
I(c_{s,t}; y_{s,t}^{(1)}, \boldsymbol{H}) &= I(c_{s,t}; y_{s,t}^{(2)}, \boldsymbol{H}) \\
&= (1 - \alpha_1 - \Delta) \log P + o(\log P),
\end{aligned}
$$

---

[2]We will henceforth maintain the same notation as before, but for simplicity we will remove the phase and time index.

to get

$$r_s^{(c)} = \frac{1}{\log P} \min\{I(c_{s,t}; y_{s,t}^{(1)}, \boldsymbol{H}), I(c_{s,t}; y_{s,t}^{(2)}, \boldsymbol{H})\}$$
$$= 1 - \alpha_1 - \Delta.$$

Similarly for the last phase S (see (18),(19),(20)), we note that

$$I(c_{S,t}; y_{S,t}^{(1)}, \boldsymbol{H}) = I(c_{S,t}; y_{S,t}^{(2)}, \boldsymbol{H}) = (1 - \alpha_2) \log P + o(\log P),$$

to get

$$r_S^{(c)} = \frac{1}{\log P} \min\{I(c_{S,t}; y_{S,t}^{(1)}, \boldsymbol{H}), I(c_{S,t}; y_{S,t}^{(2)}, \boldsymbol{H})\} = 1 - \alpha_2.$$

Regarding achievability for $r_1^{(a)} = 1$, $r_1^{(a')} = 1 - \alpha_2$, $r_1^{(b)} = 1$ and $r_1^{(b')} = 1 - \alpha_1$ (see (11),(12),(13)), we note that each element in $\{c_{2,t}\}_{t=1}^{T_2}$ has enough bits (recall that $r_2^{(c)} = 1 - \alpha_1 - \Delta$), to match the quantization rate of $\{\hat{c}_{1,t}^{(a)}, \hat{c}_{1,t}^{(b)}\}_{t=1}^{T_1}$ that is necessary in order to have a bounded quantization noise. Consequently going back to phase 1, user 1 is presented with T_1 linearly independent 2×2 equivalent MIMO channels of the form

$$\begin{bmatrix} y_{1,t}^{(1)} - \hat{c}_{1,t}^{(b)} \\ \hat{c}_{1,t}^{(a)} \end{bmatrix} = \begin{bmatrix} \boldsymbol{h}_{1,t}^{\mathsf{T}} \\ \boldsymbol{g}_{1,t}^{\mathsf{T}} \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_{1,t} & \boldsymbol{u}_{1,t}' \end{bmatrix} \begin{bmatrix} a_{1,t} \\ a_{1,t}' \end{bmatrix} + \begin{bmatrix} z_{1,t}^{(1)} + \tilde{c}_{1,t}^{(b)} \\ -\tilde{c}_{1,t}^{(a)} \end{bmatrix}$$

$(t = 1, 2, \cdots, T_1)$, where again we note that the described quantization rate results in a bounded equivalent noise, which then immediately gives that $r_1^{(a)} = 1$ and $r_1^{(a')} = 1 - \alpha_2$ are achievable. Similarly for user 2, the presented T_1 linearly independent 2×2 equivalent MIMO channels

$$\begin{bmatrix} \hat{c}_{1,t}^{(b)} \\ y_{1,t}^{(2)} - \hat{c}_{1,t}^{(a)} \end{bmatrix} = \begin{bmatrix} \boldsymbol{h}_{1,t}^{\mathsf{T}} \\ \boldsymbol{g}_{1,t}^{\mathsf{T}} \end{bmatrix} \begin{bmatrix} \boldsymbol{v}_{1,t} & \boldsymbol{v}_{1,t}' \end{bmatrix} \begin{bmatrix} b_{1,t} \\ b_{1,t}' \end{bmatrix} + \begin{bmatrix} -\tilde{c}_{1,t}^{(b)} \\ z_{1,t}^{(2)} + \tilde{c}_{1,t}^{(a)} \end{bmatrix}$$

$(t = 1, 2, \cdots, T_1)$, allow for decoding at a rate corresponding to $r_1^{(b)} = 1$ and $r_1^{(b')} = 1 - \alpha_1$.

Regarding achievability for $r_s^{(a)} = \alpha_1 + \Delta$, $r_s^{(a')} = \alpha_1 - \alpha_2 + \Delta$, $r_s^{(b)} = \alpha_1 + \Delta$ and $r_s^{(b')} = \Delta$, $(2 \le s \le S - 1$ - see (8),(14), (15)), we note that during phase s, both users can decode $c_{s,t}$, and as a result user 1 can remove $\boldsymbol{h}_{s,t}^{\mathsf{T}} \boldsymbol{w}_{s,t} c_{s,t}$ from $y_{s,t}^{(1)}$, and user 2 can remove $\boldsymbol{g}_{s,t}^{\mathsf{T}} \boldsymbol{w}_{s,t} c_{s,t}$ from $y_{s,t}^{(2)}$ ($t = 1, \cdots, T_s$). As a result user 1 is presented with T_s linearly independent 2×2 equivalent MIMO channels of the form

$$\begin{bmatrix} y_{s,t}^{(1)} - \boldsymbol{h}_{s,t}^{\mathsf{T}} \boldsymbol{w}_{s,t} c_{s,t} - \hat{c}_{s,t}^{(b)} \\ \hat{c}_{s,t}^{(a)} \end{bmatrix} = \begin{bmatrix} \boldsymbol{h}_{s,t}^{\mathsf{T}} \\ \boldsymbol{g}_{s,t}^{\mathsf{T}} \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_{s,t} & \boldsymbol{u}_{s,t}' \end{bmatrix} \begin{bmatrix} a_{s,t} \\ a_{s,t}' \end{bmatrix} + \begin{bmatrix} z_{s,t}^{(1)} + \tilde{c}_{s,t}^{(b)} \\ -\tilde{c}_{s,t}^{(a)} \end{bmatrix}$$

$(t = 1, \cdots, T_s)$. Given that the rate associated to $\{c_{s+1,t}\}_{t=1}^{T_{s+1}}$, matches the quantization rate for $\{\hat{c}_{s,t}^{(a)}, \hat{c}_{s,t}^{(b)}\}_{t=1}^{T_s}$, allows for a bounded variance of the equivalent noise, and in turn for decoding of $\{a_{s,t}, a_{s,t}'\}_{t=1}^{T_s}$ at a rate corresponding to $r_s^{(a)} = \alpha_1 + \Delta$ and $r_s^{(a')} = \alpha_1 - \alpha_2 + \Delta$. Similarly user 2 is presented with T_s independent 2×2 MIMO channels of the form

$$\begin{bmatrix} \hat{c}_{s,t}^{(b)} \\ y_{s,t}^{(2)} - \boldsymbol{g}_{s,t}^{\mathsf{T}} \boldsymbol{w}_{s,t} c_{s,t} - \hat{c}_{s,t}^{(a)} \end{bmatrix} = \begin{bmatrix} \boldsymbol{h}_{s,t}^{\mathsf{T}} \\ \boldsymbol{g}_{s,t}^{\mathsf{T}} \end{bmatrix} \begin{bmatrix} \boldsymbol{v}_{s,t} & \boldsymbol{v}_{s,t}' \end{bmatrix} \begin{bmatrix} b_{s,t} \\ b_{s,t}' \end{bmatrix} + \begin{bmatrix} -\tilde{c}_{s,t}^{(b)} \\ z_{s,t}^{(2)} + \tilde{c}_{s,t}^{(a)} \end{bmatrix}$$

allowing for decoding of $\{b_{s,t}, b_{s,t}'\}_{t=1}^{T_s}$ ($t = 1, \cdots, T_s$) at rates corresponding to $r_s^{(b)} = \alpha_1 + \Delta$ and $r_s^{(b')} = \Delta$.

Regarding achievability for $r_S^{(a)} = \alpha_2$ and $r_S^{(b)} = \alpha_2$ (see (18),(19),(20)), we note that, after decoding $c_{S,t}$, user 1 can remove $\boldsymbol{h}_{S,t}^{\mathsf{T}} \boldsymbol{w}_{S,t} c_{S,t}$ from $y_{S,t}^{(1)}$, and user 2 can remove $\boldsymbol{g}_{S,t}^{\mathsf{T}} \boldsymbol{w}_{S,t} c_{S,t}$ from $y_{S,t}^{(2)}$, ($t = 1, \cdots, T_S$). Consequently during this phase, user 1 sees T_S linearly independent SISO channels of the form

$$\tilde{y}_{S,t}^{(1)} \triangleq y_{S,t}^{(1)} - \boldsymbol{h}_{S,t}^{\mathsf{T}} \boldsymbol{w}_{S,t} c_{S,t} = \boldsymbol{h}_{S,t}^{\mathsf{T}} \boldsymbol{u}_{S,t} a_{S,t} + \tilde{\boldsymbol{h}}_{S,t}^{\mathsf{T}} \boldsymbol{v}_{S,t} b_{S,t} + z_{S,t}^{(1)}$$

$(t = 1, \cdots, T_S)$ which can be readily shown to support $r_S^{(a)} = \alpha_2$. A similar argument gives achievability for $r_S^{(b)} = \alpha_2$. $\square$

VI. APPENDIX - PROOF OF OUTER BOUND

We here adopt the outer bound approach in [9] to the asymmetric case of $\alpha_1 \ne \alpha_2$. As in [9], we first linearly convert the original BC in (1a),(1b) to an equivalent BC (see (43a),(43b)) having the same DoF region as the original BC (cf. [9]), and we then consider the degraded version of the equivalent BC in the absence of delayed feedback, which matches in capacity the degraded BC with feedback (for the memoryless case), and which exceeds the capacity of the equivalent BC. The final step considers the compound and degraded version of the equivalent BC without delayed feedback, whose DoF region will serve as an outer bound on the DoF region of the original BC.

b) The equivalent degraded compound BC: Towards the equivalent BC, directly from (1a),(1b) we have that

$$\begin{aligned} y_t^{(1)} &= \boldsymbol{h}_t^{\mathsf{T}} \boldsymbol{x}_t + z_t^{(1)} \\ &= \boldsymbol{h}_t^{\mathsf{T}} \sqrt{P} \boldsymbol{Q}_t \frac{1}{\sqrt{P}} \boldsymbol{Q}_t^{-1} \boldsymbol{x}_t + z_t^{(1)} \\ &= \boldsymbol{h}_t^{\mathsf{T}} \sqrt{P} \boldsymbol{Q}_t \boldsymbol{x}_t' + z_t^{(1)} \\ &= \sqrt{P} \boldsymbol{h}_t^{\mathsf{T}} \boldsymbol{u}_t x_t^1 + \sqrt{P} \tilde{\boldsymbol{h}}_t^{\mathsf{T}} \boldsymbol{v}_t x_t^2 + z_t^{(1)} \quad \text{(43a)} \\ y_t^{(2)} &= \boldsymbol{g}_t^{\mathsf{T}} \boldsymbol{x}_t + z_t^{(2)} \\ &= \boldsymbol{g}_t^{\mathsf{T}} \sqrt{P} \boldsymbol{Q}_t \boldsymbol{x}_t' + z_t^{(2)} \\ &= \sqrt{P} \tilde{\boldsymbol{g}}_t^{\mathsf{T}} \boldsymbol{u}_t x_t^1 + \sqrt{P} \boldsymbol{g}_t^{\mathsf{T}} \boldsymbol{v}_t x_t^2 + z_t^{(2)}, \quad \text{(43b)} \end{aligned}$$

where

$$\boldsymbol{x}_t' \triangleq [x_t^1 \ x_t^2]^T \triangleq \frac{1}{\sqrt{P}} \boldsymbol{Q}_t^{-1} \boldsymbol{x}_t,$$

where $\boldsymbol{Q}_t \triangleq [\boldsymbol{u}_t \ \boldsymbol{v}_t] \in \mathbb{C}^{2 \times 2}$ is, with probability 1, an invertible matrix, where $\boldsymbol{u}_t$ is chosen to be of unit norm and orthogonal to $\hat{\boldsymbol{g}}_t$, and where $\boldsymbol{v}_t$ is chosen to be of unit norm and orthogonal

to $\hat{\boldsymbol{h}}_t$. Furthermore each receiver normalizes to get

$$
\begin{aligned}
y_t^{'(1)} &= \frac{y_t^{(1)}}{\boldsymbol{h}_t^{\mathsf{T}}\boldsymbol{u}_t} \\
&= \sqrt{P}x_t^1 + \frac{\sqrt{P}\tilde{\boldsymbol{h}}_t^{\mathsf{T}}\boldsymbol{v}_t x_t^2}{\boldsymbol{h}_t^{\mathsf{T}}\boldsymbol{u}_t} + \frac{z_t^{(1)}}{\boldsymbol{h}_t^{\mathsf{T}}\boldsymbol{u}_t} \\
&= \sqrt{P}x_t^1 + \sqrt{P^{1-\alpha_1}}h_t' x_t^2 + z_t^{'(1)}, \quad (44\text{a})
\end{aligned}
$$

$$
\begin{aligned}
y_t^{'(2)} &= \frac{y_t^{(2)}}{\boldsymbol{g}_t^{\mathsf{T}}\boldsymbol{v}_t} \\
&= \sqrt{P}x_t^2 + \frac{\sqrt{P}\tilde{\boldsymbol{g}}_t^{\mathsf{T}}\boldsymbol{u}_t x_t^1}{\boldsymbol{g}_t^{\mathsf{T}}\boldsymbol{v}_t} + \frac{z_t^{(2)}}{\boldsymbol{g}_t^{\mathsf{T}}\boldsymbol{v}_t} \\
&= \sqrt{P}x_t^2 + \sqrt{P^{1-\alpha_2}}g_t' x_t^1 + z_t^{'(2)}, \quad (44\text{b})
\end{aligned}
$$

where $z_t^{'(1)} = \frac{z_t^{(1)}}{\boldsymbol{h}_t^{\mathsf{T}}\boldsymbol{u}_t}$, $h_t' = \frac{\sqrt{P^{\alpha_1}}\tilde{\boldsymbol{h}}_t^{\mathsf{T}}\boldsymbol{v}_t}{\boldsymbol{h}_t^{\mathsf{T}}\boldsymbol{u}_t}$, $z_t^{'(2)} = \frac{z_t^{(2)}}{\boldsymbol{g}_t^{\mathsf{T}}\boldsymbol{v}_t}$, $g_t' = \frac{\sqrt{P^{\alpha_2}}\tilde{\boldsymbol{g}}_t^{\mathsf{T}}\boldsymbol{u}_t}{\boldsymbol{g}_t^{\mathsf{T}}\boldsymbol{v}_t}$. Consequently $\sqrt{P^{\alpha_1}}\tilde{\boldsymbol{h}}_t$ and $\sqrt{P^{\alpha_2}}\tilde{\boldsymbol{g}}_t$ have identity covariance matrices, and the average power of h_t', g_t', $z_t^{'(1)}$ and $z_t^{'(2)}$ does not scale with P, i.e., in the high-SNR region this power is of order P^0. With the same CSIT knowledge mapped from the original BC, it can be shown (see [9]) that the DoF region of the equivalent BC in (44a)(44b) matches the DoF region of the original BC in (1a)(1b).

Towards designing the degraded version of the above equivalent BC, we supply the second user with knowledge of $y_t^{'(1)}$, and towards designing the compound version of the above degraded equivalent BC, we add two extra users (user 3 and 4). In this compound version, the received signals for the first two users are as in (44a)(44b), while the received signals of the added (virtual) users are given by

$$
y_t^{''(1)} = \sqrt{P}x_t^1 + \sqrt{P^{1-\alpha_1}}h_t'' x_t^2 + z_t^{''(1)}, \quad (45\text{a})
$$
$$
y_t^{''(2)} = \sqrt{P}x_t^2 + \sqrt{P^{1-\alpha_2}}g_t'' x_t^1 + z_t^{''(2)}. \quad (45\text{b})
$$

We here note that by definition, h_t'' and g_t'' are statistically equivalent to the original h_t' and g_t' respectively, and that $z_t^{''(1)}$ and $z_t^{''(2)}$ are statistically equivalent to the original $z_t^{'(1)}$ and $z_t^{'(2)}$. Furthermore we note that user 3 is interested in the same message as user 1, while user 4 is interested in the same message as user 2. Also we recall that in the specific degraded compound BC, user 1 knows $y_t^{'(1)}$, user 2 knows $y_t^{'(2)}$ and $y_t^{'(1)}$, user 3 knows $y_t^{''(1)}$, and user 4 knows $y_t^{''(2)}$ and $y_t^{''(1)}$. Finally we remove delayed feedback - a removal known to not affect the capacity of the degraded BC without memory [12].

We now proceed to calculate an outer bound on the DoF region of this degraded compound BC which at least matches the DoF of the previous degraded BC and which serves as an outer bound on the DoF region of the original BC.

c) Outer bound: We consider communication over the described equivalent degraded compound BC, letting n be the large number of fading realizations over which communication takes place, and letting R_1, R_2 be the rates of the first and second user. We also let $\boldsymbol{H}_{[n]} \triangleq \{\boldsymbol{h}_t, \boldsymbol{g}_t, \hat{\boldsymbol{h}}_t, \hat{\boldsymbol{g}}_t\}_{t=1}^n$, $y_{[n]}^{'(i)} \triangleq \{y_t^{'(i)}\}_{t=1}^n$ and $y_{[n]}^{''(i)} \triangleq \{y_t^{''(i)}\}_{t=1}^n$ for $i = 1, 2$.

Using Fano's inequality, we have

$$
\begin{aligned}
nR_1 &\le I(W_1; y_{[n]}^{'(1)} | \boldsymbol{H}_{[n]}) + no(n) \\
&\le n\log P + no(\log P) - h(y_{[n]}^{'(1)} | W_1, \boldsymbol{H}_{[n]}) + no(n), \quad (46)
\end{aligned}
$$

as well as

$$
\begin{aligned}
nR_1 &\le I(W_1; y_{[n]}^{''(1)} | \boldsymbol{H}_{[n]}) + no(n) \\
&\le n\log P + no(\log P) - h(y_{[n]}^{''(1)} | W_1, \boldsymbol{H}_{[n]}) + no(n), \quad (47)
\end{aligned}
$$

which is added to (46) to give

$$
\begin{aligned}
2nR_1 &\le 2n\log P + 2no(\log P) - h(y_{[n]}^{'(1)} | W_1, \boldsymbol{H}_{[n]}) \\
&\quad - h(y_{[n]}^{''(1)} | W_1, \boldsymbol{H}_{[n]}) + 2no(n) \\
&\le 2n\log P + 2no(\log P) \\
&\quad - h(y_{[n]}^{'(1)}, y_{[n]}^{''(1)} | W_1, \boldsymbol{H}_{[n]}) + 2no(n). \quad (48)
\end{aligned}
$$

Let

$$
\begin{aligned}
\bar{\boldsymbol{y}}_1 &\triangleq \operatorname{diag}(1, \sqrt{P^{\alpha_1}}) \begin{bmatrix} 1 & h_t' \\ 1 & h_t'' \end{bmatrix}^{-1} \begin{bmatrix} y_t^{'(1)} \\ y_t^{''(1)} \end{bmatrix} \\
&= \begin{bmatrix} \sqrt{P}x_t^1 \\ \sqrt{P}x_t^2 \end{bmatrix} + \begin{bmatrix} \frac{z_t^{'(1)}h_t'' - z_t^{''(1)}h_t'}{h_t'' - h_t'} \\ \sqrt{P^{\alpha_1}}\frac{z_t^{''(1)} - z_t^{'(1)}}{h_t'' - h_t'} \end{bmatrix} \\
&= \begin{bmatrix} \sqrt{P}x_t^1 \\ \sqrt{P}x_t^2 \end{bmatrix} + \begin{bmatrix} \bar{z}_t \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ z_t \end{bmatrix} \quad (49)
\end{aligned}
$$

where $\bar{z}_t = \frac{z_t^{'(1)}h_t'' - z_t^{''(1)}h_t'}{h_t'' - h_t'}$, $z_t = \sqrt{P^{\alpha_1}}\frac{z_t^{''(1)} - z_t^{'(1)}}{h_t'' - h_t'}$, and let $z_{[n]} \triangleq \{z_t\}_{t=1}^n$. Consequently

$$
\begin{aligned}
nR_1 + nR_2 &= h(W_1, W_2) \\
&= I(W_1, W_2; y_{[n]}^{'(1)}, y_{[n]}^{''(1)}, z_{[n]} | \boldsymbol{H}_{[n]}) \\
&\quad + h(W_1, W_2 | y_{[n]}^{'(1)}, y_{[n]}^{''(1)}, z_{[n]}, \boldsymbol{H}_{[n]}) \\
&= I(W_1, W_2; y_{[n]}^{'(1)}, y_{[n]}^{''(1)}, z_{[n]} | \boldsymbol{H}_{[n]}) \\
&\quad + no(\log P) + no(n) \quad (50) \\
&= I(W_1; y_{[n]}^{'(1)}, y_{[n]}^{''(1)}, z_{[n]} | \boldsymbol{H}_{[n]}) \\
&\quad + I(W_2; y_{[n]}^{'(1)}, y_{[n]}^{''(1)}, z_{[n]} | \boldsymbol{H}_{[n]}, W_1) \\
&\quad + no(\log P) + no(n), \quad (51)
\end{aligned}
$$

where the transition to (50) uses the fact that the high SNR variance of $\bar{z}_t$ and z_t scales as P^0 and P^{α_1} respectively, which in turn means that knowledge of $\{y_t^{'(1)}, y_t^{''(1)}, z_t, \boldsymbol{H}_{[n]}\}_{t=1}^n$, implies knowledge of W_1, W_2 and of $\{x_t^1, x_t^2\}_{t=1}^n$, up to bounded noise level.

Furthermore

$$
\begin{aligned}
nR_1 &= h(W_1) \\
&= I(W_1; y_{[n]}^{'(1)}, y_{[n]}^{''(1)}, z_{[n]} | \boldsymbol{H}_{[n]}) + h(W_1 | y_{[n]}^{'(1)}, y_{[n]}^{''(1)}, z_{[n]}, \boldsymbol{H}_{[n]}) \\
&= I(W_1; y_{[n]}^{'(1)}, y_{[n]}^{''(1)}, z_{[n]} | \boldsymbol{H}_{[n]}) + no(\log P) + no(n), \quad (52)
\end{aligned}
$$

since again knowledge of $\{y_t^{'(1)}, y_t^{''(1)}, z_t, \boldsymbol{H}_{[n]}\}_{t=1}^n$ provides for W_1 up to bounded noise level.

Now combining (51) and (52), gives

$$nR_2 = I(W_2; y_{[n]}^{'(1)}, y_{[n]}^{''(1)}, z_{[n]} | \boldsymbol{H}_{[n]}, W_1) + no(\log P) + no(n)$$

$$= I(W_2; y_{[n]}^{'(1)}, y_{[n]}^{''(1)} | \boldsymbol{H}_{[n]}, W_1)$$
$$\quad + I(W_2; z_{[n]} | y_{[n]}^{'(1)}, y_{[n]}^{''(1)}, \boldsymbol{H}_{[n]}, W_1) + no(\log P) + no(n)$$

$$= h(y_{[n]}^{'(1)}, y_{[n]}^{''(1)} | \boldsymbol{H}_{[n]}, W_1) - \underbrace{h(y_{[n]}^{'(1)}, y_{[n]}^{''(1)} | \boldsymbol{H}_{[n]}, W_1, W_2)}_{no(\log P)}$$

$$\quad - \underbrace{h(z_{[n]} | y_{[n]}^{'(1)}, y_{[n]}^{''(1)}, \boldsymbol{H}_{[n]}, W_1, W_2)}_{no(\log P)}$$

$$\quad + \underbrace{h(z_{[n]} | y_{[n]}^{'(1)}, y_{[n]}^{''(1)}, \boldsymbol{H}_{[n]}, W_1)}_{\leq h(z_{[n]})} + no(\log P) + no(n)$$

$$\leq h(y_{[n]}^{'(1)}, y_{[n]}^{''(1)} | \boldsymbol{H}_{[n]}, W_1) + h(z_{[n]}) + no(\log P) + no(n)$$

$$\leq h(y_{[n]}^{'(1)}, y_{[n]}^{''(1)} | W_1, \boldsymbol{H}_{[n]}) + n\alpha_1 \log P$$
$$\quad + no(\log P) + no(n),$$

which is combined with (48) to give

$$2nR_1 + nR_2 \leq 2n \log P + n\alpha_1 \log P + no(\log P) + no(n), \tag{53}$$

which in turn proves the outer bound

$$2d_1 + d_2 \leq 2 + \alpha_1, \tag{54}$$

as described in (6b). Finally interchanging the roles of the two users and of α_1, α_2, gives

$$d_1 + 2d_2 \leq 2 + \alpha_2. \tag{55}$$

Naturally the single antenna constraint gives that $d_1 \leq 1, d_2 \leq 1$. $\qquad\square$

REFERENCES

[1] 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.

[2] A. Ghasemi, A. S. Motahari, and A. K. Khandani, "On the degrees of freedom of X channel with delayed CSIT," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Jul. 2011.

[3] C. S. Vaze and M. K. Varanasi, "The degrees of freedom region of two-user and certain three-user MIMO broadcast channel with delayed CSI," Dec. 2011, submitted to *IEEE Trans. Inf. Theory*, available on arXiv:1101.0306.

[4] M. J. Abdoli, A. Ghasemi, and A. K. Khandani, "On the degrees of freedom of three-user MIMO broadcast channel with delayed CSIT," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Jul. 2011.

[5] A. Ghasemi, A. S. Motahari, and A. K. Khandani, "Interference alignment for the MIMO interference channel with delayed local CSIT," Feb. 2011, available on arXiv:1102.5673v1.

[6] J. Xu, J. G. Andrews, and S. A. Jafar, "Broadcast channels with delayed finite-rate feedback: Predict or observe?" *IEEE Trans. Wireless Commun.*, vol. 11, no. 4, pp. 1456 – 1467, Apr. 2012.

[7] M. Kobayashi, S. Yang, D. Gesbert, and X. Yi, "On the degrees of freedom of time correlated MISO broadcast channel with delayed CSIT," in *Proc. IEEE Int. Symp. Information Theory (ISIT)*, Jul. 2012.

[8] S. Yang, M. Kobayashi, D. Gesbert, and X. Yi, "Degrees of freedom of time correlated MISO broadcast channel with delayed CSIT," *IEEE Trans. Inf. Theory*, vol. 59, no. 1, pp. 315–328, Jan. 2013.

[9] T. Gou and S. Jafar, "Optimal use of current and outdated channel state information: Degrees of freedom of the MISO BC with mixed CSIT," *IEEE Communications Letters*, vol. 16, no. 7, pp. 1084 – 1087, Jul. 2012.

[10] H. Maleki, S. Jafar, and S. Shamai, "Retrospective interference alignment over interference networks," *IEEE Journal of Selected Topics in Signal Processing*, vol. 6, no. 3, pp. 228 – 240, Mar. 2012.

[11] T. Cover and J. Thomas, *Elements of Information Theory*, 2nd ed. New York: Wiley-Interscience, 2006.

[12] A. E. Gamal, "The feedback capacity of degraded broadcast channels," *IEEE Trans. Inf. Theory*, vol. 24, no. 3, pp. 379–381, Apr. 1978.