Mixed Joint Source-Channel Coding Schemes for the Multiple-Access Relay Channel: Error Probability Analysis

Yonathan Murin†, Ron Dabora†, Deniz Gündüz*
†Dept. of Electrical and Computer Engineering, Ben-Gurion University, Israel
*Centre Tecnologic de Telecomunicacions de Catalunya (CTTC), Barcelona, Spain
Email: moriny@ee.bgu.ac.il, ron@ee.bgu.ac.il, deniz.gunduz@cttc.es

I. NOTATIONS AND MODEL

In the following we denote random variables with upper case letters, e.g. $X$, $Y$, and their realizations with lower case letters, e.g. $x$, $y$. A discrete random variable $X$ takes values in a set $\mathcal{X}$. We use $p_X(x)$ to denote the probability mass function (p.m.f.) of a discrete RV $X$ on $\mathcal{X}$; for brevity we may omit the subscript $X$ when it is the uppercase version of the sample symbol $x$. We denote vectors with boldface letters, e.g. $x$, $y$; the $i$'th element of a vector $x$ is denoted by $x_i$ and we use $x_{i:j}$ where $i < j$ to denote $(x_i, x_{i+1}, \ldots, x_{j-1}, x_j)$; $x^j$ is a short form notation for $x_1^j$, and unless specified otherwise $x \triangleq x^n$. We use $A^{(n)}_x(X)$ to denote the set of $\epsilon$-strongly typical sequences w.r.t. distribution $p_X(x)$ on $\mathcal{X}$, as defined in [2, ch. 13.6]. When referring to a typical set we may omit the random variables from the notation, when these variables are clear from the context. We denote the empty set with $\emptyset$, and the complement of the set $B$ by $B^c$.

The MARC consists of two transmitters (sources), a receiver (destination) and a relay. Transmitter $i$ has access to the source sequence $S_i^m$, for $i = 1, 2$. The receiver is interested in the lossless reconstruction of the source sequences observed by the two transmitters. The objective of the relay is to help the receiver decode the source sequences. The relay and the receiver each has its own side information, denoted by $W_3^m$ and $W^m$, respectively, correlated with the source sequences. In the multiple-access broadcast relay channel (MABRC) both the relay and the destination are interested a lossless reconstruction of the sources. Figure 1 depicts the MARC with side information setup.

Fig. 1. Multiple-access relay channel with correlated side information. $(\hat{S}_1^m, \hat{S}_2^m)$ are the reconstructions at the destination.
The sources and the side information sequences, \( \{S_1,k, S_2,k, W_1,k, W_2,k\}_{k=1}^m \), are arbitrarily correlated according to a joint distribution \( p(s_1,k, s_2,k, w_1,k, w_2,k) \) over a finite alphabet \( S_1 \times S_2 \times W_1 \times W_2 \), and independent across different sample indices \( k \). All nodes know this joint distribution.

For transmission, a discrete memoryless MARC with inputs \( X_1, X_2, X_3 \) over finite input alphabets \( X_1, X_2, X_3 \), and outputs \( Y, Y_3 \) over finite output alphabets \( Y, Y_3 \), is available. The MARC is memoryless in the sense

\[
p(y_k, y_3,k | y_{k-1}, y_3,k-1, x_1,k, x_2,k, x_3,k) = p(y_k, y_3,k | x_1,k, x_2,k, x_3,k).
\]

An \((m,n)\) source-channel code for the MARC with correlated side information consists of two encoding functions at the transmitters: \( f_i^{(m,n)} : S_i^m \mapsto X_i^n, i = 1, 2 \), a decoding function at the destination, \( g^{(m,n)} : Y^m \times W^m \mapsto S_1^m \times S_2^m \), and a set of causal encoding functions at the relay, \( x_3,k = f_3^{(m,n)}(y_3,k-1, w_3^m, 1 \leq k \leq n) \). Let \( \hat{S}_i^m \) denote the reconstruction of \( S_i^m, i = 1, 2 \), at the receiver. The average probability of error of an \((m,n)\) code for the MARC is defined as

\[
P_e^{(m,n)} = \Pr((\hat{S}_1^m, \hat{S}_2^m) \neq (S_1^m, S_2^m)).
\]

A source-channel rate \( \kappa \) is said to be achievable for the MARC, if for every \( \epsilon > 0 \) there exist positive integers \( n_0, m_0 \) such that for all \( n > n_0, m > m_0, n/m = \kappa \), there exists an \((m,n)\) code for which \( P_e^{(m,n)} < \epsilon \).

II. Mixed Joint Source-Channel Coding Scheme for Discrete Memoryless MARCs and MABRCs

In this section we present two sets of sufficient conditions for the achievability of source-channel rate \( \kappa = 1 \) for DM MARCs with correlated sources and side information. Both achievability schemes are DF-based, use successive decoding at the relay, and use CPM for sending information to both the relay and the destination. The cooperation in both schemes is based on SW source coding. By using cooperation based on SW source coding (while CPM is used for transmitting to the destination), the constraints on the distribution chain obtained by the scheme of [1, Thm. 8] are removed. The two achievability schemes differ in the decoding method used at the destination. In the first scheme (Thm. 1) successive backward decoding of the cooperation information and the transmitted source sequences is implemented at the destination; while in the second scheme (Thm. 2) simultaneous decoding of the cooperation information and the transmitted sources sequences is implemented at the destination.

The transmission techniques (SW source coding or CPM) of the schemes implemented in [1, Thm. 7], [1, Thm. 8] and in the proposed schemes are depicted in Figure 2. It can be observed that in both Figure 2(a) and Figure 2(b) the relay and the sources use the same transmission techniques for transmitting to the destination. However, this is not the case in Figure 2(c).

A. Successive Backward Decoding

Theorem 1. For DM MARCs and MABRCs with relay and receiver side information as defined in Section I, source-channel rate \( \kappa = 1 \) is achievable if,
Fig. 2. Transmission techniques of the schemes implemented in [1, Thm. 7] (a), [1, Thm. 8] (b), and in the new proposed schemes (c). Blue solid arrows denote transmission based on SW source coding. Red dashed arrows denote CPM-based transmission.

\[
H(S_1|S_2, W_3) < I(X_1; Y_3|S_2, X_2, V_1, X_3, W_3) \tag{1a}
\]
\[
H(S_2|S_1, W_3) < I(X_2; Y_3|S_1, X_1, V_2, X_3, W_3) \tag{1b}
\]
\[
H(S_1, S_2|W_3) < I(X_1, X_2; Y_3|V_1, V_2, X_3, W_3) \tag{1c}
\]
\[
H(S_1|S_2, W) < I(X_1; Y|S_2, W, X_2, V_1, X_3, W) + I(V_1, X_3; Y|V_2, W) \tag{1d}
\]
\[
H(S_2|S_1, W) < I(X_2; Y|S_1, X_1, V_2, X_3, W) + I(V_2, X_3; Y|V_1, W) \tag{1e}
\]
\[
H(S_1, S_2|W) < I(X_1, X_2; Y|V_1, V_2, X_3, W) + I(V_1, V_2, X_3; Y|W), \tag{1f}
\]

for a joint distribution that factors as

\[
p(s_1, s_2, w_3, w)p(v_1)p(x_1|s_1, v_1) \times
p(v_2)p(x_2|s_2, v_2)p(x_3|v_1, v_2)p(y_3, y|x_1, x_2, x_3). \tag{2}
\]

\textbf{Proof:} A proof outline is given in Subsection II-C. \hfill \square

\textbf{B. Simultaneous Backward Decoding}

\textit{Theorem} 2. For DM MARCs and MABRCs with relay and receiver side information as defined in Section I, source-channel rate \( \kappa = 1 \) is achievable if,
\[ H(S_1|S_2,W_3) < I(X_1;Y_3|S_2,X_2,V_1,X_3,W_3) \]  
\[ H(S_2|S_1,W_3) < I(X_2;Y_3|S_1,X_1,V_2,X_3,W_3) \]  
\[ H(S_1,S_2|W_3) < I(X_1,X_2;Y_3|V_1,V_2,X_3,W_3) \]  
\[ H(S_1|S_2,W) < \min \left\{ I(X_1,X_3;Y|S_2,W,X_2,V_2), \right. \]
\[ \left. \quad I(X_1;Y|S_2,X_2,V_1,X_3,W) + I(X_1,X_3;Y|S_1,X_2,V_2) \right\} \]  
\[ H(S_2|S_1,W) < \min \left\{ I(X_2,X_3;Y|S_1,W,X_1,V_1), \right. \]
\[ \left. \quad I(X_2;Y|S_1,X_1,V_2,X_3,W) + I(X_2,X_3;Y|S_2,X_1,V_1) \right\} \]  
\[ H(S_1,S_2|W) < \min \left\{ I(X_1,X_2,X_3;Y|W), \right. \]
\[ \left. \quad I(X_1,X_2,X_3;Y|W,V_1) + I(X_1,X_3;Y|S_1,X_2,V_2), \right. \]
\[ \left. \quad I(X_1,X_2,X_3;Y|W,V_2) + I(X_2,X_3;Y|S_2,X_1,V_1), \right. \]
\[ \left. \quad I(X_1,X_2;Y|V_1,V_2,X_3,W) + \rho \right\}, \]  
for a joint distribution that factors as

\[
p(s_1, s_2, w_3, w)p(v_1)p(x_1|s_1, v_1) \times \]
\[ p(v_2)p(x_2|s_2, v_2)p(x_3|v_1, v_2)p(y_3|y|x_1, x_2, x_3), \]

where \( \rho \) is defined as follows

\[
\rho \triangleq \min \left\{ I(X_1, X_2, X_3; Y|S_1, S_2), I(X_1, X_3; Y|S_1, X_2, V_2) + I(X_2, X_3; Y|S_2, X_1, V_1) \right\}. \]

**Proof:** A proof outline is given in Subsection II-D. \[ \blacksquare \]

**C. Proof outline of Thm. 1**

1) **Codebook construction:** For \( i = 1, 2 \), assign every \( s_i \in S_i^n \) to one of \( 2^{R_i} \) bins independently according to a uniform distribution on \( U_i \triangleq \{1, 2, \ldots, 2^{R_i}\} \). Denote this assignment by \( f_i, i = 1, 2 \).

For \( i = 1, 2 \), generate \( 2^{R_i} \) codewords \( v_i(u_i), u_i \in U_i \), by choosing the letters \( v_{i,k}(u_i), k = 1, 2, \ldots, n \), independently with distribution \( p_{v_i}(v_{i,k}(u_i)) \). For each pair \( (s_i, u_i), i = 1, 2 \) generate one \( n \)-length codeword \( x_i(s_i, u_i), s_i \in S_i^n, u_i \in U_i \), by choosing the letters \( x_{i,k}(s_i, u_i) \) independently with distribution \( p_{X_i|S_i,V_i}(x_{i,k}|s_{i,k}, v_{i,k}(u_i)) \) for all \( 1 \leq k \leq n \). Finally, generate one \( n \)-length \( n \)-alphabet \( x_3(u_1, u_2) \) independently \( 1 \leq k \leq n \). Finally, generate one \( n \)-length \( n \)-alphabet \( x_3(u_1, u_2) \) for each pair \( (u_1, u_2) \in U_1 \times U_2 \), by choosing \( x_{3,k}(u_1, u_2) \) independently with distribution \( p_{X_3|V_1,V_2}(x_{3,k}|v_{1,k}(u_1), v_{2,k}(u_2)) \) for all \( 1 \leq k \leq n \).

2) **Encoding:** (See Table I). Consider a source sequence \( w_i^{Bn} \in S_i^{Bn}, i = 1, 2 \) of length \( Bn \). Partition this sequence into \( B \) \( n \)-length \( n \)-subsequences, \( s_i,b, b = 1, 2, \ldots, B \). Similarly, partition the side information sequences \( w_3^{Bn} \) and \( w_i^{Bn} \) into \( B \) \( n \)-length \( n \)-subsequences \( w_3,b, w_i,b, b = 1, 2, \ldots, B \) respectively. We transmit a total of \( Bn \) source samples over \( B + 1 \) blocks of \( n \) channel uses each. At block \( 1 \), source terminal \( i, i = 1, 2 \), transmits the channel codeword \( x_i(s_i,1,1) \). At block \( b, b = 2, \ldots, B \), source terminal \( i, i = 1, 2 \), transmits the channel codeword
Having obtained \( \hat{b}_i \) where \( u_i, b_i = f_i(s_{i,b-1}) \in U_i \) is the bin index of source vector \( s_{i,b-1} \). At block \( B + 1 \), source terminal \( i, i = 1, 2 \), transmits \( x_i(a_i, u_i) \) where \( a_i \in S_i^n \) is a fixed sequence.

At block \( b = 1 \), the relay transmits \( x_3(1,1) \). Assume that at block \( b, b = 2, \ldots, B, B+1 \), the relay has estimates \( \hat{s}_{1,b-1}, \hat{s}_{2,b-1} \) of \( s_{1,b-1}, s_{2,b-1} \). It then finds the corresponding bin indices \( \hat{u}_{i,b-1} = f_i(\hat{s}_{i,b-1}) \in U_i, i = 1, 2 \), and transmits the channel codeword \( x_3(\hat{u}_{1,b-1}, \hat{u}_{2,b-1}) \).

3) Decoding: The relay decodes the source sequences sequentially trying to reconstruct source block \( s_{i,b}, i = 1, 2 \), at the end of channel block \( b \) as follows: Using this \( \hat{s}_{1,b-1}, \hat{s}_{2,b-1} \), its received signal \( y_3,b \) and the side information \( w_3,b \), the relay decodes \( s_{1,b}, s_{2,b} \), by looking for a unique pair \( (\hat{s}_1,b, \hat{s}_2,b) \) such that:

\[
(\hat{s}_1,b, \hat{s}_2,b, x_1(\hat{s}_1,b, \hat{u}_{1,b-1}), x_2(\hat{s}_2,b, \hat{u}_{2,b-1}), v_1(\hat{u}_{1,b-1}), v_2(\hat{u}_{2,b-1}), x_3(\hat{u}_{1,b-1}, \hat{u}_{2,b-1}), w_{3,b}, y_{3,b}) \in A_{s}^{(n)}(S_1, S_2, X_1, X_2, V_1, V_2, X_3, W_3, Y_3). \tag{6}
\]

Decoding at the destination is done using successive backward decoding. The destination node waits until the end of channel block \( B + 1 \). It first tries to decode \( (u_1,B, u_2,B) \) using the received signal at channel block \( B + 1 \) and its side information \( w_B \). Going backwards from the last channel block to the first, we assume that the destination has estimates \( \hat{u}_{1,b}, \hat{u}_{2,b} \) of \( u_{1,b}, u_{2,b} \) and consider decoding of \( (s_{1,b}, s_{2,b}) \).

At block \( b \) the destination first estimates the bin indices \( \hat{u}_{i,b-1}, i = 1, 2 \), corresponding to \( s_{i,b-1} \) based on its received signal \( y_b \) and the side information \( w_{b-1} \). More precisely, the destination looks for a unique pair \( (\hat{u}_{1,b-1}, \hat{u}_{2,b-1}) \) such that:

\[
(v_1(\hat{u}_{1,b-1}), v_2(\hat{u}_{2,b-1}), x_3(\hat{u}_{1,b-1}, \hat{u}_{2,b-1}), w_{b-1}, y_b) \in A_{s}^{(n)}(V_1, V_2, X_3, W, Y). \tag{7}
\]

Having obtained \( (\hat{u}_{1,b-1}, \hat{u}_{2,b-1}) \), the destination generates the set \( L(b) \)

\[
L(b) = \{ (\hat{s}_1,b, \hat{s}_2,b) : (\hat{s}_1,b, \hat{s}_2,b, x_1(\hat{s}_1,b, \hat{u}_{1,b-1}), x_2(\hat{s}_2,b, \hat{u}_{2,b-1}), v_1(\hat{u}_{1,b-1}), v_2(\hat{u}_{2,b-1}), x_3(\hat{u}_{1,b-1}, \hat{u}_{2,b-1}), w_b, y_b) \in A_{s}^{(n)}(S_1, S_2, X_1, X_2, V_1, V_2, X_3, W, Y) \}.
\tag{8}
\]

The destination now looks for a unique \( (\hat{s}_1,b, \hat{s}_2,b) \) such that \( f_1(\hat{s}_1,b) = \hat{u}_{1,b}, f_2(\hat{s}_2,b) = \hat{u}_{2,b} \) and \( (\hat{s}_1,b, \hat{s}_2,b) \in L(b) \).

<table>
<thead>
<tr>
<th>Node</th>
<th>Block 1</th>
<th>Block 2</th>
<th>Block B</th>
<th>Block B + 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>User 1</td>
<td>( s_{1,1} )</td>
<td>( s_{1,2} )</td>
<td>( s_{1,B} )</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>User 2</td>
<td>( s_{2,1} )</td>
<td>( s_{2,2} )</td>
<td>( s_{2,B} )</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Relay</td>
<td>( x_3(1,1) )</td>
<td>( x_3(u_1,1) )</td>
<td>( x_3(u_1,B-1, u_2,B) )</td>
<td>( x_3(u_1,B, u_2,B) )</td>
</tr>
</tbody>
</table>

**TABLE I**

Encoding for the joint source-channel achievable rate of Thm. 1.
4) Error probability analysis and corresponding rate constraints: The error probability analysis is given in Appendix A. It is shown that decoding the source sequences at the relay (6) can be done reliably as long as (1a)–(1c) hold and destination decoding (7)–(8) can be done reliably if (1d)–(1f) hold.

D. Proof outline of Thm. 2

1) Codebook construction and encoding: The codebook construction and the encoding are identical to those detailed in Subsection II-C1.

2) Decoding: The relay decoding procedure is identical to the relay decoding procedure detailed in Subsection II-C3.

Decoding at the destination is done using simultaneous backward decoding. Going backwards from the last channel block to the first, we assume that at block \( b \) the destination has estimates \((\hat{u}_{1,b}, \hat{u}_{2,b})\) of \((u_{1,b}, u_{2,b})\). The destination decodes \((s_{1,b}, s_{2,b}, u_{1,b-1}, u_{2,b-1})\) based on the received signal \( y_b \), and the side information \( w_b \), by looking for a unique combination \((\hat{s}_{1,b}, \hat{s}_{2,b}, \hat{u}_{1,b-1}, \hat{u}_{2,b-1})\) that satisfy (8), and \( f_1(\hat{s}_{1,b}) = \hat{u}_{1,b}, f_2(\hat{s}_{2,b}) = \hat{u}_{2,b} \).

3) Error probability analysis and corresponding rate constraints: The error probability analysis is given in Appendix B. It is shown that decoding the source sequences at the relay (6) can be done reliably as long as (3a)–(3c) hold and destination decoding can be done reliably if (3d)–(3f) hold.
APPENDIX A
ERROR PROBABILITY ANALYSIS FOR THM. 1

**Relay error probability**: The relay error probability analysis is identical to the relay error probability analysis detailed in [1, Appendix C].

**Destination error probability**: The average decoding error probability at the destination in block $b$, $P_d^{(n)}$, is defined by

$$P_d^{(n)} \triangleq \sum_{(s_1,b,s_2,b) \in S_n^1 \times S_n^2} p(s_1,b,s_2,b) \Pr \{ \text{destination error}|(s_1,b,s_2,b) \text{ are the source outputs} \}$$

$$\leq \sum_{(s_1,b,s_2,b) \notin A_e^{(n)}(S_1,S_2)} p(s_1,b,s_2,b) \Pr \{ \text{destination error}|(s_1,b,s_2,b) \text{ are the source outputs} \}$$

$$\leq \epsilon + \sum_{(s_1,b,s_2,b) \in A_e^{(n)}(S_1,S_2)} \sum_{\omega_b \in W^n} p(\omega_b|s_1,b,s_2,b) \times \Pr \{ \text{destination error}|(s_1,b,s_2,b) \text{ are the source outputs}, \omega_b \text{ is the side information} \}$$

$$\leq \epsilon + \sum_{(s_1,b,s_2,b) \in A_e^{(n)}(S_1,S_2)} \sum_{\omega_b \notin A_e^{(n)}(W|s_1,b,s_2,b)} p(\omega_b|s_1,b,s_2,b)$$

$$+ \sum_{(s_1,b,s_2,b) \in A_e^{(n)}(S_1,S_2)} \sum_{\omega_b \in A_e^{(n)}(W|s_1,b,s_2,b)} p(\omega_b|s_1,b,s_2,b) \times \Pr \{ \text{destination error}|(s_1,b,s_2,b) \text{ are the source outputs}, \omega_b \text{ is the side information} \}$$

$$\leq 2\epsilon + \sum_{(s_1,b,s_2,b,\omega_b) \in A_e^{(n)}(S_1,S_2,W)} p(s_1,b,s_2,b,\omega_b) \times \Pr \{ \text{destination error}|(s_1,b,s_2,b) \text{ are the source outputs}, \omega_b \text{ is the side information} \},$$

where (A.1a), (A.1c) follows from the union bound and (A.1b), (A.1d) follows from the AEP, for sufficiently large $n$.

In the following we show that for $(s_1,b,s_2,b,\omega_b) \in A_e^{(n)}(S_1,S_2,W)$, the summation terms in (A.1) can be upper bounded independently of $(s_1,b,s_2,b,\omega_b)$. In order to show the above we assume that $(s_1,b,s_2,b,\omega_b) \in A_e^{(n)}(S_1,S_2,W)$ and let $\emptyset$ denote the event that the triplet $(s_1,b,s_2,b,\omega_b)$ is the sources outputs and the side information at the destination.

Let $\epsilon_0$ be a positive number such that $\epsilon_0 > \epsilon$ and $\epsilon_0 \to 0$ as $\epsilon \to 0$. Assuming correct decoding at block $b+1$ (hence $(u_{1,b},u_{2,b})$ are available at the destination), a destination decoding error event $E^d$, in block $b$, is the union
of the following events:

\[ E_1 \triangleq \left\{ (s_1, s_2, x_1(s_1, b), x_2(s_2, u_2, b), v_1(u_1, b), v_2(u_2, b), x_3(u_1, u_2, b), w_b, y_b) \notin A_e^{(n)} \right\}, \]

\[ E_2 \triangleq \left\{ \exists (\hat{u}_1, b, \hat{u}_2, b) : (v_1(\hat{u}_1, b), v_2(\hat{u}_2, b), x_3(\hat{u}_1, b, \hat{u}_2, b), w_b, y_b) \in A_e^{(n)}(V_1, V_2, X_3, W, Y) \right\}, \]

\[ E_3 \triangleq \left\{ \exists (\hat{s}_1, b, \hat{s}_2, b) : f_1(\hat{s}_1, b) = u_1, b, f_2(\hat{s}_2, b) = u_2, b, (\hat{s}_1, b, \hat{s}_2, b) \in L(b) \right\}. \]

From the AEP, for sufficiently large \( n \), \( \Pr \{ E_1 | \mathcal{D} \} \) can be bounded by \( \epsilon \). Therefore by the union bound

\[ \Pr \{ E^d | \mathcal{D} \} \leq \epsilon + \Pr \{ E_2 | \mathcal{D}, (E_1)^c \} + \Pr \{ E_3 | \mathcal{D}, (E_1)^c, (E_2)^c \}. \]  \hspace{1cm} (A.2)

The event \( E_2 \) is the union of the following events:

\[ E_{21} \triangleq \left\{ \exists \hat{u}_1, b, b \neq u_1, b, 1 : (v_1(\hat{u}_1, b), v_2(u_2, b), x_3(\hat{u}_1, b, u_2, b), w_b, y_b) \in A_e^{(n)}(V_1, V_2, X_3, W, Y), \right\}, \]

\[ E_{22} \triangleq \left\{ \exists \hat{u}_2, b, \hat{u}_2, b, \neq u_2, b, 1 : (v_1(u_1, b), v_2(\hat{u}_2, b), x_3(u_1, b, \hat{u}_2, b), w_b, y_b) \in A_e^{(n)}(V_1, V_2, X_3, W, Y), \right\}, \]

\[ E_{23} \triangleq \left\{ \exists \hat{u}_1, b, \hat{u}_1, b, \neq u_1, b, 1, \hat{u}_2, b, \neq u_2, b, 1 : (v_1(\hat{u}_1, b), v_2(\hat{u}_2, b), x_3(\hat{u}_1, b, \hat{u}_2, b), w_b, y_b) \in A_e^{(n)}(V_1, V_2, X_3, W, Y). \right\}

Hence, by the union bound

\[ \Pr \{ E_2 | \mathcal{D}, (E_1)^c \} \leq \sum_{j=1}^{3} \Pr \{ E_{2j} | \mathcal{D}, (E_1)^c \}. \]  \hspace{1cm} (A.3)

Using [2, Thm. 14.2.3] it can be shown that \( \Pr \{ E_{21} | \mathcal{D}, (E_1)^c \} \) can be bounded by \( \epsilon \), for large enough \( n \), if

\[ R_1 < I(V_1, X_3; Y|V_2, W) - \epsilon_0, \]  \hspace{1cm} (A.4)

\[ \Pr \{ E_{22} | \mathcal{D}, (E_1)^c \} \] can be bounded by \( \epsilon \), for large enough \( n \), if

\[ R_2 < I(V_2, X_3; Y|V_1, W) - \epsilon_0, \]  \hspace{1cm} (A.5)

and \( \Pr \{ E_{23} | \mathcal{D}, (E_1)^c \} \) can be bounded by \( \epsilon \), for large enough \( n \), if

\[ R_1 + R_2 < I(V_1, V_2, X_3; Y|W) - \epsilon_0. \]  \hspace{1cm} (A.6)

Therefore, if conditions (A.4)–(A.6) hold, for large enough \( n \),

\[ \Pr \{ E_2 | \mathcal{D}, (E_1)^c \} \leq 3 \epsilon. \]  \hspace{1cm} (A.7)
The event $E_3$ is the union of the following events:

\begin{align*}
E_{31} & \triangleq \{ \exists s_{1,b} \neq s_{1,b} : f_1(\hat{s}_{1,b}) = u_{1,b}, \text{ and } \hat{s}_{1,b} \in L_1(b) \}, \\
E_{32} & \triangleq \{ \exists s_{2,b} \neq s_{2,b} : f_2(\hat{s}_{2,b}) = u_{2,b}, \text{ and } \hat{s}_{2,b} \in L_2(b) \}, \\
E_{33} & \triangleq \{ \exists s_{1,b} \neq s_{1,b}, \hat{s}_{2,b} \neq s_{2,b} : f_1(\hat{s}_{1,b}) = u_{1,b}, f_2(\hat{s}_{2,b}) = u_{2,b}, \text{ and } (\hat{s}_{1,b}, \hat{s}_{2,b}) \in L(b) \},
\end{align*}

where $L_1(b), L_2(b)$, are the following sets

\begin{align*}
L_1(b) & \triangleq \{ \hat{s}_{1,b} : (\hat{s}_{1,b}, s_{2,b}, x_1(\hat{s}_{1,b}, u_{1,b-1}), x_2(s_{2,b}, u_{2,b-1}), v_1(u_{1,b-1}), v_2(u_{2,b-1}), \\
& \quad x_3(u_{1,b-1}, u_{2,b-1}), w_b, y_b) \in A^e(n)(S_1, S_2, X_1, X_2, V_1, V_2, X_3, W, Y) \}, \\
L_2(b) & \triangleq \{ \hat{s}_{2,b} : (s_{1,b}, \hat{s}_{2,b}, x_1(s_{1,b}, u_{1,b-1}), x_2(\hat{s}_{2,b}, u_{2,b-1}), v_1(u_{1,b-1}), v_2(u_{2,b-1}), \\
& \quad x_3(u_{1,b-1}, u_{2,b-1}), w_b, y_b) \in A^e(n)(S_1, S_2, X_1, X_2, V_1, V_2, X_3, W, Y) \}.
\end{align*}

Hence, by the union bound

\begin{equation}
\Pr\{E_3|\mathcal{D}, (E_1)^c, (E_2)^c\} \leq \sum_{j=1}^{3} \Pr\{E_{3j}|\mathcal{D}, (E_1)^c, (E_2)^c\},
\end{equation}

Pr $\{E_{31}|\mathcal{D}, (E_1)^c, (E_2)^c\}$ is bounded as follows

\begin{align*}
\Pr\{E_{31}|\mathcal{D}, (E_1)^c, (E_2)^c\} &= \Pr\{\exists s_{1,b} \neq s_{1,b} : f_1(\hat{s}_{1,b}) = u_{1,b}, \text{ and } \hat{s}_{1,b} \in L_1(b)|\mathcal{D}, (E_1)^c, (E_2)^c\} \\
& \leq \mathbb{E}_{x_2,v_1,v_2,x_3,Y} \left\{ \sum_{\hat{s}_{1,b} \neq s_{1,b} \in L_1(b)} \Pr\{f_1(\hat{s}_{1,b}) = u_{1,b}\}|\mathcal{D}, (E_1)^c \right\} \tag{A.10a} \\
& \leq \mathbb{E}_{x_2,v_1,v_2,x_3,Y} \left\{ \|L_1(b)\| 2^{-nR_1}|\mathcal{D}, (E_1)^c \right\} \\
& \leq 2^{-nR_1} \left( 1 + 2^n[H(S_1|S_2, W) - I(X_1; Y|S_2, X_2, V_1, X_3, W) + 5\epsilon_0] \right),
\end{align*}

where (A.10a) follows from the union bound and (A.10b) follows from (C.2). From (A.10) it follows that Pr $\{E_{31}|\mathcal{D}, (E_1)^c, (E_2)^c\}$ can be bounded by $\epsilon$, for large enough $n$, if

\begin{equation}
R_1 > H(S_1|S_2, W) - I(X_1; Y|S_2, X_2, V_1, X_3, W) + 5\epsilon_0.
\end{equation}

Combining conditions (A.4) and (A.11) yields

\begin{equation}
H(S_1|S_2, W) < I(X_1; Y|S_2, X_2, V_1, X_3, W) + I(V_1, X_3; Y|V_2, W) - 6\epsilon_0.
\end{equation}

Following similar arguments as in (A.10)–(A.12), we can also show that Pr $\{E_{32}|\mathcal{D}, (E_1)^c, (E_2)^c\}$ can be bounded by $\epsilon$, for large enough $n$, if

\begin{equation}
H(S_2|S_1, W) < I(X_2; Y|S_1, X_1, V_2, X_3, W) + I(V_2, X_3; Y|V_1, W) - 6\epsilon_0,
\end{equation}

and Pr $\{E_{33}|\mathcal{D}, (E_1)^c, (E_2)^c\}$ can be bounded by $\epsilon$, for large enough $n$, if

\begin{equation}
H(S_1, S_2|W) < I(X_1, X_2; Y|V_1, V_2, X_3, W) + I(V_1, V_2, X_3; Y|W) - 6\epsilon_0.
\end{equation}
Hence if conditions (A.12)–(A.14) hold, for large enough $n$,

$$\Pr \{ E_3 | \mathcal{D}, (E_1^c)^c \} \leq 3\epsilon.$$  \hfill (A.15)

Combining equations (A.1), (A.2), (A.7) and (A.15) yields

$$\bar{P}_d^{(n)} \leq 9\epsilon.$$  \hfill (A.16)

### Appendix B

**Error Probability Analysis for Thm. 2**

**Relay error probability:** The relay error probability analysis is identical to the relay error probability analysis detailed in [1, Appendix C].

**Destination error probability:** The average decoding error probability at the destination in block $b$, $\bar{P}_d^{(n)}$, is defined by

$$\bar{P}_d^{(n)} \triangleq \sum_{(s_1,b,s_2,b) \in S}\sum_{s_2,b} p(s_1,b,s_2,b) \Pr \{ \text{destination error} | (s_1,b,s_2,b) \text{ are the source outputs} \}$$

$$\leq 2\epsilon + \sum_{(s_1,b,s_2,b,w_b) \in A^{(n)}_c(S_1,S_2,W)} \Pr \{ \text{destination error} | (s_1,b,s_2,b) \text{ are the source outputs},$$

$$w_b \text{ is the side information} \},$$  \hfill (B.1)

where (B.1) follows the same arguments led to (A.1d).

In the following we show that for $(s_1,b,s_2,b,w_b) \in A^{(n)}_c(S_1,S_2,W)$, the summation terms in (A.1) can be upper bounded independently of $(s_1,b,s_2,b,w_b)$. In order to show the above we assume that $(s_1,b,s_2,b,w_b) \in A^{(n)}_c(S_1,S_2,W)$ and let $\mathcal{D}$ denote the event that the triplet $(s_1,b,s_2,b,w_b)$ is the sources outputs and the side information at the destination.

Let $\epsilon_0$ be a positive number such that $\epsilon_0 > \epsilon$ and $\epsilon_0 \to 0$ as $\epsilon \to 0$. Assuming correct decoding at block $b+1$ (hence $(u_1,b,u_2,b)$ are available at the destination), a destination decoding error event, $E_d$, in block $b$, is the union of the following events:

$$E_1 \triangleq \{ (s_1,b,s_2,b,x_1(s_1,b,u_1,b-1),x_2(s_2,b,u_2,b-1),$$

$$v_1(u_1,b-1),v_2(u_2,b-1),x_3(u_1,b-1,u_2,b-1),w_b,y_b) \notin A^{(n)}_c \},$$

$$E_2 \triangleq \{ (s_1,b,s_2,b,u_1,b-1,u_2,b-1) \neq (s_1,b,s_2,b,u_1,b-1,u_2,b-1) :$$

$$\exists (s_1,b,s_2,b,u_1,b-1,u_2,b-1) \neq (s_1,b,s_2,b,u_1,b-1,u_2,b-1) :$$

$$(s_1,b,s_2,b,x_1(s_1,b,u_1,b-1),x_2(s_2,b,u_2,b-1),v_1(u_1,b-1),v_2(u_2,b-1),$$

$$x_3(u_1,b-1,u_2,b-1),w_b,y_b) \in A^{(n)}_c(S_1,S_2,X_1,X_2,V_1,V_2,X_3,W,Y).$$

From the AEP, for sufficiently large $n$, $\Pr \{ E_1 | \mathcal{D} \}$ can be bounded by $\epsilon$. Therefore by the union bound

$$\Pr \{ E_d | \mathcal{D} \} \leq \epsilon + \Pr \{ E_2 | \mathcal{D} , (E_1)^c \}.$$  \hfill (B.2)
Let the sets \( L_{1,b}(s_1, u_1) \), \( L_{2,b}(s_2, u_2) \) and \( L_b(s_1, s_2, u_1, u_2) \) be defined as follows

\[
L_{1,b}(s_1, u_1) \triangleq \{(s_1, u_1) : (s_1, s_2, b, x_1(s_1, u_1), x_2(s_2, u_2), v_1(u_1), v_2(u_2), x_3(u_1, u_2, b), w, y) \in A_e^{(n)}\},
\]

(B.3a)

\[
L_{2,b}(s_2, u_2) \triangleq \{(s_2, u_2) : (s_1, s_2, b, x_1(s_1, u_1, b), x_2(s_2, u_2), v_1(u_1, b), v_2(u_2), x_3(u_1, u_2, b), w, y) \in A_e^{(n)}\},
\]

(B.3b)

\[
L_b(s_1, s_2, u_1, u_2) \triangleq \{(s_1, s_2, u_1, u_2) : (s_1, s_2, b, x_1(s_1, u_1), x_2(s_2, u_2), v_1(u_1), v_2(u_2), x_3(u_1, u_2, b), w, y) \in A_e^{(n)}\}.
\]

(B.3c)

The subset \( L_{1,b}^*(s_1 = s_1, b, u_1 \neq u_1, b) \) of \( L_{1,b}(s_1, u_1) \) is defined as follows

\[
L_{1,b}^*(s_1 = s_1, b, u_1 \neq u_1, b) \triangleq \{(s_1, u_1) : s_1 = s_1, b, u_1 \neq u_1, b, (s_1, u_1) \in L_{1,b}(s_1, u_1)\}.
\]

The other respective subsets of \( L_{1,b}(s_1, u_1) \), and the corresponding subsets of \( L_{2,b}(s_2, u_2) \) and \( L_b(s_1, s_2, u_1, u_2) \) are defined similarly. The event \( E_2 \) is the union of the following events:

\[
E_{21} \triangleq \{\exists (s_1, b, u_1, b) : f_1(s_1) = u_1, b, (s_1, b, u_1, b) \in L_{1,b}(s_1, b, u_1, b)\},
\]

\[
E_{22} \triangleq \{\exists (s_2, b, u_2, b) : f_2(s_2) = u_2, b, (s_2, b, u_2, b) \in L_{2,b}(s_2, b, u_2, b)\},
\]

\[
E_{23} \triangleq \{\exists (s_1, b, u_1, b) : f_1(s_1) = u_1, b, (s_1, b, u_1, b) \in L_b(s_1, b, u_1, b)\}.
\]

Hence, by the union bound

\[
\Pr \{E_2|\mathcal{D}, (E_1)^c\} \leq \sum_{j=1}^{3} \Pr \{E_{2j}|\mathcal{D}, (E_1)^c\}.
\]

(B.4)

In the following subsections we bound \( \Pr \{E_{21}|\mathcal{D}, (E_1)^c\} \), \( \Pr \{E_{22}|\mathcal{D}, (E_1)^c\} \) and \( \Pr \{E_{23}|\mathcal{D}, (E_1)^c\} \).

A. Bounding The Probabilities \( \Pr \{E_{21}|\mathcal{D}, (E_1)^c\} \), \( \Pr \{E_{22}|\mathcal{D}, (E_1)^c\} \)

The event \( E_{21} \) is the union of the following events

\[
E_{211} \triangleq \{\exists (s_1, b, u_1, b) : (s_1, b, u_1, b) \in L_{1,b}^*(s_1 = s_1, b, u_1 \neq u_1, b)\},
\]

\[
E_{212} \triangleq \{\exists (s_1, b) : f_1(s_1) = u_1, b, (s_1, b) \in L_{1,b}^*(s_1 \neq s_1, b, u_1 = u_1, b)\},
\]

\[
E_{213} \triangleq \{\exists (s_1, b) : f_1(s_1) = u_1, b, (s_1, b) \in L_{1,b}^*(s_1 \neq s_1, b, u_1 \neq u_1, b)\}.
\]
Hence, by the union bound
\[ \Pr \{ E_{21} | \mathcal{D}, (E_1)^c \} \leq \sum_{j=1}^{3} \Pr \{ E_{21j} | \mathcal{D}, (E_1)^c \}. \] (B.5)

Following similar arguments as in [1, Equations (C.16)–(C.20)] we can show that \( \Pr \{ E_{211} | \mathcal{D}, (E_1)^c \} \) can be bounded by \( \epsilon \), for large enough \( n \), if
\[ R_1 < I(X_1, X_3; Y | S_1, X_2, V_2) - \epsilon_0. \] (B.6)

Following arguments similar to (A.10)–(A.11) we can show that \( \Pr \{ E_{212} | \mathcal{D}, (E_1)^c \} \) can be bounded by \( \epsilon \), for large enough \( n \), if
\[ R_1 > H(S_1 | S_2, W) - I(X_1; Y | S_2, X_2, V_1, X_3, W) + 5\epsilon_0, \] (B.7)

which can be written as follows
\[ H(S_1 | S_2, W) - R_1 < I(X_1; Y | S_2, X_2, V_1, X_3, W) - 5\epsilon_0. \] (B.8)

\( \Pr \{ E_{213} | \mathcal{D}, (E_1)^c \} \) is bounded as follows
\[
\Pr \{ E_{213} | \mathcal{D}, (E_1)^c \} = \Pr \{ \exists(\hat{s}_{1,b}, u_{1,b-1}) : f_1(\hat{s}_{1,b}) = u_{1,b}, \text{ and } (\hat{s}_{1,b}, \hat{u}_{1,b-1}) \in \mathcal{L}_{1,b}^*(\hat{s}_1 \neq s_{1,b}, u_1 \neq u_{1,b-1}) | \mathcal{D}, (E_1)^c \}
\leq \mathbb{E}_{X_2, V_2, Y} \left\{ \sum_{(\hat{s}_{1,b}, \hat{u}_{1,b-1}) \in \mathcal{L}_{1,b}^*(\hat{s}_1 \neq s_{1,b}, u_1 \neq u_{1,b-1})} \Pr \{ f_1(\hat{s}_{1,b}) = u_{1,b} \} | \mathcal{D}, (E_1)^c \right\} \] (B.9a)
\[
\leq 2^{-nR_1} \cdot 2^{n[H(S_1 | S_2, W) + R_1 - I(X_1, X_3; Y | S_2, W, X_2, V_2) + 5\epsilon_0]}, \] (B.9b)

where (B.9a) follows from the union bound and (B.9b) follows from (D.1). From (B.9) it follows that \( \Pr \{ E_{213} | \mathcal{D}, (E_1)^c \} \) can be bounded by \( \epsilon \), for large enough \( n \), if
\[ R_1 > H(S_1 | S_2, W) + R_1 - I(X_1, X_3; Y | S_2, W, X_2, V_2) + 5\epsilon_0. \] (B.10)

which can be written as follows
\[ H(S_1 | S_2, W) < I(X_1, X_3; Y | S_2, W, X_2, V_2) - 5\epsilon_0. \] (B.11)

Following similar arguments as in (B.5)–(B.11) we can show that \( \Pr \{ E_{22} | \mathcal{D}, (E_1)^c \} \) can be bounded by \( 3\epsilon \), for large enough \( n \), if
\[
R_1 < I(X_2, X_3; Y | S_2, X_1, V_1) - \epsilon_0 \] (B.12a)
\[
H(S_2 | S_1, W) - R_2 < I(X_2; Y | S_1, W, X_1, V_2, X_3) - 5\epsilon_0 \] (B.12b)
\[
H(S_1 | S_2, W) < I(X_2, X_3; Y | S_1, W, X_1, V_1) - 5\epsilon_0. \] (B.12c)
B. Bounding The Probability \( \Pr \{ E_{23} | \mathcal{D}, (E_1)^c \} \)

The event \( E_{23} \) is the union of the following events

\[
E_{231} \triangleq \{ \exists (\hat{u}_{1,b-1}, \hat{u}_{2,b-1}) : \\
(\hat{s}_{1,b}, \hat{s}_{2,b}, \hat{u}_{1,b-1}, \hat{u}_{2,b-1}) \in \mathcal{L}_b^\ast (s_1 = s_{1,b}, s_2 = s_{2,b}, u_1 \neq u_{1,b-1}, u_2 \neq u_{2,b-1}) \},
\]

\[
E_{232} \triangleq \{ \exists (\hat{u}_{1,b-1}, \hat{s}_{2,b}) : \\
f_2(\hat{s}_{2,b}) = u_{2,b}, \text{ and } (\hat{s}_{1,b}, \hat{s}_{2,b}, \hat{u}_{1,b-1}, \hat{u}_{2,b-1}) \in \mathcal{L}_b^\ast (s_1 = s_{1,b}, s_2 \neq s_{2,b}, u_1 \neq u_{1,b-1}, u_2 = u_{2,b-1}) \},
\]

\[
E_{233} \triangleq \{ \exists (\hat{u}_{2,b-1}, \hat{s}_{1,b}) : \\
f_1(\hat{s}_{1,b}) = u_{1,b}, \text{ and } (\hat{s}_{1,b}, \hat{s}_{2,b}, \hat{u}_{1,b-1}, \hat{u}_{2,b-1}) \in \mathcal{L}_b^\ast (s_1 \neq s_{1,b}, s_2 = s_{2,b}, u_1 = u_{1,b-1}, u_2 \neq u_{2,b-1}) \},
\]

\[
E_{234} \triangleq \{ \exists (\hat{s}_{1,b}, \hat{s}_{2,b}) : \\
f_1(\hat{s}_{1,b}) = u_{1,b}, f_2(\hat{s}_{2,b}) = u_{2,b} \text{ and } \\
(\hat{s}_{1,b}, \hat{s}_{2,b}, \hat{u}_{1,b-1}, \hat{u}_{2,b-1}) \in \mathcal{L}_b^\ast (s_1 \neq s_{1,b}, s_2 \neq s_{2,b}, u_1 = u_{1,b-1}, u_2 = u_{2,b-1}) \},
\]

\[
E_{235} \triangleq \{ \exists (\hat{u}_{1,b-1}, \hat{u}_{2,b-1}, \hat{s}_{1,b}) : \\
f_1(\hat{s}_{1,b}) = u_{1,b} \text{ and } (\hat{s}_{1,b}, \hat{s}_{2,b}, \hat{u}_{1,b-1}, \hat{u}_{2,b-1}) \in \mathcal{L}_b^\ast (s_1 \neq s_{1,b}, s_2 = s_{2,b}, u_1 \neq u_{1,b-1}, u_2 \neq u_{2,b-1}) \},
\]

\[
E_{236} \triangleq \{ \exists (\hat{u}_{1,b-1}, \hat{s}_{1,b}, \hat{s}_{2,b}) : \\
f_1(\hat{s}_{1,b}) = u_{1,b}, f_2(\hat{s}_{2,b}) = u_{2,b} \text{ and } \\
(\hat{s}_{1,b}, \hat{s}_{2,b}, \hat{u}_{1,b-1}, \hat{u}_{2,b-1}) \in \mathcal{L}_b^\ast (s_1 \neq s_{1,b}, s_2 \neq s_{2,b}, u_1 \neq u_{1,b-1}, u_2 = u_{2,b-1}) \},
\]

\[
E_{237} \triangleq \{ \exists (\hat{u}_{2,b-1}, \hat{s}_{1,b}, \hat{s}_{2,b}) : \\
f_2(\hat{s}_{2,b}) = u_{2,b} \text{ and } (\hat{s}_{1,b}, \hat{s}_{2,b}, \hat{u}_{1,b-1}, \hat{u}_{2,b-1}) \in \mathcal{L}_b^\ast (s_1 = s_{1,b}, s_2 \neq s_{2,b}, u_1 \neq u_{1,b-1}, u_2 \neq u_{2,b-1}) \},
\]

\[
E_{238} \triangleq \{ \exists (\hat{u}_{2,b-1}, \hat{s}_{1,b}, \hat{s}_{2,b}) : \\
f_1(\hat{s}_{1,b}) = u_{1,b}, f_2(\hat{s}_{2,b}) = u_{2,b} \text{ and } \\
(\hat{s}_{1,b}, \hat{s}_{2,b}, \hat{u}_{1,b-1}, \hat{u}_{2,b-1}) \in \mathcal{L}_b^\ast (s_1 \neq s_{1,b}, s_2 \neq s_{2,b}, u_1 = u_{1,b-1}, u_2 \neq u_{2,b-1}) \},
\]

\[
E_{239} \triangleq \{ \exists (\hat{u}_{1,b-1}, \hat{u}_{2,b-1}, \hat{s}_{1,b}, \hat{s}_{2,b}) : \\
f_1(\hat{s}_{1,b}) = u_{1,b}, f_2(\hat{s}_{2,b}) = u_{2,b} \text{ and } \\
(\hat{s}_{1,b}, \hat{s}_{2,b}, \hat{u}_{1,b-1}, \hat{u}_{2,b-1}) \in \mathcal{L}_b^\ast (s_1 \neq s_{1,b}, s_2 \neq s_{2,b}, u_1 \neq u_{1,b-1}, u_2 \neq u_{2,b-1}) \}.
\]

Hence, by the union bound

\[
\Pr \{ E_{23} | \mathcal{D}, (E_1)^c \} \leq \sum_{j=1}^{9} \Pr \{ E_{23j} | \mathcal{D}, (E_1)^c \}. \quad (B.13)
\]
1) Bounding $\Pr \{E_{231}|\mathcal{D}, (E_1)^c\}$: Following similar arguments as in [1, Equation (C.22)] we can show that $\Pr \{E_{231}|\mathcal{D}, (E_1)^c\}$ can be bounded by $\epsilon$, for large enough $n$, if

$$R_1 + R_2 < I(X_1, X_2, X_3; Y|S_1, S_2) - \epsilon_0. \tag{B.14}$$

2) Bounding $\Pr \{E_{232}|\mathcal{D}, (E_1)^c\}$: $\Pr \{E_{232}|\mathcal{D}, (E_1)^c\}$ is bounded as follows

$$\Pr \{E_{232}|\mathcal{D}, (E_1)^c\} = \Pr \{\exists (\hat{u}_{1,b-1}, \hat{s}_{2,b}) : f_2(\hat{s}_{2,b}) = u_{2,b}, \text{ and} \}

\{(\hat{s}_{1,b}, \hat{s}_{2,b}, \hat{u}_{1,b-1}, \hat{u}_{2,b-1}) \in \mathcal{L}_b^c(s_1 = s_{1,b}, s_2 \neq s_{2,b}, u_1 \neq u_{1,b-1}, u_2 = u_{2,b-1})\}

\leq E_{v_1,Y} \left\{ \left(\mathcal{L}_b^c(s_1 = s_{1,b}, s_2 \neq s_{2,b}, u_1 \neq u_{1,b-1}, u_2 = u_{2,b-1})\right) \sum \Pr \{f_2(\hat{s}_{2,b}) = u_{2,b}\} \right\} \tag{B.15a}

\leq E_{v_1,Y} \left\{ \left(\mathcal{L}_b^c(s_1 = s_{1,b}, s_2 \neq s_{2,b}, u_1 \neq u_{1,b-1}, u_2 = u_{2,b-1})\right) \|2^{-nR_2}|\mathcal{D}, (E_1)^c\right\}

\leq 2^{-nR_2} \cdot 2^{n[H(S_2|S_1, W) + R_1 - I(X_1, X_2, X_3; Y|S_1, W, V_2) - \epsilon_0]}, \tag{B.15b}

where (B.15a) follows from the union bound and (B.15b) follows from (E.1). From (B.15) it follows that $\Pr \{E_{232}|\mathcal{D}, (E_1)^c\}$ can be bounded by $\epsilon$, for large enough $n$, if

$$R_2 > H(S_2|S_1, W) + R_1 - I(X_1, X_2, X_3; Y|S_1, W, V_2) + 5\epsilon_0. \tag{B.16}$$

which is equivalent to the following condition

$$H(S_2|S_1, W) + R_1 - R_2 < I(X_1, X_2, X_3; Y|S_1, W, V_2) - 5\epsilon_0. \tag{B.17}$$

3) Bounding $\Pr \{E_{233}|\mathcal{D}, (E_1)^c\}$: Following arguments similar to the derivation of (B.17), we can show that $\Pr \{E_{233}|\mathcal{D}, (E_1)^c\}$ can be bounded by $\epsilon$, for large enough $n$, if

$$R_1 > H(S_1|S_2, W) + R_2 - I(X_1, X_2, X_3; Y|S_2, W, V_1) + 5\epsilon_0. \tag{B.18}$$

which can be written as follows

$$H(S_1|S_2, W) + R_2 - R_1 < I(X_1, X_2, X_3; Y|S_2, W, V_1) - 5\epsilon_0. \tag{B.19}$$

4) Bounding $\Pr \{E_{234}|\mathcal{D}, (E_1)^c\}$: Following arguments similar to the derivation of (A.14), we can show that $\Pr \{E_{234}|\mathcal{D}, (E_1)^c\}$ can be bounded by $\epsilon$, for large enough $n$, if the following conditions hold

$$R_1 + R_2 > H(S_1, S_2|W) - I(X_1, X_2; Y|V_1, V_2, X_3, W) + 5\epsilon_0. \tag{B.20}$$

which can be written as follows

$$H(S_1, S_2|W) - R_1 - R_2 < I(X_1, X_2; Y|V_1, V_2, X_3, W) - 5\epsilon_0. \tag{B.21}$$
5) **Bounding** $\Pr \{ E_{235} | \mathcal{D}, (E_1)^c \}$: Following arguments similar to the derivation of (B.17), we can show that $\Pr \{ E_{235} | \mathcal{D}, (E_1)^c \}$ can be bounded by $\epsilon$, for large enough $n$, if

$$R_1 > H(S_1|S_2,W) + R_1 + R_2 - I(X_1, X_2, X_3; Y|S_2, W) + 5\epsilon_0. \quad \text{(B.22)}$$

which can be written as follows

$$H(S_1|S_2,W) + R_2 < I(X_1, X_2, X_3; Y|S_2, W) - 5\epsilon_0. \quad \text{(B.23)}$$

6) **Bounding** $\Pr \{ E_{236} | \mathcal{D}, (E_1)^c \}$: Following arguments similar to the derivation of (B.17), we can show that $\Pr \{ E_{236} | \mathcal{D}, (E_1)^c \}$ can be bounded by $\epsilon$, for large enough $n$, if

$$R_1 + R_2 > H(S_1, S_2|W) + R_1 + R_2 - I(X_1, X_2, X_3; Y|W, V_2) + 5\epsilon_0. \quad \text{(B.24)}$$

which can be written as follows

$$H(S_1, S_2|W) - R_2 < I(X_1, X_2, X_3; Y|W, V_2) - 5\epsilon_0. \quad \text{(B.25)}$$

7) **Bounding** $\Pr \{ E_{237} | \mathcal{D}, (E_1)^c \}$: Following arguments similar to the derivation of (B.22), we can show that $\Pr \{ E_{237} | \mathcal{D}, (E_1)^c \}$ can be bounded by $\epsilon$, for large enough $n$, if

$$R_2 > H(S_2|S_1,W) + R_1 + R_2 - I(X_1, X_2, X_3; Y|S_1, W) + 5\epsilon_0. \quad \text{(B.26)}$$

which can be written as follows

$$H(S_2|S_1,W) + R_1 < I(X_1, X_2, X_3; Y|S_1, W) - 5\epsilon_0. \quad \text{(B.27)}$$

8) **Bounding** $\Pr \{ E_{238} | \mathcal{D}, (E_1)^c \}$: Following arguments similar to the derivation of (B.24), we can show that $\Pr \{ E_{238} | \mathcal{D}, (E_1)^c \}$ can be bounded by $\epsilon$, for large enough $n$, if

$$R_1 + R_2 > H(S_1, S_2|W) + R_2 - I(X_1, X_2, X_3; Y|W, V_1) + 5\epsilon_0. \quad \text{(B.28)}$$

which can be written as follows

$$H(S_1, S_2|W) - R_1 < I(X_1, X_2, X_3; Y|W, V_1) - 5\epsilon_0. \quad \text{(B.29)}$$

9) **Bounding** $\Pr \{ E_{239} | \mathcal{D}, (E_1)^c \}$: Following arguments similar to the derivation of (B.24), we can show that $\Pr \{ E_{239} | \mathcal{D}, (E_1)^c \}$ can be bounded by $\epsilon$, for large enough $n$, if

$$R_1 + R_2 > H(S_1, S_2|W) + R_1 + R_2 - I(X_1, X_2, X_3; Y|W) + 5\epsilon_0. \quad \text{(B.30)}$$

which can be written as follows

$$H(S_1, S_2|W) < I(X_1, X_2, X_3; Y|W) - 5\epsilon_0. \quad \text{(B.31)}$$
C. Simplifying the Decoding Constraints

Combining the constraints for reliable decoding detailed in (B.6)–(B.31) results in the following set of constraints:

\[ R_1 < I(X_1, X_3; Y|S_1, X_2, V_2) \]  \hfill (B.32a)

\[ H(S_1|S_2, W) - R_1 < I(X_1; Y|S_2, X_2, V_1, X_3, W) \]  \hfill (B.32b)

\[ H(S_1|S_2, W) < I(X_1, X_3; Y|S_2, W, X_2, V_2) \]  \hfill (B.32c)

\[ R_2 < I(X_2, X_3; Y|S_2, X_1, V_1) \]  \hfill (B.32d)

\[ H(S_2|S_1, W) - R_2 < I(X_2; Y|S_1, X_1, V_2, X_3, W) \]  \hfill (B.32e)

\[ H(S_2|S_1, W) < I(X_2, X_3; Y|S_1, W, X_1, V_1) \]  \hfill (B.32f)

\[ R_1 + R_2 < I(X_1, X_2, X_3; Y|S_1, S_2) \]  \hfill (B.32g)

\[ H(S_2|S_1, W) + R_1 - R_2 < I(X_1, X_2, X_3; Y|S_1, W, V_2) \]  \hfill (B.32h)

\[ H(S_1|S_2, W) - R_1 + R_2 < I(X_1, X_2, X_3; Y|S_2, W, V_1) \]  \hfill (B.32i)

\[ H(S_1, S_2|W) - R_1 - R_2 < I(X_1, X_2; Y|V_1, V_2, X_3, W) \]  \hfill (B.32j)

\[ H(S_1|S_2, W) + R_2 < I(X_1, X_2, X_3; Y|S_2, W) \]  \hfill (B.32k)

\[ H(S_1, S_2|W) - R_2 < I(X_1, X_2, X_3; Y|W, V_2) \]  \hfill (B.32l)

\[ H(S_2|S_1, W) + R_1 < I(X_1, X_2, X_3; Y|S_1, W) \]  \hfill (B.32m)

\[ H(S_1, S_2|W) - R_1 < I(X_1, X_2, X_3; Y|W, V_1) \]  \hfill (B.32n)

\[ H(S_1, S_2|W) < I(X_1, X_2, X_3; Y|W) \]. \hfill (B.32o)

From constraints (B.32a), (B.32d) and (B.32g) we get the following set of constraints for \( R_1, R_2, \) and \( R_1 + R_2 \):

\[ R_1 < \min \{I(X_1, X_3; Y|S_1, X_2, V_2), I(X_1, X_2, X_3; Y|S_1, S_2)\} = I(X_1, X_3; Y|S_1, X_2, V_2) \]  \hfill (B.33a)

\[ R_2 < \min \{I(X_2, X_3; Y|S_2, X_1, V_1), I(X_1, X_2, X_3; Y|S_1, S_2)\} = I(X_2, X_3; Y|S_2, X_1, V_1) \]  \hfill (B.33b)

\[ R_1 + R_2 < \min \{I(X_1, X_2, X_3; Y|S_1, S_2), I(X_1, X_3; Y|S_1, X_2, V_2) + I(X_2, X_3; Y|S_2, X_1, V_1)\} = \rho. \]  \hfill (B.33c)

From constrains (B.32b), (B.32c), (B.32i), (B.32k) and (B.33) we get the following set of constraints for \( H(S_1|S_2, W) \):

\[ H(S_1|S_2, W) < I(X_1; Y|S_2, X_2, V_1, X_3, W) + I(X_1, X_3; Y|S_1, X_2, V_2) \]  \hfill (B.34a)

\[ H(S_1|S_2, W) < I(X_1, X_3; Y|S_2, W, X_2, V_2) \]  \hfill (B.34b)

\[ H(S_1|S_2, W) < I(X_1, X_2, X_3; Y|S_2, W, V_1) + I(X_1, X_3; Y|S_1, X_2, V_2) \]  \hfill (B.34c)

\[ H(S_1|S_2, W) < I(X_1, X_2, X_3; Y|S_2, W). \]  \hfill (B.34d)
Constraint (B.34d) is redundant due to constraint (B.34b). Constraint (B.34c) is redundant due to constraint (B.34a). Therefore, (B.34) can be simplified to

\[ H(S_1|S_2, W) < \min \left\{ I(X_1, X_3; Y|S_2, W, X_2, V_2), I(X_1; Y|S_2, X_2, V_1, X_3, W) + I(X_1, X_3; Y|S_1, X_2, V_2) \right\}. \]

(B.35)

Following arguments similar to the derivation of (B.35), we can show that

\[ H(S_2|S_1, W) < \min \left\{ I(X_2, X_3; Y|S_1, W, X_1, V_1), I(X_2; Y|S_1, X_1, V_2, X_3, W) + I(X_2, X_3; Y|S_2, X_1, V_1) \right\}. \]

(B.36)

From constrains (B.32j), (B.32l), (B.32n), (B.32o) and (B.33) we get the following set of constraints for \( H(S_1, S_2|W) \)

\[
\begin{align*}
H(S_1, S_2|W) &< I(X_1, X_2; Y|V_1, V_2, X_3, W)+\rho \\
H(S_1, S_2|W) &< I(X_1, X_2, X_3; Y|W, V_2) + I(X_2, X_3; Y|S_2, X_1, V_1) \\
H(S_1, S_2|W) &< I(X_1, X_2, X_3; Y|W, V_1) + I(X_1, X_3; Y|S_1, X_2, V_2) \tag{B.37b} \\
H(S_1, S_2|W) &< I(X_1, X_2, X_3; Y|W), \tag{B.37d}
\end{align*}
\]

which can be written as follows

\[
H(S_1, S_2|W) < \min \left\{ I(X_1, X_2, X_3; Y|W), \\
I(X_1, X_2, X_3; Y|W, V_1) + I(X_1, X_3; Y|S_1, X_2, V_2), \\
I(X_1, X_2, X_3; Y|W, V_2) + I(X_2, X_3; Y|S_2, X_1, V_1), \\
I(X_1, X_2; Y|V_1, V_2, X_3, W) + \rho \right\}. \tag{B.38}
\]

Combining constraints (B.35), (B.36) and (B.38) yield conditions (3d)–(3f).

**APPENDIX C**

**BOUNDING** \( E_{x_2, v_1, v_2, x_3, y} \{ \| L_1(b) \| \mathcal{D}, (E_1)^c \} \)

We now bound \( E_{x_2, v_1, v_2, x_3, y} \{ \| L_1(b) \| \mathcal{D}, (E_1)^c \} \). Let the function \( \varphi_b(\hat{s}_{1,b}) \) be defined as follows

\[
\varphi_b(\hat{s}_{1,b}) = \begin{cases}
1, & (\hat{s}_{1,b}, s_{2,b}, x_1(\hat{s}_{1,b}, u_{1,b-1}), x_2(s_{2,b}, u_{2,b-1}), v_1(u_{1,b-1}), v_2(u_{2,b-1}), x_3(u_{1,b-1}, u_{2,b-1}), w_b, y_b) \in A_c^*(n), \\
0, & \text{otherwise.}
\end{cases}
\]
Hence \( \|L_1(b)\| = \sum_{\hat{s}_{1,b} \in S^n} \varphi_b(\hat{s}_{1,b}) \). Hence

\[
E_{x_2, v_1, v_2, x_3, y} \left\{ \|L_1(b)\| \bigg| \mathcal{F}_t, (E_1)^c \right\}
= \sum_{\hat{s}_{1,b} \in A_1^{(n)}(S_1|s_{2,b}, w_b)} E_{x_2, v_1, v_2, x_3, y} \left\{ \varphi_b(\hat{s}_{1,b}) \bigg| (E_1)^c \right\}
= 1 + \sum_{\hat{s}_{1,b} \neq s_{1,b}} \Pr \left\{ (x_1, x_2, v_1, v_2, x_3, y) \in A_1^{(n)}(X_1, X_2, V_1, V_2, X_3, Y|\hat{s}_{1,b}, s_{2,b}, w_b) \bigg| (E_1)^c \right\}
= 1 + \sum_{\hat{s}_{1,b} \neq s_{1,b}} \sum_{(x_1, x_2, v_1, v_2, x_3, y) \in A_1^{(n)}(X_1, X_2, V_1, V_2, X_3, Y|\hat{s}_{1,b}, s_{2,b}, w_b)} \Pr \left\{ (x_1, x_2, v_1, v_2, x_3, y) \bigg| \hat{s}_{1,b}, s_{2,b}, w_b \right\}
\leq 1 + 2^n[H(S_1|S_2, W) - \epsilon_0] \cdot 2^n[H(X_1, V_1, V_2, X_3|S_1, S_2, W) - \epsilon_0] \cdot 2^{-n[H(Y|S_1, V_1, V_2, X_3) + \epsilon_0]} \cdot 2^{-n[H(V_1, V_2, X_3|S_1, S_2, W) + \epsilon_0]}
= 1 + 2^n[H(S_1|S_2, W) - I(S_2, W; X_1|S_1, V_1, V_2, X_3) - I(S_1, X_2, Y|S_2, W, V_1, V_2, X_3) - 5\epsilon_0]
= 1 + 2^n[H(S_1|S_2, W) - I(X_1; Y|S_2, X_1, X_2, X_3, W) - 5\epsilon_0].
\]

where (C.1a) follows from the AEP; and (C.1b) follows from the fact that \( I(S_2, W; X_1|S_1, V_1, V_2, X_3) = 0 \) and from the Markov chain \( V_2 - (S_2, X_2, V_1, X_3, W) - Y \). Therefore we have that

\[
E_{x_2, v_1, v_2, x_3, y} \left\{ \|L_1(b)\| \bigg| \mathcal{F}_t, (E_1)^c \right\} \leq 1 + 2^n[H(S_1|S_2, W) - I(X_1; Y|S_2, X_2, V_1, X_3, W) - 5\epsilon_0]. \tag{C.2}
\]

\section*{APPENDIX D}

**Bounding \( E_{x_2, v_2, x_3, y} \left\{ \|L_1^{(n)}(s_1 \neq s_{1,b}, u_1 \neq u_{1,b-1})\| \bigg| \mathcal{F}_t, (E_1)^c \right\} \)**

We now bound \( E_{x_2, v_2, x_3, y} \left\{ \|L_1^{(n)}(s_1 \neq s_{1,b}, u_1 \neq u_{1,b-1})\| \bigg| \mathcal{F}_t, (E_1)^c \right\} \). Let the function \( \varphi_{1,b}(s_1, u_1) \) be defined as follows

\[
\varphi_{1,b}(s_1, u_1) = \begin{cases} 
1, & (s_1, s_{2,b}, x_1(s_1, u_1), x_2(s_{2,b}, u_{2,b-1}), \\
\nu_1(u_1), \nu_2(u_{2,b-1}), \nu_3(u_1, u_{2,b-1}), w_b, y_b) \in A_1^{(n)}.
\end{cases}
\]

Therefore

\[
\|L_1^{(n)}(s_1 \neq s_{1,b}, u_1 \neq u_{1,b-1})\| = \sum_{\hat{s}_{1,b} \in S^n} \sum_{\hat{u}_{1,b-1}=1} 2^n \varphi_b(\hat{s}_{1,b}, \hat{u}_{1,b-1})
\]

Hence
\[ \mathbb{E}_{x_2, v_2, y} \left\{ \left\| L^*_1(s_1 \neq s_1, b, u_1 \neq u_1, b-1) \right\| \left| \mathcal{D}, (E_1)^c \right\} \right\} \\
= \sum_{s_{1, \hat{b}} \in A^{(n)}_1(s_1 | s_2, b, w_b)} 2^{n R_1} \sum_{\hat{s}_{1, \hat{b}} \neq s_{1, \hat{b}}} \mathbb{E}_{x_2, v_2, y} \left\{ \varphi_b(s_1, \hat{s}_{1, \hat{b}}) \right\} \left| \mathcal{D}, (E_1)^c \right\} \\
= \sum_{s_{1, \hat{b}} \in A^{(n)}_1(s_1 | s_2, b, w_b)} 2^{n R_1} \sum_{\hat{s}_{1, \hat{b}} \neq s_{1, \hat{b}}} \mathbb{E}_{x_2, v_2, y} \left\{ \varphi_b(s_1, \hat{s}_{1, \hat{b}}) \right\} \left| \mathcal{D}, (E_1)^c \right\} \\
\Pr \left\{ (x_1, x_2, v_1, v_2, x_3, y) \in A^{(n)}_1(X_1, X_2, V_1, V_2, X_3, Y | \hat{s}_{1, \hat{b}}, s_2, b, w_b) \right\} \left| (E_1)^c \right\} \\
= \sum_{s_{1, \hat{b}} \in A^{(n)}_1(s_1 | s_2, b, w_b)} 2^{n R_1} \sum_{\hat{s}_{1, \hat{b}} \neq s_{1, \hat{b}}} \mathbb{E}_{x_2, v_2, y} \left\{ \varphi_b(s_1, \hat{s}_{1, \hat{b}}) \right\} \left| \mathcal{D}, (E_1)^c \right\} \\
\leq 2^{n[H(S_1 | S_2, W) + R_1 - I(S_2, W; X_1, X_3 | S_1, V_2) - I(S_1, X_1, V_1, X_3; X_2, Y | S_2, W, V_2) + 5\epsilon_0]} \\
= 2^{n[H(S_1 | S_2, W) + R_1 - I(X_1, X_3; Y | S_2, W, X_2, V_2) - 5\epsilon_0].}

Therefore we have that

\[ \mathbb{E}_{x_2, v_2, y} \left\{ \left\| L^*_1(s_1 \neq s_1, b, u_1 \neq u_1, b-1) \right\| \left| \mathcal{D}, (E_1)^c \right\} \right\} \leq 2^{n[H(S_1 | S_2, W) + R_1 - I(X_1, X_3; Y | S_2, W, X_2, V_2) - 5\epsilon_0].} \quad (D.1) \]
Hence

\[
E_{v_2,y} \left\{ \| \mathcal{L}_c^* (s_1 = s_{1,b}, s_2 \neq s_{2,b}, u_1 \neq u_{1,b-1}, u_2 = u_{2,b-1}) \| \right\} \mathcal{D}_c, (E_1)^c \}
\]

= \sum_{s_1 \neq s_{1,b}} \sum_{s_2 \neq s_{2,b}} E_{v_2,y} \left\{ \varphi_6 (s_1, s_2, u_1, u_2, b) \right\} \mathcal{D}_c, (E_1)^c

= \sum_{s_1 \neq s_{1,b}} \sum_{s_2 \neq s_{2,b}} \mathbb{P} \left\{ (x_1, x_2, v_1, v_2, x_3, y) \in A_c^* (X_1, X_2, V_1, V_2, X_3, Y | s_1, s_2, b, w_0) \right\} \mathcal{D}_c, (E_1)^c

\leq 2^n[H(S_1 | S_2, W - \varepsilon_0)] + 2^n[H(V_2 | S_1, S_2, W_{v2}) + \varepsilon_0] + 2^n[H(Y | S_1, W_{v2}) + \varepsilon_0]

\leq 2^n[H(S_1 | S_2, W) + \varepsilon_0] + 2^n[H(V_2 | S_1, S_2, W_{v2}) + \varepsilon_0] + 2^n[H(Y | S_1, W_{v2}) + \varepsilon_0]

\leq 2^n[H(S_1 | S_2, W) + \varepsilon_0] + 2^n[H(V_2 | S_1, S_2, W_{v2}) + \varepsilon_0] + 2^n[H(Y | S_1, W_{v2}) + \varepsilon_0]

Therefore we have that

\[
E_{v_2,y} \left\{ \| \mathcal{L}_c^* (s_1 = s_{1,b}, s_2 \neq s_{2,b}, u_1 \neq u_{1,b-1}, u_2 = u_{2,b-1}) \| \right\} \mathcal{D}_c, (E_1)^c \leq 2^n[H(S_1 | S_2, W) + R_1 - I(X_1, X_2, X_3; Y | S_1, W, V_2) - 5\varepsilon_0] \quad \text{(E.1)}
\]

APPENDIX F

JOINT DISTRIBUTION OF Thm. 1 AND Thm. 2

Lemma F.1. The joint distribution of \( s_{1,b}, s_{2,b}, w_{3,b}, w_0, x_1(s_{1,b}, u_{1,b-1}), x_2(s_{2,b}, u_{2,b-1}), v_1(u_{1,b-1}), v_2(u_{2,b-1}), x_3(u_{1,b-1}, u_{2,b-1}), y_{3,b} \) and \( y_6 \) satisfies (for brevity the block indices are omitted)

\[
pr(s_{1,b}, s_{2,b}, w_{3,b}, w_0, x_1(s_{1,b}, u_{1,b-1}), x_2(s_{2,b}, u_{2,b-1}), v_1(u_{1,b-1}), v_2(u_{2,b-1}), x_3(u_{1,b-1}, u_{2,b-1}), y_{3,b}, y_6) = \prod_{j=1}^{n} p(s_{1,j}, s_{2,j}, u_{3,j}, v_{j}) p(v_{j}) p(x_{1,j} | s_{1,j}, v_{j}) p(x_{2,j} | s_{2,j}, v_{j}) p(x_{3,j} | v_{j}) \times p(y_{3,j}, y_{j} | x_{1,j}, x_{2,j}, x_{3,j}) \quad \text{(F.1)}
\]
Proof: The proof is given in Subsection F-A.

Corollary F.1. The joint distribution of \( s_{1,b}, s_{2,b}, w_{3,b}, x_1(s_{1,b}, u_{1,b-1}), x_2(s_{2,b}, u_{2,b-1}), v_1(u_{1,b-1}), v_2(u_{2,b-1}), x_3(u_{1,b-1}, u_{2,b-1}) \) and \( y_{3,b} \) satisfies (for brevity the block indices are omitted)

\[
p(s_{1,b}, s_{2,b}, w_{3,b}, x_1(s_{1,b}, u_{1,b-1}), x_2(s_{2,b}, u_{2,b-1}), v_1(u_{1,b-1}), v_2(u_{2,b-1}), x_3(u_{1,b-1}, u_{2,b-1}), y_{3,b}) = \prod_{j=1}^{n} p(s_{1,j}, s_{2,j}, w_{3,j}) p(u_{1,j}) p(x_{1,j} | s_{1,j}, v_{1,j}) p(x_{2,j} | s_{2,j}, v_{2,j}) p(x_{3,j} | v_{1,j}, v_{2,j}) \times p(y_{3,j} | x_{1,j}, x_{2,j}, x_{3,j}).
\] (F.2)

Proof: This result is a direct consequence of Lemma F.1.

Corollary F.2. The joint distribution of \( s_{1,b}, s_{2,b}, w_{b}, x_1(s_{1,b}, u_{1,b-1}), x_2(s_{2,b}, u_{2,b-1}), v_1(u_{1,b-1}), v_2(u_{2,b-1}), x_3(u_{1,b-1}, u_{2,b-1}) \) and \( y_{b} \) satisfies (for brevity the block indices are omitted)

\[
p(s_{1,b}, s_{2,b}, w_{b}, x_1(s_{1,b}, u_{1,b-1}), x_2(s_{2,b}, u_{2,b-1}), v_1(u_{1,b-1}), v_2(u_{2,b-1}), x_3(u_{1,b-1}, u_{2,b-1}), y_{b}) = \prod_{j=1}^{n} p(s_{1,j}, s_{2,j}, w_{j}) p(u_{1,j}) p(x_{1,j} | s_{1,j}, v_{1,j}) p(x_{2,j} | s_{2,j}, v_{2,j}) p(x_{3,j} | v_{1,j}, v_{2,j}) \times p(y_{b,j} | x_{1,j}, x_{2,j}, x_{3,j}).
\] (F.3)

Proof: This result is a direct consequence of Lemma F.1.
A. Proof of Lemma F.1

For brevity we omit the block indices.

\[
p(s_{1,b}, s_{2,b}, w_{3,b}, w_6, x_1(s_{1,b}, u_{1,b}^{-1}), x_2(s_{2,b}, u_{2,b}^{-1}), v_1(u_{1,b}^{-1}), v_2(u_{2,b}^{-1}), x_3(u_{1,b}^{-1}, u_{2,b}^{-1}), y_{3,b}, y_b)
\]

\[
= \prod_{j=1}^n p(s_{1,j}, s_{2,j}, w_{3,j}, w_j, x_{1,j}, x_{2,j}, v_{1,j}, v_{2,j}, x_{3,j}, y_{3,j}, y_j)
\]

\[
\quad \quad \quad \quad \left| s_{1,j}^{-1}, s_{2,j}^{-1}, w_{3,j}^{-1}, w_j^{-1}, x_{1,j}^{-1}, x_{2,j}^{-1}, v_{1,j}^{-1}, v_{2,j}^{-1}, x_{3,j}^{-1}, y_{3,j}^{-1}, y_j^{-1} \right|
\]

\[
= \prod_{j=1}^n p(y_{3,j}, y_j | s_{1,j}^j, s_{2,j}^j, w_{3,j}^j, w_j^j, x_{1,j}^j, x_{2,j}^j, v_{1,j}^j, v_{2,j}^j, x_{3,j}^j, y_{3,j}^j, y_j^j) \times
\]

\[
p(s_{1,j}, s_{2,j}, w_{3,j}, w_j, x_{1,j}, x_{2,j}, v_{1,j}, v_{2,j}, x_{3,j} | s_{1,j}^{-1}, s_{2,j}^{-1}, w_{3,j}^{-1}, w_j^{-1}, x_{1,j}^{-1}, x_{2,j}^{-1}, v_{1,j}^{-1}, v_{2,j}^{-1}, x_{3,j}^{-1}, y_{3,j}^{-1}, y_j^{-1})
\]

\[
= \prod_{j=1}^n p(y_{3,j}, y_j | x_{1,j}, x_{2,j}, x_{3,j}) \times
\]

\[
p(s_{1,j}, s_{2,j}, w_{3,j}, w_j, x_{1,j}, x_{2,j}, v_{1,j}, v_{2,j}, x_{3,j} | s_{1,j}^{-1}, s_{2,j}^{-1}, w_{3,j}^{-1}, w_j^{-1}, x_{1,j}^{-1}, x_{2,j}^{-1}, v_{1,j}^{-1}, v_{2,j}^{-1}, x_{3,j}^{-1}, y_{3,j}^{-1}, y_j^{-1})
\]

\[
= \prod_{j=1}^n p(y_{3,j}, y_j | x_{1,j}, x_{2,j}, x_{3,j}) \times
\]

\[
p(x_{3,j} | s_{1,j}^j, s_{2,j}^j, w_{3,j}^j, w_j^j, x_{1,j}^j, x_{2,j}^j, v_{1,j}^j, v_{2,j}^j, x_{3,j}^j, y_{3,j}^j, y_j^j)
\]

\[
= \prod_{j=1}^n p(y_{3,j}, y_j | x_{1,j}, x_{2,j}, x_{3,j}) \times
\]

\[
p(x_{3,j} | s_{1,j}^j, s_{2,j}^j, w_{3,j}^j, w_j^j, x_{1,j}^j, x_{2,j}^j, v_{1,j}^j, v_{2,j}^j, x_{3,j}^j, y_{3,j}^j)
\]

\[
= \prod_{j=1}^n p(y_{3,j}, y_j | x_{1,j}, x_{2,j}, x_{3,j}) \times
\]

\[
p(x_{3,j} | v_{1,j}^j, v_{2,j}^j, x_{3,j}^j)
\]

\[
= \prod_{j=1}^n p(y_{3,j}, y_j | x_{1,j}, x_{2,j}, x_{3,j}) \times
\]

\[
p(x_{3,j} | v_{1,j}^j, v_{2,j}^j, x_{3,j}^j, v_{3,j}^{-1})
\]

\[
= \prod_{j=1}^n p(y_{3,j}, y_j | x_{1,j}, x_{2,j}, x_{3,j}) \times
\]

\[
p(x_{3,j} | v_{1,j}^j, v_{2,j}^j, x_{3,j}^j)
\]

\[
= \prod_{j=1}^n p(y_{3,j}, y_j | x_{1,j}, x_{2,j}, x_{3,j}) \times
\]

\[
p(x_{3,j} | v_{1,j}^j, v_{2,j}^j, x_{3,j}^j, v_{3,j}^{-1})
\]

\[
= \prod_{j=1}^n p(y_{3,j}, y_j | x_{1,j}, x_{2,j}, x_{3,j}) \times
\]

\[
p(x_{3,j} | v_{1,j}^j, v_{2,j}^j, x_{3,j}^j, v_{3,j}^{-1})
\]

\[
= \prod_{j=1}^n p(y_{3,j}, y_j | x_{1,j}, x_{2,j}, x_{3,j}) \times
\]

\[
p(x_{3,j} | v_{1,j}^j, v_{2,j}^j, x_{3,j}^j, v_{3,j}^{-1})
\]

\[
= \prod_{j=1}^n p(y_{3,j}, y_j | x_{1,j}, x_{2,j}, x_{3,j}) \times
\]

\[
p(x_{3,j} | v_{1,j}^j, v_{2,j}^j, x_{3,j}^j, v_{3,j}^{-1})
\]
\[= \prod_{j=1}^{n} p(y_{3,j}, y_j | x_{1,j}, x_{2,j}, x_{3,j}) p(s_{1,j}, s_{2,j}, w_3,j, w_j | s_{1,1}^{j-1}, s_{2,1}^{j-1}, w_3^{j-1}, w_1^{j-1}) \times \]

\[p(v_{1,j}, v_{2,j} | s_1^{j-1}, s_2^{j-1}, w_{3,1}, w_1^{j-1}, x_1^{j-1}, x_2^{j-1}, v_1^{j-1}, v_2^{j-1}, x_3^{j-1}) \times \]

\[p(x_{1,j}, x_{2,j} | s_1^{j-1}, s_2^{j-1}, w_{3,1}, w_1^{j-1}, x_1^{j-1}, x_2^{j-1}, v_1^{j-1}, v_2^{j-1}, x_3^{j-1}) p(x_{3,j} | v_{1,1}, v_{2,1}, x_{3,1}^{j-1}) \]

\[= \prod_{j=1}^{n} p(y_{3,j}, y_j | x_{1,j}, x_{2,j}, x_{3,j}) p(s_{1,j}, s_{2,j}, w_3,j, w_j | s_{1,1}^{j-1}, s_{2,1}^{j-1}, w_3^{j-1}, w_1^{j-1}) \times \]

\[p(v_{1,j}, v_{2,j} | v_{1,1}, v_{2,1}) p(x_{1,j} | s_{1,1}^{j-1}, x_1^{j-1}, v_{1,1}) p(x_{2,j} | s_{2,1}^{j-1}, x_2^{j-1}, v_{2,1}) p(x_{3,j} | v_{1,1}, v_{2,1}, x_{3,1}^{j-1}) \]

\[= \prod_{j=1}^{n} p(y_{3,j}, y_j | x_{1,j}, x_{2,j}, x_{3,j}) p(s_{1,j}, s_{2,j}, w_3,j, w_j) \times \]

\[p(v_{1,j}, p(v_{2,j}) p(x_{1,j} | s_{1,j}, v_{1,j}) p(x_{2,j} | s_{2,j}, v_{2,j}) p(x_{3,j} | v_{1,j}, v_{2,j}) \]

where (a) holds because there is no feedback from the destination to the other nodes and due to causality memorylessness of the channel [3]; (b) holds because \( s_{1,j}, s_{2,j}, w_3,j, w_j, x_{1,j}, x_{2,j}, v_{1,j}, v_{2,j} \) given the history \( s_{1,1}^{j-1}, s_{2,1}^{j-1}, w_3^{j-1}, w_1^{j-1}, x_1^{j-1}, x_2^{j-1}, v_1^{j-1}, v_2^{j-1}, x_3^{j-1} \) are independent of \( y_3^{j-1} \), since there is no feedback from the relay to the sources; (c) follows from the construction of \( x_3(u_{1,b-1}, u_{2,b-1}) \); (d) follows from the input distribution of the sources and side information; (e) follows from the construction of \( v_1(u_{1,b-1}), v_2(u_{2,b-1}), x_1(s_{1,b}, u_{1,b-1}) \) and \( x_2(s_{2,b}, u_{2,b-1}) \); and (f) holds because all the sources and the codewords are generated in an i.i.d manner so their \( j \)'th symbol does not depend on the past.

REFERENCES

