A Note on the Broadcast Channel With Stale State Information at the Transmitter

Hyeji Kim, *Student Member, IEEE*, Yeow-Khiang Chia, and Abbas El Gamal, *Fellow, IEEE*

***Abstract*— This paper shows that the Maddah-Ali–Tse (MAT) scheme, which achieves the symmetric capacity of two example broadcast channels with strictly causal state information at the transmitter, is a simple special case of the Shayevitz–Wigger (SW) scheme for the broadcast channel with generalized feedback, which involves block Markov coding, Gray–Wyner compression, superposition coding, and Marton coding. Focusing on the class of symmetric broadcast channels with state, we derive an expression for the maximum achievable symmetric rate using the SW scheme. We show that the MAT results for the two-receiver case can be recovered by evaluating this expression for the special case in which superposition coding and Marton coding are not used. We then introduce a new broadcast channel example that shares many features of the MAT examples. We show that another special case of our maximum symmetric rate expression in which superposition coding is also used attains a higher symmetric rate than the MAT scheme. The symmetric capacity of this new example is not known, however.**

***Index Terms*— Broadcast channel, channel with state, feedback, source coding with side information.**

I. Introduction

IT IS well known that a broadcast channel with random state $p(y_1, y_2|x, s)p(s)$ when the state S is known at the decoders can be viewed as a broadcast channel with the same input X but with outputs (Y_1, S) and (Y_2, S) [1, Ch. 7]. If the state is also known *strictly causally* at the encoder, i.e., the encoder at time i knows S^{i-1}, then the setup can be viewed as a broadcast channel with outputs (Y_1, S) and (Y_2, S) and *causal feedback* of part of the outputs (see Fig. 1). Hence the broadcast channel with state known at the decoders and strictly casually at the encoder is intimately related to the broadcast channel with *generalized* feedback [2], and it is expected that results for one of these two settings can be readily translated into results for the other.

Dueck was the first to show via an insightful example [3] that feedback can enlarge the capacity region of the broadcast channel. The key idea in Dueck's example is for the encoder to broadcast past common information about the channel obtained through feedback. Even though the channel is memoryless, knowledge of this *stale* common information at the decoders helps them recover previous messages at a higher rate than without feedback. This key idea inspired Shayevitz and Wigger to develop a block Markov coding scheme for the broadcast channel with generalized feedback [2]. In their scheme, new messages are sent in each transmission block together with *refinement* information about the previous messages based on the channel information obtained through feedback. The refinement information is obtained by compressing the previous codewords in a manner similar to the Gray-Wyner system with side information [4]. The encoder uses Marton coding, superposition coding, and coded time sharing to encode the messages and the refinement information. Decoding is performed backwards with the refinement information recovered in a block used to recover the messages and the refinement information sent in the previous block.

In a separate line of investigation which is motivated by fading broadcast channels and network coding, Maddah-Ali and Tse [5] demonstrated via two beautiful example channels that strictly causal (stale) state information at the encoder can enlarge the capacity region of the broadcast channel with state when the state is also known at the decoders. In the 2-receiver special case of their scheme, which establishes the symmetric capacity of the example channels, transmission is performed over three blocks. In the first block, the message intended for the first receiver is sent at a rate higher than what it can reliably decode. In the second block, the message for the second receiver is sent again at a rate higher than what it can decode. In the third block, refinement information about the messages that depends on the state information from the first two blocks is sent to both receivers to enable them to decode their respective messages.

In this paper, we show that this Maddah-Ali–Tse (MAT) scheme is a simple special case of a straightforward adaptation of the Shayevitz–Wigger (SW) scheme. We consider a class of symmetric broadcast channels with state that includes the MAT 2-receiver examples as special cases and derive an expression for its maximum achievable symmetric rate using the SW scheme. We then consider the symmetric *deterministic* broadcast channels with state, which again includes the MAT examples as special cases. We further specialize our expression for the maximum symmetric rate to the SW scheme with no superposition coding or Marton coding, henceforth referred to as the *time-sharing scheme*. We show that the maximum symmetric rate for this time-sharing scheme is optimal for the MAT 2-receiver examples and is in fact a simple extension of their scheme. Observing that in both of the MAT examples the channel is deterministic for each state (in addition to being symmetric), we investigate the

Manuscript received October 1, 2013; revised October 8, 2014; accepted January 30, 2015. Date of publication May 19, 2015; date of current version June 12, 2015. This work was supported by the Air Force under Grant FA9550-10-1-0124.

H. Kim and A. El Gamal are with the Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA (e-mail: hyejikim@ stanford.edu; abbas@ee.stanford.edu).

Y.-K. Chia is with the Institute for Infocomm Research, Singapore 138632 (e-mail: yeowkhiang@gmail.com).

Communicated by Y. Liang, Associate Editor for Shannon Theory.

Digital Object Identifier 10.1109/TIT.2015.2428237

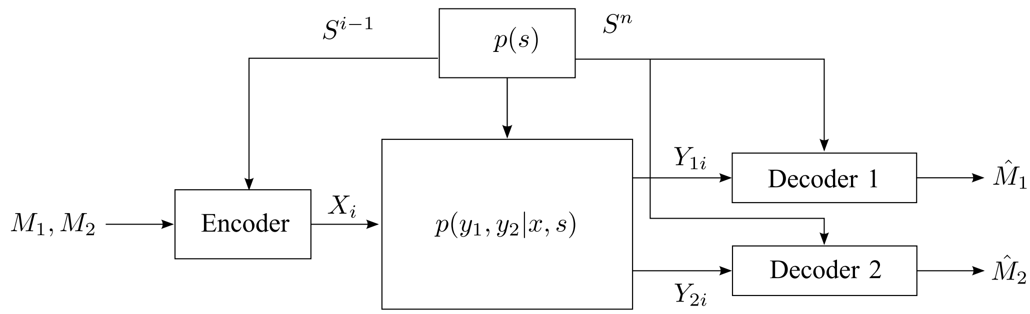

Fig. 1. Two receiver DM-BC with stale state information at the transmitter.

question of whether the time-sharing scheme is optimal for all such deterministic channels. We construct a new example in which the channel switches between a Blackwell broadcast channel [6] and a skew symmetric version of it, and show that another special case of the SW scheme that includes superposition coding, henceforth referred to as the *superposition coding scheme*, achieves a higher symmetric rate than the time-sharing scheme. We do not know, however, if the SW scheme in its full generality is optimal for this channel, or for the aforementioned deterministic class in general.

The rest of the paper is organized as follows. In the following section, we provide the needed definitions. In Section III, we adapt the Shayevitz–Wigger scheme to the broadcast channel with stale state information and derive an expression for the maximum achievable symmetric rate when the channel is symmetric. In Section IV, we specialize this expression to the time-sharing scheme for the symmetric deterministic channels and evaluate the expression to show that the time-sharing scheme is optimal for the Maddah-Ali–Tse examples. In Section V, we specialize our maximum symmetric rate expression to the superposition coding scheme for symmetric deterministic channels and introduce the Blackwell broadcast channel with state. We show that the maximum symmetric rate using the superposition coding scheme is strictly higher than using the time-sharing scheme. We also obtain an upper bound on the symmetric capacity for this example.

II. Definitions

A 2-receiver DM-BC with generalized feedback consists of an input alphabet $\mathcal{X}$, two output alphabets $(\mathcal{Y}_1, \mathcal{Y}_2)$, a feedback alphabet $\tilde{\mathcal{Y}}$, and a conditional pmf $p(y_1, y_2, \tilde{y}|x)$. A $(2^{nR_1}, 2^{nR_2}, n)$ code for the DM-BC with generalized feedback consists of (i) two message sets $[1 : 2^{nR_1}]$ and $[1 : 2^{nR_2}]$; (ii) an encoder that assigns a symbol $x_i(m_1, m_2, \tilde{y}^{i-1})$ to each message tuple $(m_1, m_2) \in [1 : 2^{nR_1}] \times [1 : 2^{nR_2}]$ and received sequence $\tilde{y}^{i-1}$ for $i \in [1 : n]$, and (iii) two decoders. Decoder 1 assigns an estimate $\hat{m}_1 \in [1 : 2^{nR_1}]$ or an error message e to each received sequence y_1^n. Decoder 2 assigns $\hat{m}_2 \in [1 : 2^{nR_2}]$ or an error message e to each received sequence y_2^n.

A 2-receiver DM-BC with random state consists of an input alphabet $\mathcal{X}$, two output alphabets $(\mathcal{Y}_1, \mathcal{Y}_2)$, a discrete memoryless state $S \sim p(s)$, and a conditional pmf $p(y_1, y_2|x, s)$. We consider the case in which the decoders know the state and the encoder knows the state strictly causally (or *stale* state in short) depicted in Fig. 1.

A $(2^{nR_1}, 2^{nR_2}, n)$ code for this setup consists of (i) two message sets $[1 : 2^{nR_1}]$ and $[1 : 2^{nR_2}]$, (ii) an encoder that assigns a symbol $x_i(m_1, m_2, s^{i-1})$ to each message tuple $(m_1, m_2) \in [1 : 2^{nR_1}] \times [1 : 2^{nR_2}]$ and received sequence s^{i-1} for $i \in [1 : n]$, and (iii) two decoders. Decoder 1 assigns an estimate $\hat{m}_1 \in [1 : 2^{nR_1}]$ or an error message e to each received sequence (y_1^n, s^n). Decoder 2 assigns $\hat{m}_2 \in [1 : 2^{nR_2}]$ or an error message e to each received sequence (y_2^n, s^n).

For both setups, the probability of error is defined as

$$P_e^{(n)} = \mathsf{P}\{\hat{M}_1 \neq M_1 \text{ or } \hat{M}_2 \neq M_2\}.$$

Similarly, in both cases, a rate tuple (R_1, R_2) is said to be achievable if there exists a sequence of $(2^{nR_1}, 2^{nR_2}, n)$ codes such that $P_e^{(n)} \to 0$ as $n \to \infty$. The capacity region is defined as the closure of the set of all achievable rate tuples.

Remark 1: From the above definitions, the latter setup can be viewed as a special case of the former. To see this, let $(X, Y_1, Y_2, \tilde{Y})$ be the random variables associated with the first setup, and (X', Y_1', Y_2', S) be the random variables associated with the second setup. Then set $X = X'$, $Y_1 = (Y_1', S)$, $Y_2 = (Y_2', S)$, and $\tilde{Y} = S$. Under this mapping, any coding scheme for the latter case is also a coding scheme for the former case.

In this paper, we will consider only the following special classes of channels.

Definition 1 (Symmetric 2-Receiver DM-BC With Random State): A 2-receiver DM-BC with random state is said to be symmetric if $\mathcal{Y}_1 = \mathcal{Y}_2 = \mathcal{Y}$, $\mathcal{S} = \{1, \ldots, |\mathcal{S}|\}$, and there exists a bijective function $\pi : \mathcal{S} \to \mathcal{S}$ such that

$$p_S(s) = p_S(\pi(s)),$$
$$p_{Y_1|X,S}(y|x, s) = p_{Y_2|X,S}(y|x, \pi(s)).$$

Definition 2 (Symmetric Deterministic 2-Receiver DM-BC With Random State): A symmetric 2-receiver DM-BC with random state is said to be deterministic if the outputs are deterministic functions of the input and the state, i.e., $Y_1 = y_1(X, S)$ and $Y_2 = y_2(X, S)$.

The examples in [5] and our new example in Section V all belong to this class of symmetric deterministic DM-BC with random state.

III. Maximum Symmetric Rate for the Shayevitz–Wigger Scheme

The Shayevitz–Wigger achievable rate region for the 2-receiver DM-BC with generalized feedback is given in the following.

Theorem 1: A rate pair (R_1, R_2) *is achievable for the 2-receiver DM-BC with generalized feedback if*

$$\begin{aligned} R_1 &\le I(U_0, U_1; Y_1, V_1|Q) \\ &\quad - I(U_0, U_1, U_2, \tilde{Y}; V_0, V_1|Q, Y_1), \\ R_2 &\le I(U_0, U_2; Y_2, V_2|Q) \\ &\quad - I(U_0, U_1, U_2, \tilde{Y}; V_0, V_2|Q, Y_2), \\ R_1 + R_2 &\le \min_{i\in\{1,2\}} I(U_0; Y_i, V_i|Q) \\ &\quad + \sum_{i=1,2} I(U_i; Y_i, V_i|Q, U_0) - I(U_1; U_2|Q, U_0) \\ &\quad - \max_{i\in\{1,2\}} I(U_0, U_1, U_2, \tilde{Y}; V_0|Q, Y_i) \\ &\quad - \sum_{i=1,2} I(U_0, U_1, U_2, \tilde{Y}; V_i|Q, V_0, Y_i), \\ R_1 + R_2 &\le \sum_{i=1,2} I(U_0, U_i; Y_i, V_i|Q) - I(U_1; U_2|Q, U_0) \\ &\quad - \sum_{i=1,2} I(U_0, U_1, U_2, \tilde{Y}; V_0, V_i|Q, Y_i) \end{aligned}$$

for some function $x(u_0, u_1, u_2, q)$ *and pmf*

$$p(q)p(u_0, u_1, u_2|q)p(\tilde{y}|x, y_1, y_2, q) \cdot p(v_0, v_1, v_2|u_0, u_1, u_2, \tilde{y}, q).$$

We now consider the following straightforward corollary of this theorem.

Corollary 1: A rate pair (R_1, R_2) *is achievable for the 2-receiver DM-BC with random state when the state is known at the decoders and strictly causally known at the encoder if*

$$R_1 \le I(U_0, U_1; Y_1, V_1|Q, S) - I(U_0, U_1, U_2; V_0, V_1|Q, Y_1, S), \tag{1}$$

$$R_2 \le I(U_0, U_2; Y_2, V_2|Q, S) - I(U_0, U_1, U_2; V_0, V_2|Q, Y_2, S), \tag{2}$$

$$\begin{aligned} R_1 + R_2 \le \min_{i\in\{1,2\}} I(U_0; Y_i, V_i|Q, S) &+ \sum_{i=1,2} I(U_i; Y_i, V_i|Q, U_0, S) - I(U_1; U_2|Q, U_0) \\ &- \max_{i\in\{1,2\}} I(U_0, U_1, U_2; V_0|Q, Y_i, S) \\ &- \sum_{i=1,2} I(U_0, U_1, U_2; V_i|Q, V_0, Y_i, S), \end{aligned} \tag{3}$$

$$R_1 + R_2 \le \sum_{i=1,2} I(U_0, U_i; Y_i, V_i|Q, S) - I(U_1; U_2|Q, U_0) - \sum_{i=1,2} I(U_0, U_1, U_2; V_0, V_i|Q, Y_i, S) \tag{4}$$

for some function $x(u_0, u_1, u_2, q)$ *and pmf*

$$p(q)p(u_0, u_1, u_2|q)p(v_0, v_1, v_2|u_0, u_1, u_2, s, q).$$

This corollary follows immediately from Remark 1. We give a brief outline of the achievability scheme for the region in the corollary. The details follow [2, proof of Theorem 1]. In Sections IV and V we give more detailed descriptions of two special cases of this scheme. For simplicity, we describe the scheme only for $Q = \emptyset$. The scheme uses block Markov coding to communicate $b - 1$ messages to the first receiver and $b - 1$ messages to the second receiver in b n-transmission blocks. In each block, superposition coding and Marton coding are used to communicate a new message pair with refinement information about the previous codewords. The auxiliary random variable U_0 represents the "cloud center" in superposition coding, and encodes the part of the messages and refinements that is recovered by both decoders. The auxiliary random variables U_1 and U_2 represent the part of the messages and refinements that are recovered only by decoder 1 and decoder 2, respectively. The common and private refinement messages are generated by compressing the previous codewords and feedback (U_0, U_1, U_2, S) in a manner similar to the lossy Gray–Wyner system with side information (Y_1, S) and (Y_2, S). The goal of compression here, however, is for each decoder to generate V_1 and V_2. The auxiliary random variable V_0 helps in recovering V_1 and V_2. Backward decoding [1, Ch. 16] is used to recover the refinements and messages. In each block, decoder $j = 1, 2$ generates an output V_j for the current block using its observation (Y_j, S) as side information and the refinement messages recovered in the next block. It then uses (Y_j, V_j, S) to recover the message in the current block and the refinement messages for the previous block. This procedure continues until each decoder recovers all its intended $b - 1$ messages.

For the rest of this paper, we consider only the symmetric rate for the 2-receiver symmetric DM-BC with random state defined in Section II.

Definition 3 (Maximum Symmetric Rate): Let $\mathcal{R}$ be the achievable rate region in Corollary 1 and R_{sym} be the *maximum symmetric rate* achievable with the scheme of Corollary 1, that is, the supremum of R such that $(R, R) \in \mathcal{R}$. Also, let R_{sum} be the *maximum sum-rate*, that is, the supremum of $R_1 + R_2$ such that $(R_1, R_2) \in \mathcal{R}$.

Because of the restriction to symmetric channels and their symmetric rates, we show in Theorem 2 that it suffices to consider only auxiliary random variables and functions that satisfy the following symmetry property.

Definition 4 (Symmetric Auxiliary Random Variables): Assume without loss of generality that $\mathcal{U}_1 = \mathcal{U}_2 = \mathcal{U}$ and $\mathcal{V}_1 = \mathcal{V}_2 = \mathcal{V}$. A set of auxiliary random variables $(U_0, U_1, U_2, V_0, V_1, V_2, Q)$ and function $X = x(U_0, U_1, U_2, Q)$ is said to be symmetric for a symmetric 2-receiver DM-BC with random state if $\mathcal{Q} = \{1, \ldots, |\mathcal{Q}|\}$ and there exists a bijective function $\tilde{\pi} : \mathcal{Q} \to \mathcal{Q}$ such that

$$\begin{aligned} p_Q(q) &= p_Q(\tilde{\pi}(q)), \\ p(u_0, u_1, u_2|q) &= p(u_0, u_2, u_1|\tilde{\pi}(q)), \\ x(u_0, u_1, u_2, q) &= x(u_0, u_2, u_1, \tilde{\pi}(q)), \\ p(v_0, v_1, v_2|u_0, u_1, u_2, q, s) &= p(v_0, v_2, v_1|u_0, u_2, u_1, \tilde{\pi}(q), \pi(s)), \end{aligned}$$

where $\pi(s)$ is as defined in Definition 1.

For the symmetric 2-receiver DM-BC, the maximum symmetric rate achievable using the coding scheme of Corollary 1 can be greatly simplified.

Theorem 2: The maximum achievable symmetric rate for the symmetric 2-receiver DM-BC with stale state using the coding scheme of Corollary 1 is

$$
\begin{aligned}
R_{\text{sym}} = \max\min
&\Big\{0.5I(U_0; Y_1, V_1|Q_{\text{sym}}, S)\\
&+ I(U_1; Y_1, V_1|Q_{\text{sym}}, U_0, S)\\
&- 0.5I(U_1; U_2|Q_{\text{sym}}, U_0)\\
&- 0.5I(U_0, U_1, U_2; V_0|Q_{\text{sym}}, Y_1, S)\\
&- I(U_0, U_1, U_2; V_1|Q_{\text{sym}}, V_0, Y_1, S),\\
&I(U_0, U_1; Y_1, V_1|Q_{\text{sym}}, S) - 0.5I(U_1;U_2|Q_{\text{sym}}, U_0)\\
&- I(U_0, U_1, U_2; V_0, V_1|Q_{\text{sym}}, Y_1, S)\Big\}, \qquad (5)
\end{aligned}
$$

where the maximization is over symmetric auxiliary random variables and function satisfying the structure given in Corollary 1.

Proof: Suppose R_{sum} is achievable with a set of auxiliary random variables $(U_0, U_1, U_2, V_0, V_1, V_2, Q)$ with $Q \in [1 : N]$ and a function $X = x(U_0, U_1, U_2, Q)$. The sum of the individual bounds (1) and (2) is always greater than or equal to bound (4) on the sum-rate. Hence,

$$
\begin{aligned}
R_{\text{sum}} = \min
&\Big\{\min_{i\in\{1,2\}} I(U_0; Y_i, V_i|Q, S)\\
&+ \sum_{i=1,2} I(U_i; Y_i, V_i|Q, U_0, S) - I(U_1; U_2|Q, U_0)\\
&- \max_{i\in\{1,2\}} I(U_0, U_1, U_2; V_0|Q, Y_i, S)\\
&- \sum_{i=1,2} I(U_0, U_1, U_2; V_i|Q, V_0, Y_i, S),\\
&\sum_{i=1,2} I(U_0, U_i; Y_i, V_i|Q, S) - I(U_1; U_2|Q, U_0)\\
&- \sum_{i=1,2} I(U_0, U_1, U_2; V_0, V_i|Q, Y_i, S)\Big\}. \qquad (6)
\end{aligned}
$$

To show that there exists a set of symmetric auxiliary random variables that attains this sum-rate, first construct the following set of auxiliary random variables and function

$$
\begin{aligned}
&Q' \in [N + 1 : 2N],\\
&p_{Q'}(q) = p_Q(q - N) \text{ for } q \in [N + 1 : 2N],\\
&p_{U_0',U_1',U_2',V_0',V_1',V_2'|Q',S}(u_0, u_1, u_2, v_0, v_1, v_2|q, s)\\
&\quad = p_{U_0,U_1,U_2,V_0,V_1,V_2|Q,S}(u_0,u_2,u_1,v_0,v_2,v_1|q-N, \pi(s)),\\
&X = x(U_0', U_2', U_1', Q' - N). \qquad (7)
\end{aligned}
$$

We show that the following equalities hold.

$$
\begin{aligned}
I(U_0', U_1'; Y_1, V_1'|Q', S) &= I(U_0, U_2; Y_2, V_2|Q, S),\\
I(U_1'; Y_1, V_1'|Q', U_0, S) &= I(U_2; Y_2, V_2|Q, U_0, S),\\
I(U_0'; Y_1, V_1'|Q', S) &= I(U_0; Y_2, V_2|Q, S),\\
I(U_1'; U_2'|Q', U_0') &= I(U_1; U_2|Q, U_0),\\
I(U_0', U_1', U_2'; V_0', V_1'|Q', Y_1, S)&\\
&= I(U_0, U_1, U_2; V_0, V_2|Q, Y_2, S),\\
I(U_0', U_1', U_2'; V_1'|Q', V_0', Y_1, S)&\\
&= I(U_0, U_1, U_2; V_2|Q, V_0, Y_2, S),\\
I(U_0', U_1', U_2'; V_0'|Q', Y_1, S)&\\
&= I(U_0, U_1, U_2; V_0|Q, Y_2, S). \qquad (8)
\end{aligned}
$$

Consider the first equality in (8),

$$
\begin{aligned}
&I(U_0', U_1'; Y_1, V_1'|Q', S)\\
&= \sum_{s\in\mathcal{S}} \sum_{q=N+1}^{2N} p_{Q'}(q)p_S(s)I(U_0', U_1'; Y_1, V_1'|Q'=q, S=s)\\
&\overset{(a)}{=} \sum_{s\in\mathcal{S}} \sum_{q=N+1}^{2N} p_{Q'}(q)p_S(s)I(U_0,U_2;Y_2,V_2|Q=q-N, S=\pi(s))\\
&= \sum_{\pi(s)\in\mathcal{S}} \sum_{q=1}^{N} p_Q(q)p_S(\pi(s))I(U_0, U_2; Y_2, V_2|Q=q, S=\pi(s))\\
&= I(U_0, U_2; Y_2, V_2|Q, S).
\end{aligned}
$$

We prove step (a) by showing that the conditional pmf of $((U_0', U_1', V_1', Y_1)|(Q', S) = (q, s))$ is equal to the conditional pmf of $((U_0, U_2, V_2, Y_2)|(Q, S) = (q - N, \pi(s)))$. We establish a more general result that the conditional pmf of $((U_0', U_1', U_2', V_0', V_1', V_2', X, Y_1, Y_2)|(Q', S) = (q, s))$ is the same as the conditional pmf of $((U_0, U_2, U_1, V_0, V_2, V_1, X, Y_2, Y_1)|(Q, S) = (q-N, \pi(s))$, which can be also used to show the rest of the equalities in (8). Since

$$(V_0', V_1', V_2') \to (U_0', U_1', U_2', Q') \to X \to (Y_1, Y_2)$$

form a Markov chain when S is given, the conditional pmf of $((U_0', U_1', U_2', V_0', V_1', V_2',\ X, Y_1, Y_2)|(Q', S) = (q, s))$ is

$$
\begin{aligned}
&p_{U_0',U_1',U_2'|Q',S}(u_0, u_1, u_2|q, s)\\
&\quad\cdot p_{V_0',V_1',V_2'|U_0',U_1',U_2',Q',S}(v_0, v_1, v_2|u_0, u_1, u_2, q, s)\\
&\quad\cdot p_{X|U_0',U_1',U_2',Q',S}(x|u_0, u_1, u_2, q, s)\\
&\quad\cdot p_{Y_1,Y_2|X,S}(y_1, y_2|x, s)\\
&\overset{(a)}{=} p_{U_0,U_1,U_2|Q,S}(u_0, u_2, u_1|q-N, \pi(s))\\
&\quad\cdot p_{V_0,V_1,V_2|U_0,U_1,U_2,Q,S}(v_0, v_2, v_1|u_0, u_2, u_1, q-N, \pi(s))\\
&\quad\cdot p_{X|U_0,U_1,U_2,Q,S}(x|u_0, u_2, u_1, q-N, \pi(s))\\
&\quad\cdot p_{Y_1,Y_2|X,S}(y_2, y_1|x, \pi(s))\\
&\overset{(b)}{=} p_{U_0,U_2,U_1|Q,S}(u_0, u_1, u_2|q-N, \pi(s))\\
&\quad\cdot p_{V_0,V_2,V_1|U_0,U_2,U_1,Q,S}(v_0, v_1, v_2|u_0, u_1, u_2, q-N, \pi(s))\\
&\quad\cdot p_{X|U_0,U_2,U_1,Q,S}(x|u_0, u_1, u_2, q-N, \pi(s))\\
&\quad\cdot p_{Y_2,Y_1|X,S}(y_1, y_2|x, \pi(s)).
\end{aligned}
$$

Step (a) holds by the symmetry of channels in Definition 1 and the definition of the auxiliary random variables in (7). By changing the order of variables as in step (b), we get the conditional pmf of $((U_0, U_2, U_1, V_0, V_2, V_1, X, Y_2, Y_1)|(Q, S) = (q-N, \pi(s)))$.

The rest of the equalities in (8) can be proved in a similar manner.

Now we compose a new set of auxiliaries that "time-share" between Q and Q'. Let

$$Q_{\text{sym}} = \begin{cases} Q & \text{with probability } 0.5, \\ Q' & \text{with probability } 0.5. \end{cases}$$

It can be easily shown that the resulting set of auxiliary random variables is symmetric.

It follows from (8) that for $j \in \{1, 2\}$,

$$\begin{aligned} I(U_0, U_j; Y_j, V_j|Q_{\text{sym}}, S) &= 0.5 \sum_{i=1,2} I(U_0, U_i; Y_i, V_i|Q, S), \\ I(U_j; Y_j, V_j|Q_{\text{sym}}, U_0, S) &= 0.5 \sum_{i=1,2} I(U_i; Y_i, V_i|Q, U_0, S), \\ I(U_0; Y_j, V_j|Q_{\text{sym}}, S) &= 0.5 \sum_{i=1,2} I(U_0; Y_i, V_i|Q, S), \\ I(U_1; U_2|Q_{\text{sym}}, U_0) &= I(U_1; U_2|Q, U_0), \\ I(U_0, U_1, U_2; V_j|Q_{\text{sym}}, V_0, Y_j, S) &= 0.5 \sum_{i=1,2} I(U_0, U_1, U_2; V_i|Q, V_0, Y_i, S), \\ I(U_0, U_1, U_2; V_0|Q_{\text{sym}}, Y_j, S) &= 0.5 \sum_{i=1,2} I(U_0, U_1, U_2; V_0|Q, Y_i, S). \end{aligned} \tag{9}$$

The sum-rate achievable with Q_{sym} is

$$\begin{aligned} \text{SR} = \min \Bigg\{ & \min_{i\in\{1,2\}} I(U_0; Y_i, V_i|Q_{\text{sym}}, S) \\ & + \sum_{i=1,2} I(U_i; Y_i, V_i|Q_{\text{sym}}, U_0, S) - I(U_1; U_2|Q_{\text{sym}}, U_0) \\ & - \max_{i\in\{1,2\}} I(U_0, U_1, U_2; V_0|Q_{\text{sym}}, Y_i, S) \\ & - \sum_{i=1,2} I(U_0, U_1, U_2; V_i|Q_{\text{sym}}, V_0, Y_i, S), \\ & \sum_{i=1,2} I(U_0, U_i; Y_i, V_i|Q_{\text{sym}}, S) - I(U_1; U_2|Q_{\text{sym}}, U_0) \\ & - \sum_{i=1,2} I(U_0, U_1, U_2; V_0, V_i|Q_{\text{sym}}, Y_i, S) \Bigg\}. \end{aligned} \tag{10}$$

We now show that $\text{SR} \geq R_{\text{sum}}$. By using (9), it can be easily shown that the second term inside the minimum in (10) is the same as the second term inside the minimum in (6). We now show that the first term inside the minimum in (10) is greater than or equal to the first term inside the minimum in (6). We start with the first term in (10).

$$\begin{aligned} & \min_{i\in\{1,2\}} I(U_0; Y_i, V_i|Q_{\text{sym}}, S) \\ & \quad + \sum_{i=1,2} I(U_i; Y_i, V_i|Q_{\text{sym}}, U_0, S) - I(U_1; U_2|Q_{\text{sym}}, U_0) \\ & \quad - \max_{i\in\{1,2\}} I(U_0, U_1, U_2; V_0|Q_{\text{sym}}, Y_i, S) \\ & \quad - \sum_{i=1,2} I(U_0, U_1, U_2; V_i|Q_{\text{sym}}, V_0, Y_i, S) \\ & \overset{(a)}{=} 0.5 \sum_{i=1,2} I(U_0; Y_i, V_i|Q, S) \\ & \qquad + \sum_{i=1,2} I(U_i; Y_i, V_i|Q, U_0, S) - I(U_1; U_2|Q, U_0) \\ & \qquad - 0.5 \sum_{i=1,2} I(U_0, U_1, U_2; V_0|Q, Y_i, S) \\ & \qquad - \sum_{i=1,2} I(U_0, U_1, U_2; V_i|Q, V_0, Y_i, S) \\ & \geq \min_{i\in\{1,2\}} I(U_0; Y_i, V_i|Q, S) \\ & \qquad + \sum_{i=1,2} I(U_i; Y_i, V_i|Q, U_0, S) - I(U_1; U_2|Q, U_0) \\ & \qquad - \max_{i\in\{1,2\}} I(U_0, U_1, U_2; V_0|Q, Y_i, S) \\ & \qquad - \sum_{i=1,2} I(U_0, U_1, U_2; V_i|Q, V_0, Y_i, S), \end{aligned}$$

where (a) holds from (9). Therefore, the maximum sum-rate is achievable with symmetric auxiliary random variables.

Finally we show that $R_{\text{sym}} = 0.5R_{\text{sum}}$ and is achievable with symmetric auxiliary random variables.

In general, $0.5R_{\text{sum}} \geq R_{\text{sym}}$. So we only need to show that $0.5R_{\text{sum}} \leq R_{\text{sym}}$. Equivalently, we show that the rate pair $(0.5R_{\text{sum}}, 0.5R_{\text{sum}})$ is achievable. Using (8), we can show that with symmetric auxiliaries and function, the upper bound on R_1 in (1) and the upper bound on R_2 in (2) are equal. Let them be denoted by R_m, that is,

$$\begin{aligned} R_1 &\leq I(U_0, U_1; Y_1, V_1|Q_{\text{sym}}, S) \\ &\quad - I(U_0, U_1, U_2; V_0, V_1|Q_{\text{sym}}, Y_1, S) := R_m, \\ R_2 &\leq I(U_0, U_2; Y_2, V_2|Q_{\text{sym}}, S) \\ &\quad - I(U_0, U_1, U_2; V_0, V_2|Q_{\text{sym}}, Y_2, S) := R_m. \end{aligned}$$

To show that the rate pair $(0.5R_{\text{sum}}, 0.5R_{\text{sum}})$ is achievable, we only need to show that $0.5R_{\text{sum}} \leq R_m$. The inequalities on the rate sum in (3) and (4) are automatically satisfied because we are given that rate-sum R_{sum} is achievable. To show $0.5R_{\text{sum}} \leq R_m$, first add two inequalities (1) and (2) and get one inequality for any achievable sum-rates, i.e. $R_1 + R_2 \leq 2R_m$. Since R_{sum} is achievable, we get $R_{\text{sum}} \leq 2R_m$.

Therefore, $R_{\text{sym}} = R_{\text{sum}}/2$ and can be written as

$$R_{\text{sym}} = \frac{1}{2} \max \min \Big\{ \min_{i\in\{1,2\}} I(U_0; Y_i, V_i | Q_{\text{sym}}, S) + \sum_{i=1,2} I(U_i; Y_i, V_i | Q_{\text{sym}}, U_0, S) - I(U_1; U_2 | Q_{\text{sym}}, U_0) - \max_{i\in\{1,2\}} I(U_0, U_1, U_2; V_0 | Q_{\text{sym}}, Y_i, S) - \sum_{i=1,2} I(U_0, U_1, U_2; V_i | Q_{\text{sym}}, V_0, Y_i, S), \sum_{i=1,2} I(U_0, U_i; Y_i, V_i | Q_{\text{sym}}, S) - I(U_1; U_2 | Q_{\text{sym}}, U_0) - \sum_{i=1,2} I(U_0, U_1, U_2; V_0, V_i | Q_{\text{sym}}, Y_i, S) \Big\}, \quad (11)$$

where the maximization is over symmetric auxiliaries and functions satisfying the structure in Corollary 1. The proof is completed by further simplifying R_{sym} using the equalities in (9). ■

For the rest of this paper, we consider only symmetric *deterministic* 2-receiver DM-BC with random state as defined in Section II.

IV. Time-Sharing Scheme

In this section, we show that the MAT coding scheme [5] is a special case of the SW coding scheme [2] when adapted to the symmetric deterministic DM-BC with random state without superposition coding or Marton coding. We refer to this special case as the *time-sharing* scheme. Specifically, we specialize the auxiliary random variables in Theorem 2 as follows. Let $Q \in \{1, 2, 3\}$ and $p_Q(1) = p_Q(2) = \alpha$, and $p_Q(3) = 1 - 2\alpha$, $0 \le \alpha \le 0.5$. Let $p(q, u_0, u_1, u_2) = p(q)p(u_0)p(u_1)p(u_2)$ and $p_{U_1}(u) = p_{U_2}(u)$. Define

$$V_1 = V_2 = V_0 = \begin{cases} Y_2 & \text{if } Q = 1, \\ Y_1 & \text{if } Q = 2, \\ \emptyset & \text{if } Q = 3, \end{cases} \quad (12)$$

$$X = \begin{cases} U_1 & \text{if } Q = 1, \\ U_2 & \text{if } Q = 2, \\ U_0 & \text{if } Q = 3. \end{cases} \quad (13)$$

Denote the maximum symmetric rate achievable with the above auxiliary random variables identification as $R_{\text{sym-ts}}$. We now specialize Theorem 2 to establish the following simplified expression for this maximum symmetric rate.

Proposition 1: The maximum symmetric rate for the symmetric deterministic 2-receiver DM-BC with stale state using the time-sharing scheme is

$$R_{\text{sym-ts}} = \max_{p(x)} \frac{C_1 I(X; Y_1, Y_2 | S)}{2C_1 + I(X; Y_2 | Y_1, S)},$$

where $C_1 = \max_{p(x)} I(X; Y_1 | S)$.

Proof: Substituting (12) and (13) into (5), we obtain

$$R_{\text{sym-ts}} = \max_{\alpha, p(u_0), p(u_1)} \min \Big\{ \Big(\frac{1}{2} - \alpha\Big) I(U_0; Y_1 | S) + \alpha I(U_1; Y_1 | S) + \frac{\alpha}{2} I(U_1; Y_2 | Y_1, S), (1 - 2\alpha) I(U_0; Y_1 | S) + \alpha I(U_1; Y_1 | S) \Big\}. \quad (14)$$

We now find α and $p(u_0)$, $p(u_1)$ that achieve (14). Since $\alpha \le 0.5$, the coefficients in front of the $I(U_0; Y_1 | S)$ terms in (14) are nonnegative and without loss of optimality we can set $p(u_0) = \arg\max I(U_0; Y_1 | S)$. Then,

$$R_{\text{sym-ts}} = \max_{\alpha, p(x)} \min\{L(\alpha), R(\alpha)\}, \quad (15)$$

where

$$L(\alpha) = \Big(\frac{1}{2} - \alpha\Big) C_1 + \alpha I(X; Y_1 | S) + \frac{\alpha}{2} I(X; Y_2 | Y_1, S),$$
$$R(\alpha) = (1 - 2\alpha) C_1 + \alpha I(X; Y_1 | S).$$

To find α and $p(x)$ that maximize the minimum of the two terms, we first fix the pmf $p(x)$ and find α^* that maximizes the minimum of the two terms in terms of $p(x)$. We then optimize $R_{\text{sym-ts}}$ in $p(x)$.

For a fixed $p(x)$, both $L(\alpha)$ and $R(\alpha)$ are linear functions of α, and $L(0) \le R(0)$ and $L(0.5) \ge R(0.5)$. Thus, $\min\{L(\alpha), R(\alpha)\}$ attains its maximum value at α^* such that $L(\alpha^*) = R(\alpha^*)$, namely,

$$\alpha^* = \frac{C_1}{2C_1 + I(X; Y_2 | Y_1, S)}. \quad (16)$$

Replacing α^* in (15) by (16) completes the proof. ■

To be self contained, we give an outline of the coding scheme that achieves $R_{\text{sym-ts}}$. The time-sharing scheme uses block Markov coding to send $b-1$ independent message pairs $(M_{1j}, M_{2j}) \in [1 : 2^{nR_1}] \times [1 : 2^{nR_2}]$, $j \in [1 : b-1]$ in b n-transmission blocks.

Codebook Generation: We generate two codebooks, one for the refinement messages and the other for the new messages and the refinement messages.

Fix a pmf $p(q, u_0, u_1, u_2) = p(q)p(u_0)p(u_1)p(u_2)$, where $p_{U_0} = \arg\max_{p(x)} I(X; Y_1 | S)$ and $p_{U_1} = p_{U_2} = \arg\max_{p(x)} C_1 I(X; Y_1, Y_2 | S)/(2C_1 + I(X; Y_2 | Y_1, S))$ (i.e., the pmf that attains the upper bound in Proposition 1). Let $p_Q(1) = p_Q(2) = \alpha^*$ and $p_Q(3) = 1 - 2\alpha^*$, where $\alpha^* = C_1/(2C_1 + I(X; Y_2 | Y_1, S))$ evaluated with $p_X = p_{U_1}$. Let the functions $v_0(x, s, q)$ and $x(u_0, u_1, u_2, q)$ be defined as in (12) and in (13), respectively.

Randomly generate a time-sharing sequence $q^n \sim \prod_{i=1}^n p_Q(q_i)$. To generate the codebook for compression, randomly and independently generate $2^{n\tilde{R}'_0}$ sequences $v_0^n(l)$, $l \in [1 : 2^{n\tilde{R}'_0}]$, each according to $\prod_{i=1}^n p_{V_0|Q}(v_{0i}|q_i)$. Partition the sequences into $2^{nR'_0}$ equal size bins $\mathcal{B}$ indexed by $k \in [1 : 2^{nR'_0}]$.

To generate the codebook for transmission, randomly and independently generate $2^{nR'_0}$ sequences $u_0^n(k)$, $k \in [1 : 2^{nR'_0}]$, each according to $\prod_{i=1}^n p(u_{0i})$.

TABLE I

TIME-SHARING SCHEME

Block	1	2	$\cdots$	j	$\cdots$
X		$v_0^n(l_1)$	$\cdots$	$v_0^n(l_{j-1})$	$\cdots$
	$k_0=1$	k_1	$\cdots$	k_{j-1}	$\cdots$
	$u_0^n(k_0), u_1^n(m_{11}), u_2^n(m_{21})$	$u_0^n(k_1), u_1^n(m_{12}), u_2^n(m_{22})$	$\cdots$	$u_0^n(k_{j-1}), u_1^n(m_{1j}), u_2^n(m_{2j})$	$\cdots$
	$x^n(u_0^n, u_1^n, u_2^n, q^n)$	$x^n(u_0^n, u_1^n, u_2^n, q^n)$	$\cdots$	$x^n(u_0^n, u_1^n, u_2^n, q^n)$	$\cdots$
Y_1	$\hat{m}_{11}$	$\leftarrow (\hat{l}_1, \hat{k}_1), \hat{m}_{12}$	$\cdots$	$\leftarrow (\hat{l}_{j-1}, \hat{k}_{j-1}), \hat{m}_{1j}$	$\cdots$
Y_2	$\hat{m}_{21}$	$\leftarrow (\hat{l}_1, \hat{k}_1), \hat{m}_{22}$	$\cdots$	$\leftarrow (\hat{l}_{j-1}, \hat{k}_{j-1}), \hat{m}_{2j}$	$\cdots$

TABLE II

THE MAT 2-RECEIVERS SCHEME

Sub-block	1	2	3
X			$v_0^n(l)=(y_2^{n_1}, y_{1,n_1+1}^{n_1+n_2}, \emptyset_{n_1+n_2+1}^n)$
			k
	$u_1^n(m_1)$	$u_2^n(m_2)$	$u_0^n(k)$
	$x^{n_1}=u_1^{n_1}$	$x_{n_1+1}^{n_1+n_2}=u_{2,n_1+1}^{n_1+n_2}$	$x_{n_1+n_2+1}^n=u_{0,n_1+n_2+1}^n$
Y_1			$\hat{l}, \hat{k}, \hat{m}_1$
Y_2			$\hat{l}, \hat{k}, \hat{m}_2$

Similarly, generate 2^{nR_1} sequences $u_1^n(m_1)$, $m_1 \in [1:2^{nR_1}]$, each according to $\prod_{i=1}^n p_U(u_{1i})$, and 2^{nR_2} sequences $u_2^n(m_2)$, $m_2 \in [1:2^{nR_2}]$, each according to $\prod_{i=1}^n p_U(u_{2i})$. Finally generate $x_i(k, m_1, m_2) = x(u_{0i}(k), u_{1i}(m_1), u_{2i}(m_2), q_i)$, $i \in [1:n]$.

Encoding and Decoding: Encoding and decoding are described with the help of Table I. The encoder transmits $x^n(1, m_{11}, m_{21})$ in block 1. In block $j \in [2:b]$, the refinement message k_{j-1} is the bin index such that $v_0^n(x_{j-1}^n, s_{j-1}^n, q^n) \in \mathcal{B}(k_{j-1})$, where $v_{0i} = v_0(x_i, s_i, q_i)$, $i \in [1:n]$. If no such sequence exists, choose $k_{j-1} = 1$. The encoder transmits $x^n(k_{j-1}, m_{1j}, m_{2j})$.

Backward decoding is used at each decoder (we set $(m_{1b}, m_{2b}) = (1,1)$).

1) Decoder 1 declares $\hat{k}_{b-1}$ is sent if it is the unique message such that $(u_0^n(\hat{k}_{b-1}), y_{1b}^n, s_b^n, q^n) \in \mathcal{T}_\epsilon^{(n)}$. The probability of error in recovering k_{b-1} tends to zero as $n \to \infty$ if $R_0' < (1-2\alpha^*)C_1$.
2) Decoder 1 finds a sequence $v_0^n(\hat{l}_{b-1}) \in \mathcal{B}(\hat{k}_{b-1})$ such that $(v_0^n(\hat{l}_{b-1}), y_{1,b-1}^n, s_{b-1}^n, q^n) \in \mathcal{T}_\epsilon^{(n)}$.
3) Decoder 1 declares $\hat{m}_{1,b-1}$ is sent if it is the unique message such that $(u_1^n(\hat{m}_{1,b-1}), v_0^n(\hat{l}_{b-1}), y_{1,b-1}^n, s_{b-1}^n, q^n) \in \mathcal{T}_\epsilon^{(n)}$. The probability of error for this step tends to zero as $n \to \infty$ if $R_0' > \alpha^* I(X; Y_1|Y_2, S)$ and $R_1 < \alpha^* I(X; Y_1, Y_2|S)$.
4) Decoder 1 repeats the same procedure until it recovers all $b-1$ messages, $M_{1j},\ j \in [1:b-1]$.

Decoder 2 performs decoding similarly.

Remark 2: Although the SW scheme, which achieves the maximum symmetric rate in (14), uses block Markov coding, coded time sharing, and backward decoding, it is not difficult to see that this maximum symmetric rate can also be achieved using the MAT scheme as illustrated in Table II. Instead of using coded time sharing, the MAT scheme for 2-receiver case [5, Sec. III] uses time-sharing over three sub-blocks such that

$$V_1 = V_2 = V_0 = \begin{cases} Y_2 & \text{in sub-block 1,} \\ Y_1 & \text{in sub-block 2,} \\ \emptyset & \text{in sub-block 3,} \end{cases}$$

$$X = \begin{cases} U_1 & \text{in sub-block 1,} \\ U_2 & \text{in sub-block 2,} \\ U_0 & \text{in sub-block 3.} \end{cases}$$

Recall that in the SW scheme with coded time sharing, the encoder needs to wait until the end of each block to generate the refinement message k (since it needs the entire state sequence s^n). The message k is then sent it in the following block. In the time-sharing scheme over three sub-blocks, the refinement message k is the bin index such that $v_0^n(x^{n_1+n_2}, s^{n_1+n_2}) \in \mathcal{B}(k)$, where n_1 and n_2 denote the lengths of sub-blocks 1 and 2, respectively. The state sequence $s^{n_1+n_2}$ is available to the encoder at the end of sub-block 2. Hence, the encoder can generate k at the end of sunblock 2 and send it in sub-block 3.

Remark 3: The MAT scheme for the 2-receiver case can be extended to any number of receivers [5]. Although the SW scheme is only for the 2-receiver case, it can be readily extended to more receivers using the sub-blocks and auxiliary random variables as we described for the 2-receiver case.

We now apply the time-sharing scheme to the two examples in [5]. Note that these two examples satisfy the additional condition that $\arg\max_{p(x)} I(X; Y_1|S) = \arg\max_{p(x)} I(X; Y_1, Y_2|S)$, hence

$$R_{\text{sym-ts}} = \frac{C_1 C_{12}}{C_1 + C_{12}}, \tag{17}$$

where

$$C_1 = \max_{p(x)} I(X; Y_1|S) \text{ and } C_{12} = \max_{p(x)} I(X; Y_1, Y_2|S),$$

and is achievable with U_0, U_1, and U_2 each distributed according to $\arg\max_{p(x)} I(X; Y_1|S)$, and $p_Q(1) = p_Q(2) = C_1/(C_1+C_{12})$, $p_Q(3) = 1-2C_1/(C_1+C_{12})$ in (12) and (13).

Example 1 (Broadcast Erasure Channel [5], [7]): Consider a DM-BC with random state with $X \in \{0,1\}$, $p(y_1, y_2|x) = p(y_1|x)p(y_2|x)$, where $Y_i = X$ with probability $1-\epsilon$ and $Y_i = \mathrm{e}$ with probability ϵ for $i = 1,2$, and $S = (S_1, S_2)$, where $S_i = 0$ if $Y_i = X$ and $S_i = 1$ if $Y_i = \mathrm{e}$ for $i = 1,2$.

To evaluate the maximum symmetric rate in (17), note that $C_1 = 1 - \epsilon$, $C_{12} = 1 - \epsilon^2$. Then,

$$R_{\text{sym-ts}} = \frac{1-\epsilon^2}{2+\epsilon}. \tag{18}$$

In [8], an outer bound on the capacity region for this example was obtained based on the observation that this capacity region cannot be larger than that of the physically degraded broadcast channel with input X, outputs Y_1 and (Y_1, Y_2), and with causal feedback. Using the same technique, it was shown in [5] that the bound on the symmetric capacity coincides with (18).

Example 2 (Finite Field Deterministic Channel [5]): Consider the DM-BC

$$\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = HX,$$

where

$$H = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix}, \quad X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}, \quad \text{and } S = H.$$

Assume that H is chosen uniformly at random from the set of full-rank matrices over a finite field. Further assume that $|\mathcal{Y}_1| = |\mathcal{Y}_2| = |\mathcal{Y}|$.

To evaluate the maximum symmetric rate in (17), note that $C_1 = \log|\mathcal{Y}|$, $C_{12} = 2\log|\mathcal{Y}|$. Then,

$$R_{\text{sym-ts}} = \frac{2\log|\mathcal{Y}|}{3}. \tag{19}$$

Using the same converse technique as for Example 1, it was shown in [5] that (19) is the symmetric capacity of this channel.

Note that in the above two examples, the channel is deterministic for each state. Is the time-sharing scheme then optimal for all such channels? The example in the following section shows that time-sharing scheme is not in general optimal for this class of channels.

V. Superposition Coding Scheme

In the time-sharing scheme, we separately transmit new message and their refinement. In this section, we consider another special case of the scheme in Corollary 1 in which we also use superposition coding. Specifically, we specialize the auxiliary random variables in Theorem 2 as follows. Let $Q \in \{1, 2\}$ and $P_Q(1) = P_Q(2) = 0.5$. Let $p(q, u_0, u_1, u_2) = p(q)p(u_0)p(u_1|u_0)p(u_2|u_0)$ and $p_{U_1|U_0}(u|u_0) = p_{U_2|U_0}(u|u_0)$. Define

$$V_1 = V_2 = V_0 = \begin{cases} Y_2 & \text{if } Q = 1, \\ Y_1 & \text{if } Q = 2, \end{cases} \tag{20}$$

$$X = \begin{cases} U_1 & \text{if } Q = 1, \\ U_2 & \text{if } Q = 2. \end{cases} \tag{21}$$

Denote the maximum symmetric rate achievable with the above auxiliary random variables identification by $R_{\text{sym-sp}}$. We now specialize Theorem 2 to establish the following simplified expression for this maximum symmetric rate.

Proposition 2: The maximum achievable symmetric rate for the symmetric deterministic 2-receiver DM-BC with stale state using the superposition coding scheme is

$$R_{\text{sym-sp}} = \max_{p(u_0,x)} \min \Big\{ 0.5I(X;Y_1|S) + 0.5I(U_0;Y_1|S), \; 0.5I(X;Y_1|S) + 0.25I(X;Y_2|Y_1,U_0,S) \Big\}.$$

Proof: This proposition is obtained by substituting (20) and (21) into (5). ■

To be self contained, we give an outline of the coding scheme that achieves $R_{\text{sym-sp}}$. The superposition coding scheme again uses a block Markov coding in which $b-1$ independent message pairs $(M_{1j}, M_{2j}) \in [1 : 2^{nR_1}] \times [1 : 2^{nR_2}]$, $j \in [1 : b-1]$, are sent in b n-transmission blocks.

Codebook Generation: We generate two codebooks, one for the refinement messages and the other for the new messages and the refinement messages. Fix a pmf $p(q, u_0, u_1, u_2) = p(q)p(u_0)p(u_1|u_0)p(u_2|u_0)$, where $p_{U_0,U_1} = p_{U_0,U_2} = p_{U,X}$ that attains the upper bound in Proposition 2. Let $Q \in \{1, 2\}$ and $p_Q(1) = p_Q(2) = 0.5$. Let the functions $v_0(x, s, q)$ and $x(u_1, u_2, q)$ be defined as in (20) and in (21), respectively.

Randomly generate a time-sharing sequence q^n according to $\prod_{i=1}^n p_Q(q_i)$. To generate the codebook for compression, randomly and independently generate $2^{n\tilde{R}_0'}$ sequences $v_0^n(l)$, $l \in [1 : 2^{n\tilde{R}_0'}]$, each according to $\prod_{i=1}^n p_{V|Q}(v_i|q_i)$. Partition the sequences into $2^{nR_0'}$ equal size bins $\mathcal{B}$ indexed by $k \in [1 : 2^{nR_0'}]$.

To generate the codebook for transmission, rate splitting is used. Divide M_1 into two independent messages M_{1c} at rate R_{1c} and M_{1p} at rate R_{1p}. Similarly, divide M_2 into M_{2c} at rate R_{2c} and M_{2p} at rate R_{2p}. Hence $R_1 = R_{1c} + R_{1p}$ and $R_2 = R_{2c} + R_{2p}$. Randomly and independently generate $2^{n(R_0'+R_{1c}+R_{2c})}$ sequences $u_0^n(k, m_{1c}, m_{2c})$, each according to $\prod_{i=1}^n p_{U_0}(u_{0i})$. For each (k, m_{1c}, m_{2c}), randomly and conditionally independently generate $2^{nR_{1p}}$ sequences $u_1^n(k, m_{1c}, m_{2c}, m_{1p})$, $m_{1p} \in [1 : 2^{nR_{1p}}]$, each according to $\prod_{i=1}^n p_{U|U_0}(u_{1i}|u_{0i}(k, m_{1c}, m_{2c}))$. Similarly, randomly and conditionally independently generate $2^{nR_{2p}}$ sequences $u_2^n(k, m_{1c}, m_{2c}, m_{2p})$, $m_{2p} \in [1 : 2^{nR_{2p}}]$, each according to $\prod_{i=1}^n p_{U|U_0}(u_{2i}|u_{0i}(k, m_{1c}, m_{2c}))$. Finally generate $x_i(k, m_1, m_2) = x(u_{1i}(k, m_{1c}, m_{2c}, m_{1p}), u_{2i}(k, m_{1c}, m_{2c}, m_{2p}), q_i)$, $i \in [1 : n]$.

Encoding and Decoding: Encoding and decoding are described with the help of Table III. The encoder transmits $x^n(1, m_{11}, m_{21})$ in block 1. In block $j \in [2 : b]$, the refinement message k_{j-1} is the bin index such that $v_0^n(x_{j-1}^n, s_{j-1}^n, q^n) \in \mathcal{B}(k_{j-1})$, where $v_{0i} = v_0(x_i, s_i, q_i)$, $i \in [1 : n]$. If no such sequence exists, choose $k_{j-1} = 1$. The encoder transmits $x^n(k_{j-1}, m_{1j}, m_{2j})$.

Backward decoding is used at each decoder (we set $(m_{1b}, m_{2b}) = (1, 1)$).

1) Decoder 1 recovers $\hat{k}_{b-1}$ from block b (block b operation is treated differently as detailed in [2]).

TABLE III
SUPERPOSITION CODING SCHEME

Block	1	2	$\cdots$	j	$\cdots$
		$v_0^n(l_1)$	$\cdots$	$v_0^n(l_{j-1})$	$\cdots$
X	$k_0 = 1$	k_1	$\cdots$	k_{j-1}	$\cdots$
	$u_0^n(k_0, m_{1c,1}, m_{2c,1})$	$u_0^n(k_1, m_{1c,2}, m_{2c,2})$	$\cdots$	$u_0^n(k_{j-1}, m_{1c,j}, m_{2c,j})$	$\cdots$
	$u_1^n(k_0, m_{1c,1}, m_{2c,1}, m_{1p,1})$	$u_1^n(k_1, m_{1c,2}, m_{2c,2}, m_{1p,2})$	$\cdots$	$u_1^n(k_{j-1}, m_{1c,j}, m_{2c,j}, m_{1p,j})$	$\cdots$
	$u_2^n(k_0, m_{1c,1}, m_{2c,1}, m_{2p,1})$	$u_2^n(k_1, m_{1c,2}, m_{2c,2}, m_{2p,2})$	$\cdots$	$u_2^n(k_{j-1}, m_{1c,j}, m_{2c,j}, m_{2p,j})$	$\cdots$
	$x^n(u_1^n, u_2^n, q^n)$	$x^n(u_1^n, u_2^n, q^n)$	$\cdots$	$x^n(u_1^n, u_2^n, q^n)$	$\cdots$
Y_1	$\hat{m}_{11}$	$\leftarrow (\hat{l}_1, \hat{k}_1), \hat{m}_{12}$	$\cdots$	$\leftarrow (\hat{l}_{j-1}, \hat{k}_{j-1}), \hat{m}_{1j}$	$\cdots$
Y_2	$\hat{m}_{21}$	$\leftarrow (\hat{l}_1, \hat{k}_1), \hat{m}_{22}$	$\cdots$	$\leftarrow (\hat{l}_{j-1}, \hat{k}_{j-1}), \hat{m}_{2j}$	$\cdots$

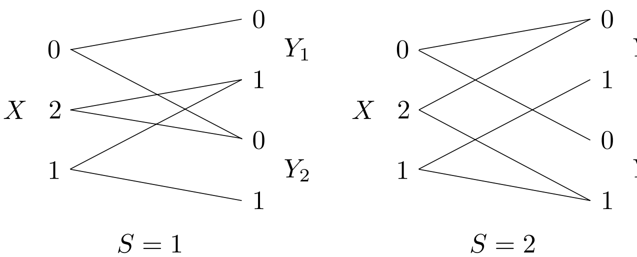

Fig. 2. Blackwell channel with state.

2) Decoder 1 finds a sequence $v_0^n(\hat{l}_{b-1}) \in \mathcal{B}(\hat{k}_{b-1})$ such that $(v_0^n(\hat{l}_{b-1}), y_{1,b-1}^n, s_{b-1}^n, q^n) \in \mathcal{T}_\epsilon^{(n)}$.
3) Decoder 1 declares that $\hat{m}_{1,b-1} = (\hat{m}_{1c,b-1}, \hat{m}_{1p,b-1})$ and $\hat{k}_{b-2}$ are sent if they are the unique message pair such that $(u_0^n(\hat{k}_{b-2}, \hat{m}_{1c,b-1}, \hat{m}_{2c,b-1}), u_1^n(\hat{k}_{b-2}, \hat{m}_{1c,b-1}, \hat{m}_{2c,b-1}, \hat{m}_{1p,b-1}), v_0^n(\hat{l}_{b-1}), y_{1,b-1}^n, s_{b-1}^n, q^n) \in \mathcal{T}_\epsilon^{(n)}$ for some $\hat{m}_{2c,b-1}$. The probability of error for this step tends to zero as $n \to \infty$ if $R_0' > 0.5I(X; Y_2|Y_1, S)$ and

$$\begin{aligned} R_0' + R_1 &< 0.5I(U_0; Y_1|S) + 0.5I(X; Y_1, Y_2|S), \\ R_0' + R_2 &< 0.5I(U_0; Y_2|S) + 0.5I(X; Y_1, Y_2|S), \\ R_0' + R_1 + R_2 &< 0.5I(U_0; Y_1|S) + 0.5I(X; Y_1, Y_2|S) \\ &\quad + 0.5I(X; Y_1, Y_2|U_0, S), \\ R_0' + R_1 + R_2 &< 0.5I(U_0; Y_2|S) + 0.5I(X; Y_1, Y_2|S) \\ &\quad + 0.5I(X; Y_1, Y_2|U_0, S), \\ 2R_0' + R_1 + R_2 &< 0.5I(U_0; Y_1|S) + 0.5I(U_0; Y_2|S) \\ &\quad + I(X; Y_1, Y_2|S). \end{aligned}$$

4) Decoder 1 repeats the same process until decoder 1 recovers all $b-1$ messages, $M_{1j},\ j \in [1 : b-1]$.

Decoder 2 performs decoding similarly.

We now introduce a new example of a symmetric deterministic broadcast channel with stale state for which the superposition coding scheme outperforms the time-sharing scheme.

Example 1 (Blackwell Channel With State): Consider the symmetric DM-BC with random state depicted in Fig. 2, where $p_S(1) = p_S(2) = 0.5$.

We first evaluate $R_{\text{sym-ts}}$. Let $U_0 \sim \text{Bern}(0.5)$ and U_1 and U_2 be independently and identically distributed according to $p_{U_1}(0) = p_0,\ p_{U_1}(2) = p_2,\ p_{U_1}(1) = 1 - p_0 - p_2$. We numerically maximize the expression for the maximum symmetric rate in Proposition 1 in (p_0, p_2) to obtain

$$\begin{aligned} R_{\text{sym-ts}} &= \max_{p_0, p_2} \frac{H(p_0, 1-p_0-p_2, p_2)}{2 + 0.5\bar{p}_0 H_b(p_2/\bar{p}_0) + 0.5\bar{p}_2 H_b(p_0/\bar{p}_2)} \\ &= 0.5989 \end{aligned}$$

for $p_0^* = p_2^* = 0.37325$. Here, $\bar{p}_0 = 1 - p_0$, $\bar{p}_2 = 1 - p_2$, and $H(p_0, 1-p_0-p_2, p_2)$ is the entropy of U_1.

We now show that superposition coding can achieve a symmetric rate greater than $R_{\text{sym-ts}}$. Let $U_0 \in \{0, 1, 2, 3\}$ and $p_{U_0}(0) = p_{U_0}(1) = q_1, p_{U_0}(2) = p_{U_0}(3) = (1-2q_1)/2$, $0 \le q_1 \le 0.5$. Let $p_{U_1|U_0}(u|u_0) = p_{U_2|U_0}(u|u_0)$ and

$$p_{U_1|U_0}(u|u_0) = \begin{cases} \alpha_1\bar{\beta}_1 & \text{if } (u, u_0) = (0, 0) \text{ or } (1, 1), \\ \beta_1 & \text{if } (u, u_0) = (2, 0) \text{ or } (2, 1), \\ \bar{\alpha}_1\bar{\beta}_1 & \text{if } (u, u_0) = (1, 0) \text{ or } (0, 1), \\ \alpha_2\bar{\beta}_2 & \text{if } (u, u_0) = (0, 2) \text{ or } (1, 3), \\ \beta_2 & \text{if } (u, u_0) = (2, 2) \text{ or } (2, 3), \\ \bar{\alpha}_2\bar{\beta}_2 & \text{if } (u, u_0) = (1, 2) \text{ or } (0, 3). \end{cases}$$

Choose (V_0, V_1, V_2) and X as in (20)-(21).

Maximizing the symmetric rate in Proposition 2 over q_1 and $0 \le \alpha_1, \alpha_2, \beta_1, \beta_2 \le 1$, we obtain $R_{\text{sym-sp}} \ge 0.6103$ at $q_1^* = 0.5$, $\alpha_1^* = 0.13628$ and $\beta_1^* = 0.23025$, which is greater than the symmetric rate achieved using the time-sharing scheme.

To investigate the optimality of the achievable symmetric rate using superposition coding, we consider the same physically degraded broadcast channel with state in [5] with input X and outputs (Y_1, Y_2) and Y_2. The capacity region of this channel is the set of rate pairs (R_1, R_2) such that

$$\begin{aligned} R_1 &\le I(X; Y_1, Y_2|S, U), \\ R_2 &\le I(U; Y_2|S), \end{aligned}$$

where $|\mathcal{U}| \le \min\{|\mathcal{X}|, |\mathcal{Y}_1||\mathcal{S}|, |\mathcal{Y}_2||\mathcal{S}|\} + 1$.

Hence, the symmetric capacity is upper bounded as

$$C_{\text{sym}} \le \max_{\substack{p(u)p(x|u), \\ |\mathcal{U}| \le 4}} \min\{I(U; Y_2|S), I(X; Y_1, Y_2|S, U)\}. \quad (22)$$

We can show that the upper bound is greater than 0.653 numerically using the substitutions: $U \sim \text{Bern}(0.5)$, $p_{X|U}(0|0) = p_{X|U}(1|1) = 0.832$, $p_{X|U}(2|0) = p_{X|U}(2|1) = 0.168$, $p_{X|U}(1|0) = p_{X|U}(0|1) = 0$. Thus, the upper bound in (22) is greater than the inner bound using the superposition coding scheme of $R_{\text{sym-sp}} \geq 0.6103$.

We now show that this upper bound in (22) is less than or equal to $2/3$ but equality cannot hold, which implies that the upper bound is strictly less than $2/3$. Consider

$$\begin{aligned}
&\max_{p(u,x)} \min\{I(U;Y_2|S), I(X;Y_1,Y_2|S,U)\} \\
&\stackrel{(a)}{=} \max_{p(u,x)} \min\{H(Y_2|S)-H(Y_2|S,U), H(Y_1,Y_2|S,U)\} \\
&\stackrel{(b)}{\leq} \max_{p(u,x)} \min\{H(Y_2|S)-H(Y_2|S,U), 2H(Y_2|S,U)\} \\
&\stackrel{(c)}{\leq} \max_{p(x)} \frac{2H(Y_2|S)}{3} \\
&\stackrel{(d)}{\leq} \frac{2}{3},
\end{aligned}$$

where (a) holds because the Blackwell channel with state is deterministic and (b) holds since $H(Y_1, Y_2|S, U) \leq H(Y_1|S, U) + H(Y_2|S, U) = 2H(Y_2|S, U)$. Step (c) can be shown as follows. Suppose $H(Y_2|S) - H(Y_2|S, U) > 2H(Y_2|S)/3$, then $2H(Y_2|S, U) < 2H(Y_2|S)/3$. Therefore at least one of the two terms is less than or equal to $2/3H(Y_2|S)$. Step (d) holds since $|\mathcal{Y}_2| = 2$, and equality holds iff $Y_2 \sim \text{Bern}(0.5)$. Now suppose equality holds for (b), (c), and (d), and then, from equality for (d), $Y_2 \sim \text{Bern}(0.5)$, which implies that $X = Y_1 = Y_2 \sim \text{Bern}(0.5)$. Then, from the equality for (c), $H(X|S, U) = 1/3$ and from equality for (b), $H(X|S, U) = 2H(X|S, U) = 0$, which is a contradiction. Thus, equality cannot hold for (b), (c), and (d). We conclude that $C_{\text{sym}} < 2/3$.

VI. Conclusion

We derived a simplified expression for the maximum symmetric rate achievable using the Shayevitz–Wigger scheme for the symmetric broadcast channel with random state when the state is known at the receivers and only strictly causally at the transmitter. We considered a time-sharing special case of the SW scheme for symmetric deterministic broadcast channels and showed that it attains the symmetric capacity of the Maddah-Ali–Tse examples. We then introduced the Blackwell channel with state example and showed that a superposition coding special case of the SW scheme can achieve a higher symmetric rate than the time-sharing scheme.

There are several open questions that would be interesting to explore further, including the following.

- We showed that the time-sharing scheme is not optimal for the class of deterministic channels as defined in Section II. For what general class of channels is it optimal?
- Is the symmetric rate achieved using the superposition coding scheme for the Blackwell channel with state example optimal? Can a higher symmetric rate be achieved using Marton coding?
- For what general class of channels is the symmetric rate achieved using the SW scheme optimal?

References

[1] A. El Gamal and Y. H. Kim, *Network Information Theory*, 1st ed. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[2] 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.

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

[4] R. M. Gray and A. D. Wyner, "Source coding for a simple network," *Bell Syst. Tech. J.*, vol. 53, no. 9, pp. 1681–1721, Nov. 1974.

[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, Jul. 2012.

[6] D. Blackwell, L. Breiman, and A. J. Thomasian, "Proof of Shannon's transmission theorem for finite-state indecomposable channels," *Ann. Math. Statist.*, vol. 29, no. 4, pp. 1209–1220, Dec. 1958.

[7] L. Georgiadis and L. Tassiulas, "Broadcast erasure channel with feedback—Capacity and algorithms," in *Proc. Workshop Netw. Coding, Theory, Appl.*, Lausanne, Switzerland, Jun. 2009, pp. 54–61.

[8] P. Larsson and N. Johansson, "Multi-user ARQ," in *Proc. IEEE 63rd Veh. Technol. Conf. (VTC)*, vol. 4. May 2006, pp. 2052–2057.

Hyeji Kim (S'14) is a Ph.D. candidate in the Department of Electrical Engineering at Stanford University. She received the B.S. degree with honors in Electrical Engineering from the Korea Advanced Institute of Science and Technology (KAIST) in 2011 and the M.S. degree in Electrical Engineering from Stanford University in 2013. She is a recipient of the Stanford Graduate Fellowship (SGF). Her research interest include information theory, communication systems, and statistical learning.

Yeow-Khiang Chia is currently a research scientist at the Institute for Infocomm Research, Singapore. He received his M.Eng (with first class honours) in Electrical Engineering from Imperial College, London. He received his M.S. and PhD, both in Electrical Engineering, from Stanford University in 2012. He was a recipient of the Stanford Graduate Fellowship (SGF) and the Agency for Science, Technology and Research (A*STAR), Singapore National Science Scholarship (NSS) for his graduate studies at Stanford University. His research interests include information theory, optimization, machine learning and signal processing.

Abbas El Gamal (S'71–M'73–SM'83–F'00) is the Hitachi America Professor in the School of Engineering and the Chair of the Department of Electrical Engineering at Stanford University. He received his B.Sc. Honors degree from Cairo University in 1972, and his M.S. in Statistics and Ph.D. in Electrical Engineering both from Stanford University in 1977 and 1978, respectively. From 1978 to 1980, he was an Assistant Professor of Electrical Engineering at USC. From 2004 to 2009, he was the Director of the Information Systems Laboratory at Stanford University. His research contributions have been in network information theory, FPGAs, and digital imaging devices and systems. He has authored or coauthored over 200 papers and holds 35 patents in these areas. He is coauthor of the book *Network Information Theory* (Cambridge Press 2011). He received several honors and awards for his research contributions, including the 2012 Claude E. Shannon Award, the 2009 Padovani Lecture, and the 2004 INFOCOM Paper Award. He is a member of the US National Academy of Engineering and a Fellow of the IEEE. He has been serving on the Board of Governors of the Information Theory Society since 2009 and is currently its Junior Past President. He has played key roles in several semiconductor, EDA, and biotechnology startup companies.