A non-linear transmission scheme for the K-user Broadcast Channel with state feedback

Chao He, Sheng Yang, Pablo Piantanida

A Non-Linear Transmission Scheme for the K-User Broadcast Channel with State Feedback

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

***Abstract*—The state-dependent K-User Broadcast Channel (BC) with memoryless state feedback is investigated. We propose a novel transmission scheme and derive its corresponding achievable rate region which has the advantage of being simple and thus easy to evaluate. In particular, it is first shown that this scheme achieves the capacity of the Symmetric Erasure BC with an arbitrary size of input alphabet. Then, we further study the Fading Gaussian BC and derive a symmetric achievable rate. Besides achieving the optimal degrees of freedom, numerical results show that the proposed scheme attains higher rates than the previous schemes at finite SNR.**

I. Introduction

It is well known that feedback can enlarge the capacity region in communication networks, including the Broadcast Channel (BC) [1], [2]. The role of feedback in such channels can be understood as an enabler for cooperation between users. However, as many problems in network information theory, the capacity of the general BC with feedback is still open, even for two-user case. In fact, for the K-user BC case, less is known. Nevertheless, due to its practical importance in wireless networks, a large amount of recent works have been focused on such channels where considerable progresses have been made, e.g., linear schemes using state feedback in Erasure BC (EBC) [3], [4] and in Fading Gaussian BC (GBC) [5].

The capacity region of the EBC was fully determined in the three-user case and partially characterized for $K > 3$ cases, across independent works [3], [4]. The main idea behind the schemes proposed in [3], [4] is fundamentally the same: first send out the uncoded packets, and then transmit adequate "linear combinations" of the lost but overheard packets according to the state feedback. Given sufficient linearly independent combinations at each receiver, the desired packets are always decodable. However, the schemes in [3], [4] only works for a packet alphabet of size 2^q with $q \in \mathbb{N}$ and $2^q \geq K$. This is to guarantee the existence of a desired number of linearly independent vectors in the corresponding vector space. As a consequence, the capacity region is still open for other alphabet sizes.

In a parallel line of works on the GBC, the optimal Degrees of Freedom (DoF) region was shown to be achieved by the proposed MAT scheme in [5] for the K-user MISO (multiple-input-single-output) case when the number of transmit antennas is larger than K. The MAT scheme works similarly as the schemes in [3], [4]: first send out the signal uncoded and then the linear combinations of overheard signal, so that enough equations are available to each user. Since the MAT scheme in [5] has a fixed structure built on a dimension counting argument, the DoF optimality at high SNR may not ensure its efficiency in finite SNR due to its insensitivity to such parameters. In an independent work [6], Shayevitz and Wigger studied the two-user broadcast channel with generalized feedback and proposed an achievable rate region based on double binning and block Markov coding. Later on, Kim *et al.* showed that, in the two-user setting, the Shayevitz-Wigger scheme actually includes the MAT scheme as a special case in [7]. Finite SNR performance of the MAT scheme has also been investigated in [8] for the two-user case and in [9] for the K-user case.

In this work, we focus on the K-user ($K > 2$) case and investigate a class of state-dependent BCs with state feedback which includes both the EBC and the Fading GBC as special cases. The main contributions are the following. First, we propose a non-linear scheme based on distributed compression and derive the corresponding rate region that is simple to evaluate. It is worth mentioning that our goal here is not to derive a general rate region that includes all the known regions (e.g., the two-user Shayevitz-Wigger region). Instead, we are interested in schemes with relatively simple structure in order to have a numerically tractable rate region. To that end, we make some reasonable choices such as excluding binning at the transmitter (since the instantaneous state is unknown, we know that binning has limited benefit in GBC.). Then, we show that the proposed scheme achieves the symmetric capacity of a Symmetric EBC for *any* input alphabet size. Further, for the Symmetric Fading GBC, we derive an achievable symmetric rate as a maximization over K positive values. Analytical results show that the proposed scheme achieves the optimal DoF under the same setting as in [5], whereas a numerical example of the three-user symmetric rate reveals a superior performance over existing schemes in the literature.

The rest of the paper is organized as follows. We describe the system model in Section II and provide the main results in Section III. The novel scheme is presented in Section IV. An example on the Fading Gaussian BC is given in Section V. Finally, Section VI concludes the paper. Some of the technical details are relegated to the appendix.

Notation: First, for random quantities, we use upper case letters, e.g., X, for scalars, upper case letters with bold and non-italic fonts, e.g., $\mathbf{V}$, for vectors, and upper case letter with

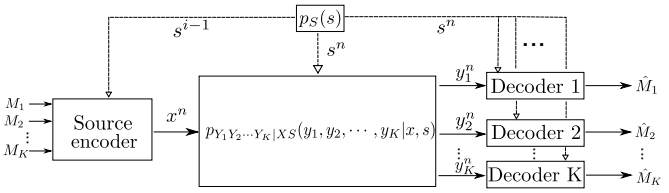

Fig. 1. Broadcast channel with strictly causal state feedback.

bold and sans serif fonts, e.g., M, for matrices. Deterministic quantities are denoted with italic letters, e.g., a scalar x, a vector $\boldsymbol{v}$, and a matrix $\boldsymbol{M}$. Logarithms are in base 2. Calligraphic letters are used for sets. In particular, we use $\mathcal{I}$ and $\mathcal{J}$ for some subsets of users with implicit size constraints $|\mathcal{I}| = i$ and $|\mathcal{J}| = j$, respectively. The constraints are made explicit when necessary. $\mathcal{U}$ is also used as subset of users but without size constraint. We let $\mathcal{K} \triangleq \{1, \ldots, K\}$ be the set of all users. Hence, $\{V_\mathcal{I}\}_\mathcal{I} \equiv \{V_\mathcal{I} : \mathcal{I} \subseteq \mathcal{K}, |\mathcal{I}| = i\}$, $\{V_\mathcal{U}\}_\mathcal{U} \equiv \{V_\mathcal{U} : \mathcal{U} \subseteq \mathcal{K}\}$, and $\{V_\mathcal{I}\}_{\mathcal{I}\subset\mathcal{J}} \equiv \{V_\mathcal{I} : \mathcal{I} \subset \mathcal{J}, |\mathcal{I}| = i\}$.

II. System Model

We consider a memoryless state-dependent K-user BC through which the source communicates, in n slots, K independent messages to the K receivers, respectively. The channel can be described by the joint probability mass function (pmf),

$$p(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_K|\boldsymbol{x}, \boldsymbol{s})p(\boldsymbol{s}) = \prod_{i=1}^{n} p(y_{1i}, \ldots, y_{Ki}|x_i, s_i)p(s_i) \quad (1)$$

where $\boldsymbol{x} \in \mathcal{X}^n$, $\boldsymbol{y}_k \in \mathcal{Y}_k^n$, and $\boldsymbol{s} \in \mathcal{S}^n$ are the sequences of the channel input, the k-th channel output, and the channel state, respectively. The channel state is known instantaneously to the receivers but only strictly causally to the source, e.g., via feedback. The channel model is shown in Fig. 1.

Let $M_k \in \mathcal{M}_k \triangleq [1 : 2^{nR_k}]$ be the message for user k, $k \in \mathcal{K}$. We say that the rate tuple $(R_1, \ldots, R_K)$ is achievable if there exist a sequence of encoding functions $\{f_t : \mathcal{M}_1 \times \cdots \times \mathcal{M}_K \times \mathcal{S}^{t-1} \to \mathcal{X}\}_{t=1}^n$ and K decoding functions $\{g_k : \mathcal{Y}_k^n \times \mathcal{S}^n \to \mathcal{M}_k\}_{k=1}^K$, such that $\max_k \Pr\big(g_k(Y_k^n, S^n) \neq M_k\big) \to 0$ when $n \to \infty$. The symmetric rate R_{sym} is achievable if the rate tuple $(R_{\text{sym}}, \ldots, R_{\text{sym}})$ is achievable. In particular, we are interested in the following two specific channels.

A. Erasure Broadcast Channel

The Erasure Broadcast Channel is a state-dependent deterministic channel which is defined at each slot by

$$Y_k = \begin{cases} X, & S_k = 1 \\ e, & S_k = 0 \end{cases} \quad (2)$$

which is completely characterized by the set of probabilities of the state realization

$$\delta_\mathcal{U} \triangleq \Pr(S_\mathcal{U} = \mathbf{0},\ S_{\bar{\mathcal{U}}} = \mathbf{1}), \quad \mathcal{U} \subseteq \{1, \ldots, K\} \quad (3)$$

with $\sum_\mathcal{U} \delta_\mathcal{U} = 1$. Throughout the paper, we use $S_\mathcal{U} = \mathbf{0}$ (*resp.* $S_\mathcal{U} = \mathbf{1}$) to mean that $S_k = 0$ (*resp.* $S_k = 1$), $\forall\ k \in \mathcal{U}$. When the erasures across users are independent and identically distributed (i.i.d.), we have:

$$\delta_\mathcal{U} = \delta^{|\mathcal{U}|}(1-\delta)^{K-|\mathcal{U}|}, \quad \mathcal{U} \subseteq \{1, \ldots, K\} \quad (4)$$

where $\delta \in [0, 1]$ is the erasure probability for each user. We focus on the i.i.d. case in this paper.

B. Fading Gaussian Broadcast Channel

The Fading Gaussian BC is a state-dependent AWGN channel, which is defined at each slot by

$$\mathbf{Y}_k = \mathsf{H}_k\mathbf{X} + \mathbf{Z}_k, \quad k = 1, \ldots, K \quad (5)$$

where $\mathbf{Z}_k \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I}_{n_{\mathrm{r},k}})$, $\mathbf{X} \in \mathbb{C}^{n_\mathrm{t}\times 1}$, $\mathbf{Y}_k \in \mathbb{C}^{n_{\mathrm{r},k}\times 1}$, and $\mathsf{H}_k \in \mathbb{C}^{n_{\mathrm{r},k}\times n_\mathrm{t}}$, with n_t, $n_{\mathrm{r},k}$ being antenna number at transmitter and receiver k, respectively. We identify the state S_k with the channel matrix H_k. Similarly as for $S_\mathcal{U}$, we use $\mathsf{H}_\mathcal{U}$ to denote a matrix from a vertical concatenation of all the matrices $\{\mathsf{H}_k\}_{k\in\mathcal{U}}$, same notation applies for $\mathbf{Y}_\mathcal{U}$ and $\mathbf{Z}_\mathcal{U}$. Hence, it follows that $\mathbf{Y}_\mathcal{U} = \mathsf{H}_\mathcal{U}\mathbf{X} + \mathbf{Z}_\mathcal{U}$. The channel is subject to the input power constraint $\frac{1}{n}\sum_{i=1}^n \|\boldsymbol{x}_i\|^2 \leq P$ for any input sequence $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$. The Signal-to-Noise Ratio (SNR) is defined as $\mathsf{snr} \triangleq \frac{P}{n_\mathrm{t}\sigma^2}$. Unless indicated otherwise, we assume that the channel matrices are independent across users.

III. Main Results

We first introduce the following random variables (RVs).

- $X^{(j)}$, $\{Y_k^{(j)}\}_k$, $S^{(j)}$, and $Q^{(j)}$ are the input, output, state, and time sharing RVs, respectively, for phases $j = 1, \ldots, K$, with pmf

$$\prod_{j=1}^{K} p(y_1^{(j)}, \ldots, y_K^{(j)} \,|\, x^{(j)}, s^{(j)})p(s^{(j)})p(q^{(j)}). \quad (6)$$

- $(\hat{Y}_{i\to\mathcal{J}}, V_{i\to\mathcal{J}})$, $i < j$ and $|\mathcal{J}| = j$, are the side information intended for users in $\mathcal{J}$ from phase i, and the signal that carries such information, respectively, with pmf

$$\prod_{j=1}^{K}\prod_{i=0}^{j-1}\prod_{\mathcal{J}} p(v_{i\to\mathcal{J}})p(\hat{y}_{i\to\mathcal{J}} \,|\, \{v_\mathcal{I}\}_{\mathcal{I}\subset\mathcal{J}}, s^{(i)}, q^{(i)}), \quad (7)$$

where we define $v_\mathcal{U} \triangleq \{v_{k\to\mathcal{U}} :\ k < |\mathcal{U}|\}$ and $V_\mathcal{U} \triangleq \{V_{k\to\mathcal{U}} :\ k < |\mathcal{U}|\}$ for brevity. And the input at phase j, $j = 1, \ldots, K$, is generated with the pmf

$$\prod_{j=1}^{K} p(x^{(j)} \,|\, \{v_\mathcal{J}\}_\mathcal{J}, q^{(j)}). \quad (8)$$

The main result of this paper is stated below.

Theorem 1. *For some K-tuple $(\alpha_1, \ldots, \alpha_K) \in \mathbb{R}_+^K$ with $\sum_k \alpha_k = 1$, and some pmf as described in* (7) *and* (8)*, a rate tuple $(R_1, \ldots, R_K)$ is achievable for the K-user BC with state feedback if*

$$R_k \leq \alpha_1 I(V_k; Y_k^{(1)}, \{\hat{Y}_{1\to\mathcal{U}}\}_{\mathcal{U}\ni k} \,|\, S^{(1)}, Q^{(1)}), \quad (9)$$

$$0 \leq \min_{\substack{i,j,k,\mathcal{J}:\\ i<j,k\in\mathcal{J}}} \Big\{\alpha_j I(V_{i\to\mathcal{J}}; Y_k^{(j)}, \{\hat{Y}_{j\to\mathcal{U}}\}_{\mathcal{U}\supset\mathcal{J}} \,|\, S^{(j)}, Q^{(j)}) - \alpha_i I(\{V_\mathcal{I}\}_{\mathcal{I}\subset\mathcal{J}}; \hat{Y}_{i\to\mathcal{J}} \,|\, Y_k^{(i)}, S^{(i)}, Q^{(i)})\Big\}. \quad (10)$$

Proof. The proof is relegated to Section IV. □

A. Symmetric Capacity of the Erasure Broadcast Channel

Corollary 1. *The symmetric capacity of the i.i.d. EBC can be achieved with the proposed scheme for any input alphabet size, namely,*

$$C_{sym} = \left(\sum_{k=1}^{K} \frac{1}{1-\delta^k}\right)^{-1} \log|\mathcal{X}|. \tag{11}$$

Proof. The converse proof is provided in [3], [4]. To prove the achievability, we apply the following auxiliary RVs to the achievable rate region in Theorem 1.

- Time-sharing RVs $Q^{(j)} \triangleq (Q_1^{(j)}, Q_2^{(j)}) \in \mathcal{Q}^{(j)}$ where

  $$\mathcal{Q}^{(j)} \triangleq \begin{cases} \{(q_1,q_2): \quad q_1 = 0, q_2 \in [1:K]\}, & j=1 \\ \{(q_1,q_2): \quad q_1 \in [1:j-1], q_2 = \mathcal{J}\}, & j \geq 2 \end{cases}$$

  with $|\mathcal{Q}^{(1)}| = K$ and $|\mathcal{Q}^{(j)}| = (j-1)\binom{K}{j}$, for $j \geq 2$. We let

  $$p(q^{(j)}) = p(q_1^{(j)})\binom{K}{j}^{-1}, \quad \text{if } q^{(j)} \in \mathcal{Q}^{(j)} \tag{12}$$

  and 0 otherwise.
- Uniform V's in $\mathcal{X}$: for any $i < j$ and $\mathcal{J}$, we set $\mathcal{V}_{i\to\mathcal{J}} = \mathcal{X}$, and

  $$p(v_{i\to\mathcal{J}}) = |\mathcal{X}|^{-1}, \quad \text{if } v_{i\to\mathcal{J}} \in \mathcal{X} \tag{13}$$

  and 0 otherwise.
- Input as deterministic function of V and Q:

  $$X^{(j)} = V_{Q_1^{(j)} \to Q_2^{(j)}}. \tag{14}$$

- Side information $\hat{Y}$ as deterministic function of (X,S,Q): for any $i<j$ and $\mathcal{J}$,

  $$\hat{Y}_{i\to\mathcal{J}} = X^{(i)}, \tag{15}$$

  if there exists $\mathcal{I}$ such that $Q_2^{(i)} = \mathcal{I}$, $S_{\mathcal{J}\setminus\mathcal{I}} = \mathbf{1}$, $S_{\mathcal{K}\setminus\mathcal{J}} = \mathbf{0}$, and $S_{\mathcal{I}} \neq \mathbf{1}$, and $\hat{Y}_{i\to\mathcal{J}} = 0$ otherwise.

The computation of the symmetric capacity is quite involved. A sketch of proof is provided in the Appendix. □

The intuition behind the setting of the RV $\hat{Y}_{i\to\mathcal{J}}$ is the following. When a signal $X^{(i)}$ intended for a user group $\mathcal{I}$ is sent, if some of the users in $\mathcal{I}$ do not receive the signal *and* some other users (denoted by $\mathcal{J}\setminus\mathcal{I}$) overhear it, then this signal becomes a side information for user group $\mathcal{J}$ and will be compressed and transmitted in phase j. This idea is inspired by the schemes in [3], [4]. The main difference is that here we only care about *what* side information to be shared among users in *which* user group, but do not deal with *how*. With the linear schemes in [3], [4], however, one has to decide how to combine different side information based on what each user already has and needs. In our scheme, these details are treated with the tools from distributed compression.

B. Symmetric Rate and DoF of Fading Gaussian BC

Corollary 2. *For the Symmetric Fading GBC, the proposed scheme achieves the following symmetric rate:*

$$R_{sym} = \max_{\beta_i \geq 0} \left(K + \sum_{j=2}^{K}\binom{K}{j}\prod_{t=2}^{j}\frac{\sum_{l\leq t} b_{l,t}}{a_t}\right)^{-1} a_1 \tag{16}$$

where, for $t = 1,\ldots,K$,

$$a_t \triangleq \mathbb{E}\log\det\left(\mathbf{I} + \mathsf{snr}\,\mathbf{H}_{\mathcal{T}}^H \mathbf{\Lambda}_t \mathbf{H}_{\mathcal{T}}\right) \tag{17}$$

$$b_{l,t} \triangleq \mathbb{E}\log\det\left(\mathbf{I} + \frac{\mathsf{snr}}{\beta_t}\mathbf{H}_l(\mathbf{I} + \mathsf{snr}\,\mathbf{H}_1^H\mathbf{H}_1)^{-1}\mathbf{H}_l^H\right) \tag{18}$$

with $\mathcal{T} \triangleq \{1\}\cup\{t+1,\ldots,K\}$, $\mathbf{\Lambda}_t \triangleq \mathrm{diag}\{\mathbf{I}_{n_r}, \beta_t^{-1}\mathbf{I}_{(K-t)n_r}\}$.

Proof. We use the following auxiliary RVs.

- Time sharing RVs $Q^{(j)} \triangleq (Q_1^{(j)}, Q_2^{(j)}) \in \mathcal{Q}^{(j)}$ with

  $$p(q^{(j)}) = \binom{K}{j}^{-1}, \quad \text{if } q_1^{(j)} = j-1,\ q_2^{(j)} = \mathcal{J} \tag{19}$$

  and 0 otherwise.
- Gaussian distributed V's: for any $i = j-1$ and $\mathcal{J}$, we let $V_{i\to\mathcal{J}} \sim \mathcal{CN}(\mathbf{0}, \frac{P}{n_t}\mathbf{I}_{n_t})$.
- Input as deterministic function of V and Q:

  $$\mathbf{X}^{(j)} = V_{Q_1^{(j)}\to Q_2^{(j)}}. \tag{20}$$

- Side information $\hat{Y}$ as noisy function of (X,S,Q), for any $i = j-1$ and $\mathcal{J}$,

  $$\hat{\mathbf{Y}}_{i\to\mathcal{J}} = \mathbf{H}_{\mathcal{J}\setminus\mathcal{I}}\mathbf{X}^{(i)} + \hat{\mathbf{Z}}_{\mathcal{J}\setminus\mathcal{I}},\ \hat{\mathbf{Z}}_{\mathcal{J}\setminus\mathcal{I}} \sim \mathcal{CN}(\mathbf{0}, \beta_i\sigma^2\mathbf{I})$$

  if $Q_2^{(i)} = \mathcal{I}$, and $\hat{\mathbf{Y}}_{i\to\mathcal{J}} = 0$ otherwise.

A sketch of proof of (16) is provided in the Appendix. □

As in the EBC case, we explain the intuition behind the choice of the side information $\hat{Y}_{i\to\mathcal{J}}$. When a signal $\mathbf{X}^{(i)}$ that is intended for a user group $\mathcal{I}$ is transmitted, user k, outside of the group, overhears the noisy version of the signal $\mathbf{H}_k\mathbf{X}^{(i)}$ which is useful to all users in group $\mathcal{I}$. Therefore, $\mathbf{H}_k\mathbf{X}^{(i)}$ becomes side information to be shared among users in the group $\mathcal{J} \triangleq \{k\}\cup\mathcal{I}$. Since $\mathbf{H}_k\mathbf{X}^{(i)}$ can be costly to send, only a noisy (compressed) version, namely, $\mathbf{H}_k\mathbf{X}^{(i)} + \hat{\mathbf{Z}}_k$ is used instead. Again, as mentioned earlier, we do not need to explicitly deal with *how* to share the side information.

By setting $\beta_i = 1$, $\forall\, i$, in (16) and letting $\mathsf{snr} \to \infty$, we can verify that the following optimal symmetric DoF can be achieved in the MISO case [5],

$$\mathsf{DoF}_{\mathrm{sym}} = \left(\sum_{k=1}^{K}\frac{1}{k}\right)^{-1}, \tag{21}$$

when $n_{\mathrm{t}} \geq K$ and $n_{\mathrm{r},1} = \cdots = n_{\mathrm{r},K} = 1$. Details are provided in the Appendix.

IV. Description of the Proposed Scheme

We divide the n-slot transmission into K phases with phase j of length n_j and define $\alpha_j \triangleq \frac{n_j}{n}$ with $\sum_{j=1}^{K} \alpha_j = 1$. In phase 1, K independent source messages, $M_k \in \mathcal{M}_k \triangleq [1 : 2^{nR_k}]$, $k \in \mathcal{K}$, are sent. In phase j, $j = 2, \ldots, K$, $(j-1)\binom{K}{j}$ messages, $M_{i\to\mathcal{J}} \in \mathcal{M}_{i\to\mathcal{J}} \triangleq [1 : 2^{n_j R_{i\to\mathcal{J}}}]$, carrying side information on the past phase i and intended for users in $\mathcal{J}$, are created and transmitted.

Codebook Generation

1) Randomly generate the time sharing sequence according to $\prod_{j=1}^{K}\prod_{t=1}^{n_j} q_t^{(j)}$.
2) Randomly generate 2^{nR_k} independent sequences $\boldsymbol{v}_k$ according to $\prod_{i=1}^{n_1} p(v_{k,i})$, and index them as $\boldsymbol{v}_k(m_k)$ with $m_k \in [1 : 2^{nR_k}]$, $k \in \mathcal{K}$.
3) At the end of phase i, $i = 1, \ldots, K-1$, randomly generate $2^{n_i R_{i\to\mathcal{J}}}$ independent sequences $\hat{\boldsymbol{y}}_{i\to\mathcal{J}}$ and $\boldsymbol{v}_{i\to\mathcal{J}}$, for each $j > i$ and each $|\mathcal{J}| = j$, according to $\prod_{t=1}^{n_i} p(\hat{y}_{i\to\mathcal{J}}[t] \,|\, s_t^{(i)}, q_t^{(i)})$ and $\prod_{t=1}^{n_j} p(v_{i\to\mathcal{J}}[t])$, respectively, and index them as $\hat{\boldsymbol{y}}_{i\to\mathcal{J}}(m_{i\to\mathcal{J}})$ and as $\boldsymbol{v}_{i\to\mathcal{J}}(m_{i\to\mathcal{J}})$, respectively, with $m_{i\to\mathcal{J}} \in [1 : 2^{n_i R_{i\to\mathcal{J}}}]$.

Encoding

1) In phase 1, the encoder selects an arbitrary $\boldsymbol{x}^{(1)}$ from the set of sequences that are jointly typical with $(\boldsymbol{v}_1(M_1), \cdots, \boldsymbol{v}_K(M_K), \boldsymbol{q}^{(1)})$, and then sends out $\boldsymbol{x}^{(1)}$.
2) At the end of phase i, $i = 1, \ldots, K-1$, and for each $j > i$ and each $|\mathcal{J}| = j$, the source searches for some $M_{i\to\mathcal{J}}$ such that $(\hat{\boldsymbol{y}}_{i\to\mathcal{J}}(M_{i\to\mathcal{J}}), \{\boldsymbol{v}_{l\to\mathcal{I}}\}_{l<i,\mathcal{I}\subset\mathcal{J}}, \boldsymbol{s}^{(i)}, \boldsymbol{q}^{(i)})$ are jointly typical, which has high probability if

$$n_i R_{i\to\mathcal{J}} \geq n_i I(\hat{Y}_{i\to\mathcal{J}}; \{V_{\mathcal{I}}\}_{\mathcal{I}\subset\mathcal{J}} \,|\, S^{(i)}, Q^{(i)}) + n_i \epsilon_{n_i}.$$

In phase j, $j = 2, \ldots, K$, the source selects and transmits an arbitrary $\boldsymbol{x}^{(j)}$ from the set of sequences that are jointly typical with $\big(\{\boldsymbol{v}_{i\to\mathcal{J}}(M_{i\to\mathcal{J}})\}_{i<j,\mathcal{J}}, \boldsymbol{q}^{(j)}\big)$.

Decoding

We focus on user k, who performs backward decoding on the desired messages $\{M_{\mathcal{U}}\}_{\mathcal{U}\ni k}$, starting from phase K.

1) At phase j, $j = K, \ldots, 2$, and for each $\mathcal{J} \ni k$ and $i < j$, the decoder looks for a unique $\hat{M}_{i\to\mathcal{J}}$ such that $\big(\boldsymbol{v}_{i\to\mathcal{J}}(\hat{M}_{i\to\mathcal{J}}), \boldsymbol{y}_k^{(j)}, \{\hat{\boldsymbol{y}}_{j\to\mathcal{U}}(\hat{M}_{j\to\mathcal{U}})\}_{\mathcal{U}\supset\mathcal{J}}, \boldsymbol{s}^{(j)}, \boldsymbol{q}^{(j)}\big)$ are jointly typical and simultaneously that $\big(\hat{\boldsymbol{y}}_{i\to\mathcal{J}}(\hat{M}_{i\to\mathcal{J}}), \boldsymbol{y}_k^{(i)}, \boldsymbol{s}^{(i)}, \boldsymbol{q}^{(i)}\big)$ are jointly typical. It can be shown that such $\hat{M}_{i\to\mathcal{J}}$ can be found and $\hat{M}_{i\to\mathcal{J}} = M_{i\to\mathcal{J}}$ with high probability provided that [10]

$$\begin{aligned} n_i R_{i\to\mathcal{J}} \leq & n_j I(V_{i\to\mathcal{J}}; Y_k^{(j)}, \{\hat{Y}_{j\to\mathcal{U}}\}_{\mathcal{U}\supset\mathcal{J}} \,|\, S^{(j)}, Q^{(j)}) \\ & + n_i I(\hat{Y}_{i\to\mathcal{J}}; Y_k^{(i)} \,|\, S^{(i)}, Q^{(i)}) - n_i \epsilon_{n_i}. \quad (22) \end{aligned}$$

2) Finally, the decoder searches for a unique $\hat{M}_k$ such that $\big(\boldsymbol{v}_k(\hat{M}_k), \boldsymbol{y}_k^{(1)}, \{\hat{\boldsymbol{y}}_{1\to\mathcal{U}}(\hat{M}_{1\to\mathcal{U}})\}_{\mathcal{U}\ni k}, \boldsymbol{s}^{(1)}, \boldsymbol{q}^{(1)}\big)$ are jointly typical. Such $\hat{M}_k$ can be found and $\hat{M}_k = M_k$ with high probability if

$$nR_k \leq n_1 I(V_k; Y_k^{(1)}, \{\hat{Y}_{1\to\mathcal{U}}\}_{\mathcal{U}\ni k} \,|\, S^{(1)}, Q^{(1)}) - n_1 \epsilon_{n_1}.$$

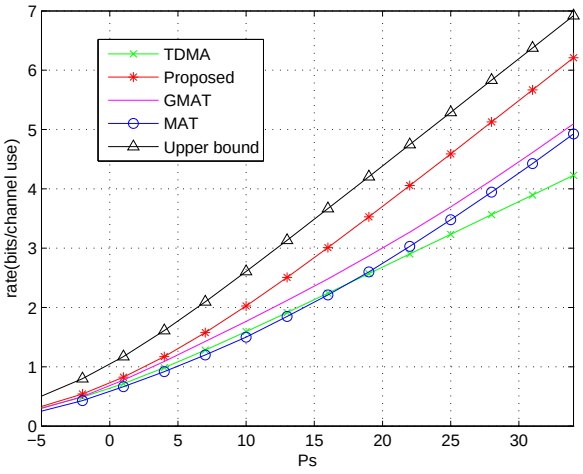

Fig. 2. Symmetric rate versus SNR in the symmetric three-user fading MISO channel, with $n_{\mathrm{t}} = 3$ transmit antennas.

We let $n_i \to \infty$, $i \in \mathcal{K}$, and apply Fourier-Motzkin elimination to all constraints to get the rate region given in Theorem 1.

It is worth mentioning that Marton coding is not used in our scheme, which makes the achievable region less general than [6] for two-user case. As a matter of fact, the choice has been made due to our setup of K users. Introducing Marton coding would lead to a rate region intractable and hard to evaluate.

V. Numerical Example

In this section, we consider the three-user Gaussian MISO channel with i.i.d. Rayleigh fading. We let $n_{\mathrm{t}} = K = 3$ and focus on the symmetric rate (16). The maximization over $\{\beta_i\}$ is done numerically. In evaluation, we compare the proposed scheme with four cases: 1) TDMA (optimal for no state feedback case), 2) the analog MAT scheme [5], 3) the generalized MAT (GMAT) scheme [9], and 4) an upper bound. In both MAT and GMAT scheme, we apply the precoding so that the power constraint is respected in each time slot. For the outer bound, we apply the genie-aided argument as in [5]. Namely, by providing the output Y_1 to receiver 2, and providing both outputs Y_1 and Y_2 to receiver 3, we obtain an outer bound from the degraded channel whose region can be described by single-letter constraints: $R_1 \leq I(U_1; Y_1)$, $R_2 \leq I(U_2; Y_1 Y_2 \,|\, U_1)$, $R_3 \leq I(X; Y_1 Y_2 Y_3 \,|\, U_1 U_2)$. Then, in the symmetric Rayleigh channel, one can show that $6R_1 + 3R_2 + 2R_3 \leq 6I(X; Y_1)$ which implies that $C_{\mathrm{sym}} \leq \frac{6}{11} C_{\mathrm{SU}}$. Note that the TDMA scheme achieves $\frac{1}{3} C_{\mathrm{SU}}$ as symmetric rate.

In Fig. 2, the symmetric rate versus SNR is plotted. In all SNR regions, the proposed scheme outperforms three reference schemes. Specifically in medium-to-high SNR regime, our scheme has a non-negligible power gain over MAT and GMAT schemes, shown by the gap between curves. This is mainly due to the flexibility over the duration of each phase (time-slot) and the compression parameters as a function of the SNR, which is not possible with the MAT/GMAT schemes as they are described in [9].

VI. Summary and Discussion

We proposed a general scheme for the state-dependent K-user BC with state feedback. This scheme, with a proper choice of auxiliary RVs, achieves the symmetric capacity in the Erasure BC, the optimal DoF in a Fading Gaussian BC as well as a higher rate than the existing schemes. Different from existing schemes in the same setting, the main ingredient of the proposed scheme is distributed compression of side information. Owing to its simplicity, the proposed region may be applied to other channels in a rather straightforward way.

Appendix

A. Sketch of Proof of Expression (11)

We first evaluate the quantities in (9) and (10), that is:

$$I(V_k; Y_k^{(1)}, \{\hat{Y}_{1\to\mathcal{U}}\}_{\mathcal{U}\ni k} \,|\, S^{(1)}, Q^{(1)}) = P(Q_2^{(1)} = k)(1-\delta^K)\log|\mathcal{X}| \quad (23)$$

$$I(V_{i\to\mathcal{J}}; Y_k^{(j)}, \{\hat{Y}_{j\to\mathcal{U}}\}_{\mathcal{U}\supset\mathcal{J}} \,|\, S^{(j)}, Q^{(j)}) = P(Q_1^{(j)} = i)\binom{K}{j}^{-1}(1-\delta^{K-j+1})\log|\mathcal{X}| \quad (24)$$

$$I(\{V_{\mathcal{I}}\}_{\mathcal{I}\subset\mathcal{J}}; \hat{Y}_{i\to\mathcal{J}} \,|\, Y_k^{(i)}, S^{(i)}, Q^{(i)}) = \binom{j-1}{i-1}\binom{K}{i}^{-1}\delta^{K-j+1}(1-\delta)^{j-i}\log|\mathcal{X}|. \quad (25)$$

Using $\sum_{i=1}^{j-1} P(Q_1^{(j)} = i) = 1$ and applying (10), we can eliminate $P(Q_1^{(j)} = i)$ to obtain $K-1$ constraints on the α's, namely, for $j = 2, \ldots, K$,

$$\frac{\alpha_j}{\binom{K}{j}} \geq \frac{\delta^{K-j+1}}{(1-\delta^{K-j+1})}\sum_{i=1}^{j-1}\binom{j-1}{i-1}(1-\delta)^{j-i}\frac{\alpha_i}{\binom{K}{i}}, \quad (26)$$

which is equivalent to $K-1$ lower bounds on the $K-1$ ratios $\frac{\alpha_2}{\alpha_1}, \ldots, \frac{\alpha_K}{\alpha_1}$. Since we have the sum constraint $\sum_{j=1}^{K}\alpha_j = 1$, and that the goal is to maximize α_1 in order to maximize the rate R_k in (9), it follows that the ratios $\left\{\frac{\alpha_j}{\alpha_1}\right\}_{j>1}$ should be minimized. Thus, each ratio should meet its own lower limit, i.e., equality must be attained in (26). In this case, the $K-1$ ratios can be completely determined from the $K-1$ equalities. Finally, α_1 can be obtained by

$$\alpha_1 = \left(1 + \sum_{j>1}\frac{\alpha_j}{\alpha_1}\right)^{-1} = \frac{K}{(1-\delta^K)\sum_{j=1}^{K}(1-\delta^j)^{-1}}. \quad (27)$$

Due to the lack of space, details of the above steps will be reported in an extended version of the paper.

B. Sketch of Proof of Expression (16)

We apply the RVs defined in Section III-B, and compute the quantities in (9) and (10), that is:

$$I(V_k; Y_k^{(1)}, \{\hat{Y}_{1\to\mathcal{U}}\}_{\mathcal{U}\ni k} \,|\, S^{(1)}, Q^{(1)}) = K^{-1}a_1 \quad (28)$$

$$I(V_{i\to\mathcal{J}}; Y_k^{(j)}, \{\hat{Y}_{j\to\mathcal{U}}\}_{\mathcal{U}\supset\mathcal{J}} \,|\, S^{(j)}, Q^{(j)}) = \binom{K}{j}^{-1}a_j \quad (29)$$

$$I(\{V_{\mathcal{I}}\}_{\mathcal{I}\subset\mathcal{J}}; \hat{Y}_{i\to\mathcal{J}} \,|\, Y_k^{(i)}, S^{(i)}, Q^{(i)}) = \sum_{l=1}^{i}\binom{K}{i}^{-1}b_{l,i} \quad (30)$$

where the last two equalities hold only when $j = i+1$; for other i, j, both quantities are zeros according to the values of $Q^{(i)}$ and $Q^{(j)}$. Thus, we obtain:

$$\alpha_j \geq \frac{\binom{K}{j}\sum_{l\leq j} b_{l,j}}{\binom{K}{j-1}a_j}\alpha_{j-1} \geq \cdots \geq \frac{\binom{K}{j}}{K}\prod_{t=2}^{j}\frac{\sum_{l\leq t} b_{l,t}}{a_t}\alpha_1$$

Applying the previous reasoning, each α_j should meet its lower bound to maximize α_1, which means that the above inequality provides the exact optimal ratio $\frac{\alpha_j}{\alpha_1}$. From (27), we obtain the exact optimal α_1, which combining with (28) and (9) leads to (16).

C. Proof of Expression (21)

We let $\beta_i = 1, \forall\, i$. One can verify, from (17) and (18), that, at high SNR,

$$a_t = |\mathcal{T}|\log\mathsf{snr} + o(\log\mathsf{snr}), \quad (31)$$

$$b_{l,t} = \begin{cases} o(\log\mathsf{snr}), & l = 1 \\ \log\mathsf{snr} + o(\log\mathsf{snr}), & l \neq 1 \end{cases} \quad (32)$$

provided that $n_{\mathrm{t}} \geq K$. Note that $|\mathcal{T}| = K - t + 1$, we have

$$\frac{R_{\mathrm{sym}}}{\log\mathsf{snr}} \overset{\mathsf{snr}\to\infty}{\longrightarrow} \left(K + \sum_{j=2}^{K}\binom{K}{j}\prod_{t=2}^{j}\frac{t-1}{K-t+1}\right)^{-1}K \quad (33)$$

$$= \left(1 + \frac{1}{K}\sum_{j=2}^{K}\binom{K}{j}\binom{K-1}{j-1}^{-1}\right)^{-1} \quad (34)$$

$$= \left(\sum_{j=1}^{k} j^{-1}\right)^{-1}. \quad (35)$$

References

[1] G. Dueck, "Partial feedback for two-way and broadcast channels," *Information and Control*, vol. 46, no. 1, pp. 1–15, 1980.

[2] L. Ozarow and S. Leung-Yan-Cheong, "An achievable region and outer bound for the Gaussian broadcast channel with feedback (corresp.)," *IEEE Trans. Inf. Theor.*, vol. 30, no. 4, pp. 667–671, July 1984.

[3] C.-C. Wang, "On the capacity of 1-to-K broadcast packet erasure channels with channel output feedback," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 931–956, Feb 2012.

[4] M. Gatzianas, L. Georgiadis, and L. Tassiulas, "Multiuser broadcast erasure channel with feedback-capacity and algorithms," *IEEE Trans. Inf. Theory*, vol. 59, no. 9, pp. 5779–5804, May 2013.

[5] 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.

[6] O. Shayevitz and M. Wigger, "On the capacity of the discrete memoryless broadcast channel with feedback," *IEEE Trans. Inf. Theory*, vol. 59, no. 3, pp. 1329–1345, Mar 2013.

[7] H. Kim, Y.-K. Chia, and A. El Gamal, "A note on the broadcast channel with stale state information at the transmitter," *IEEE Trans. Inf. Theory*, vol. 61, no. 7, pp. 3622–3631, May 2015.

[8] A. Vahid, M. A. Maddah-Ali, and A. S. Avestimehr, "Approximate capacity of the two-user MISO broadcast channel with delayed CSIT," in *51st Annual Allerton Conference on Communication, Control, and Computing*, Oct 2013, pp. 1136–1143.

[9] X. Yi and D. Gesbert, "Precoding methods for the MISO broadcast channel with delayed CSIT," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 1–11, May 2013.

[10] E. Tuncel, "Slepian-Wolf coding over broadcast channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1469–1482, April 2006.