State-Dependent Z Channel

S. Hajizadeh, M. Monemizadeh, E. Bahmani, G. A. Hodtani1, and M. Joneidi2

1 Department of Electrical Engineering, Ferdowsi University of Mashhad, Mashhad, Iran
Saeed.hajizadeh1367@gmail.com; ghodtani@gmail.com; mostafamonemizadeh@gmail.com; bahmani.e@gmail.com
2 Department of Electrical Engineering, Sharif University of Technology, Tehran, Iran
joneidi@ee.sharif.ir

Abstract—In this paper we study the “Z” channel with side information non-causally available at the encoders. We use the notion of Han-Kobayashi rate splitting along with Gelf'and-Pinsker random binning scheme and Chong-Motani-Garg-El Gamal (CMGE) jointly decoding to find the achievable rate region. We will see that our achievable rate region gives the achievable rate of the multiple access channel with side information and also degraded broadcast channel with side information. We will also derive an inner bound and an outer bound on the capacity region of the state-dependent degraded discrete memoryless Z channel. We will also see that using Costa dirty paper coding, we can remove the negative effect of the interference from the direction of one transmitter-receiver pair. Also, by assuming the high signal to noise ratio and strong interference regime, and using the lattice strategies, we derive an achievable rate region for the Gaussian degraded “Z” channel with additive interference non-causally available at both of the encoders. Our method is based on lattice transmission scheme, jointly decoding at the first decoder and successive decoding at the second decoder. Using such coding scheme we remove the effect of the interference completely.

Keywords: Z channel; Side information; rate splitting; Dirty paper coding; Lattice strategies

I. INTRODUCTION

The Z channel is a two-transmitter two-receiver model shown in Fig. 1 where the first sender only wishes to send information to the first receiver whereas the second transmitter sends information to both of the receivers. The Z channel was first studied by Viswanath et al. [1] where they introduced the model and found the capacity region of a special class of Z channels and the achievable rate of a special case of the Gaussian Z channel (GZC). In [2], Liu and Ulukus obtained several capacity bounds for a class of GZC. Chong-Motani-Garg (CMGE) [3] studied three different types of degraded Z channel and characterized the capacity region in one type. They also characterized the capacity region of GZC with moderately strong crossover link.

The capacity region of the general Z channel is still an open problem. The best achievable rate region for the discrete memoryless Z channel until today is due to Do et al. [4].

Channels with side information were first studied by Shannon [5] where he characterized the capacity of a point-to-point channel with side information causally available at the transmitters. Gelf'and and Pinsker [6] found the capacity of a single-user channel with side information non-causally available at the encoders.

In this paper we study the Z channel with channel state information non-causally available at the encoders that is depicted in Fig. 2. Here the first transmitter sends \( m_{11} \in [1:2^{R_{11}}] \) to \( Y_1 \) while the second transmitter first splits its messages, \( m_{21} \in [1:2^{R_{21}}] \) and \( m_{22} \in [1:2^{R_{22}}] \), to two independent parts; i.e. \( M_{21} = (M_{21}^c, M_{21}^p) \) and \( M_{22} = (M_{22}^c, M_{22}^p) \) with rates \( R_{22} = R_{22}^c + R_{22}^p \) and \( R_{22} = R_{22}^c + R_{22}^p \) respectively, and then encodes its messages to send through the channel. The channel state information is non-causally available at the transmitters. The messages \( (M_{21}^k, M_{22}^k) \) can be decoded by both receivers, while \( M_{2k}^c \) is decoded by its respective receiver, \( k = 1,2 \).

We propose an achievable rate region using the lattice based coding for the Gaussian degraded “Z” channel with additive interference non-causally available at both of the encoders under high-SNR and strong interference regime. Our method is based on lattice transmission scheme, jointly decoding at the first decoder and successive decoding at the second decoder. Using such coding scheme we remove the effect of the interference completely.
The rest of the paper is as follows. In section II, definitions are provided. In section III, we derive an achievable rate region for the general discrete memoryless Z channel with side information non-causally available at the encoders. In section IV, we derive an inner and an outer bound on the capacity region of degraded discrete memoryless Z channel and show that using dirty paper coding, we can remove the negative effect of the interference in the direction of one transmitter-receiver pair in the derived inner bound. In section V, we derive an achievable rate region for the Gaussian degraded “Z” channel with additive interference non-causally available at both of the encoders using lattice strategies and show that using lattice strategy we can completely remove the interference. The conclusion is given in section VI.

II. DEFINITIONS

The discrete memoryless “Z” channel with channel state information non-causally available at the transmitter, depicted in Fig. 2, consists of five finite sets \( \mathcal{S}, \mathcal{X}_1, \mathcal{X}_2, \mathcal{Y}_1, \mathcal{Y}_2 \) and two marginal probability distributions \( p(y_2|x_2, s) \) and \( p(y_1|x_1, x_2, s) \). The memorylessness nature of the channel imposes the following additional constraint on the channel transition probability

\[
p(y^n_1, y^n_2 | x^n_1, x^n_2, s^n) = \prod_{i=1}^n p(y_{2i}|x_{2i}, s_i)p(y_{1i}|x_{1i}, x_{2i}, s_i) \quad (1)
\]

A \( (2^{nR_1}, 2^{nR_{21}}, 2^{nR_{22}}, n, \epsilon) \) code for the discrete memoryless Z channel with side information consists of two sets of encoding mappings

\[
e_1: \{1,2,\ldots, 2^{nR_1}\} \times \mathcal{S} \rightarrow \mathcal{X}_1^n
\]

\[
e_2: \{1,2,\ldots, 2^{nR_{21}}\} \times \{1,2,\ldots, 2^{nR_{22}}\} \times \mathcal{S} \rightarrow \mathcal{X}_2^n
\]

and two sets of decoding mappings

\[
d_1: \mathcal{Y}_1^n \rightarrow \{1,2,\ldots, 2^{nR_1}\} \times \{1,2,\ldots, 2^{nR_{21}}\}
\]

\[
d_2: \mathcal{Y}_2^n \rightarrow \{1,2,\ldots, 2^{nR_{22}}\}
\]

and an average probability of error defined as the probability that the decoded message does not equal the transmitted message such that

\[
p(d_1(y^n_1) \neq (m_{11}, m_{21})) \text{ or } d_2(y^n_2) \neq m_{22}) \leq \epsilon
\]

where the messages are assumed to be uniformly distributed on their respective sets.

A rate triple \((R_{11}, R_{21}, R_{22})\) is said to be achievable for the discrete memoryless “Z” channel with side information if there exists a sequence of \((2^{nR_1}, 2^{nR_{21}}, 2^{nR_{22}}, n, \epsilon)\) codes.

III. THE MAIN RESULT

In this section, we derive an achievable rate region for the general Z channel with side information. At the first transmitter we apply the Gel’fand-Pinsker random binning and at the second transmitter we use a combination of superposition coding, Han-Kobayashi [7] rate splitting, Gel’fand-Pinsker coding, and CMGE [8] jointly decoding.

**Definition 1:** Define \( \mathcal{P}_{\text{ZCSI}} \) as the set of all random variables \((S, W, X_1, U, U_1, U_2, X_2, Y_1, Y_2)\) such that

\[
p(s, w, x_1, u, u_1, u_2, x_2, y_1, y_2) = p(s)p(w|s)p(x_1|w,s)p(u_1|u, s)p(u_2|u, s). \quad (2)
\]

\[
p(x_2|u_1, u_1, u_2, s)p(y_1, y_2|x_1, x_2, s)
\]

where \((p(x_1|w, s), p(x_2|u_1, u_1, u_2, s)) \in (0,1)^2\)

**Theorem 1:** An achievable rate region for the discrete memoryless Z channel with side information non-causally available at the transmitters, depicted in Fig. 2, is the closure of the convex hull of the set \( \mathcal{R}_{\text{ZCSI}} = \bigcup_{p \in \mathcal{P}_{\text{ZCSI}}} \mathcal{R}_{\text{ZCSI}}(p) \) where

\[
\mathcal{R}_{\text{ZCSI}}(p) = \{(R_{11}, R_{21}, R_{22}) : R_{11} \leq A \quad (3-1)
\]

\[
R_{21} + R_{22} \leq \min\{B, C\} \quad (3-2)
\]

\[
R_{11} + R_{21} + R_{22} \leq \min\{D, E\} \quad (3-3)
\]

for some \((S, W, X_1, U, U_1, U_2, X_2, Y_1, Y_2) \in \mathcal{P}_{\text{ZCSI}}\)

where we have

\[
A = I(W; Y_1|U_1) - I(W; S|U_1)
\]

\[
B = I(U_1; Y_1|WU) + I(U_2; Y_2) - I(U_1; S|W) - I(UU_2; S)
\]

\[
C = I(UU_1; Y_1) + I(U_2; Y_2|U) - I(UU_1; S|W) - I(U_2; S|U)
\]

\[
D = I(WUU_1; Y_1) + I(U_2; Y_2|U) - I(WUU_1; S) - I(U_2; SU|U)
\]

\[
E = I(WUU_1; Y_1|U) + I(U_2; Y_2) - I(WUU_1; S|U) - I(U_2; SU|S)
\]

**Corollary 1.1:** If we put \( S \equiv \emptyset \) in (3), then we have the achievable rate of the Z channel provided by [3].

**Corollary 1.2:** If we let no information to be sent to the second receiver, we obtain the achievable rate for the state-dependent multiple access channel with independent sources, i.e. if we set \( R_{21} = R_{22} = 0, R_{11} = R, R_{21} = R_2, U = U_2 = \emptyset, V_1 = U_2, W = U_1, \) and \( Y_1 = Y \) in (15 – 21), we obtain the closure of the convex hull of the set of all rate pairs \((R, R_2)\) satisfying

\[
R_1 \leq I(U_1; Y_1|U_2) - I(U_1; S|U_2) \quad (22)
\]

\[
R_2 \leq I(U_2; Y_1|U_2) - I(U_2; S|U_2) \quad (23)
\]

\[
R_1 + R_2 \leq I(U_1, U_2; Y_1) - I(U_1, U_2; S) \quad (24)
\]

**Corollary 1.3:** If we set \( R_{11} = R_{21} = R_{22} = 0, R_{21} = R_1, \) and \( W = U_2 = \emptyset \) in (15 – 21), and assuming that the receiver \( Y_2 \) is a degraded version of \( Y_1 \), we obtain the achievable rate region of the degraded broadcast channel with side information provided in [9].

\[
R_1 \leq I(U_1; Y_1|U) - I(U_1; S|U) \quad (25)
\]

\[
R_2 \leq I(U_2; Y_2) - (U; S) \quad (26)
\]

**Proof:** Fix a distribution of the form

\[
p(s)p(w|s)p(x_1|w, s)p(u_1|u, s)p(u_2|u, s).
\]

The second transmitter splits its messages as mentioned in section I. We then generate the codebook as follows

Randomly and independently generate \( 2^n(m_{11}, m_{12}) \) sequences \( w^n(m_{11}, m_{12}) \) each one i.i.d according to \( \prod_{i=1}^n p(w_i) \) and randomly partition them into \( 2^{nR_{11}} \) bins.
Randomly and independently generate $2^n(R_1^n + R_2^n + R_3^n)$ sequences $u^n(m_{21}, m_{22}^{(2)}, m_{21}^{(2)}, m_{22}^{(1)})$ each one i.i.d according to $\prod_{i=1}^n p(u_i)$ and randomly partition them into $2^n(R_1^n + R_2^n)$ bins.

For each pair $(m_{11}, m_{22}^{(2)})$, independently generate $2^n(R_2^n + R_3^n)$ sequences $u_k^n(m_{21}, m_{22}^{(2)}, m_{21}^{(2)}, m_{22}^{(1)}, m_{22}^{(2)}, k = 1, 2$, each one i.i.d according to $\prod_{i=1}^n p(u_k(u_i|u_j)$ and randomly partition them into $2^nR_2^n$ bins.

**Encoding:** Assume that the transmitters want to send the triple $(m_{11}, m_{21}, m_{22}^{(2)})$ with $m_{22}^{(2)} = (m_{21}^{(2)}, m_{22}^{(2)})$.

TX1 looks in bin $m_{11}$ to find some $m_{11}^{(2)}$ such that $(w^n(m_{11}, m_{11}^{(2)}), s_1^n)$ are jointly typical. Assume that the chosen index is $M_{11}^{(2)}$.

TX2, in the meantime, looks in bin $(m_{11}, m_{22}^{(2)})$ to find some pair $(m_{21}, m_{22}^{(2)}, m_{21}^{(2)}, m_{22}^{(2)})$ such that the pair $(w^n(m_{21}, m_{22}^{(2)}, m_{21}^{(2)}, m_{22}^{(2)}), s_2^n)$ is jointly typical. Assume that the chosen pair is $(M_{21}, M_{22}^{(2)})$.

TX2 then looks in bin $m_{22}^{(2)}$ to find some $m_{22}^{(2)}$ such that the pair $(u_k^n(m_{21}, m_{22}^{(2)}, m_{21}^{(2)}, m_{22}^{(2)}, m_{22}^{(2)}, k) = 1, 2$, is conditionally jointly typical. Assume that the indices chosen are $M_{2k}^{(2)}$, $k = 1, 2$.

TX1 and TX2 then send $x_{11} = x_{11}(w_i, s_i)$

$x_{21} = x_{21}(u_i u_j, u_{12}, s_i)$

**Decoding:** Without loss of generality, assume that the triple $(1, 1, 1)$ was sent through the channel. The first receiver receives $y_1^n$ and looks for the unique pair $(m_{11}, m_{22}^{(2)})$ such that

$w^n(m_{11}, M_{11}^{(2)}), u^n(m_{21}, m_{22}^{(2)}, M_{21}^{(2)}), u_1^n(m_{21}, m_{22}^{(2)}, M_{11}^{(2)}, M_{22}^{(2)}, M_{21}^{(2)}, y_1^n) \in A_n^{(e)}$

where $A_n^{(e)}$ is the set of jointly typical sequences.

The second receiver, meanwhile, receives $y_2^n$ looks for the unique message index $m_{22}^{(2)}$ such that

$u^n(m_{21}, m_{22}^{(2)}, M_{21}^{(2)}, M_{22}^{(2)}), u_2^n(m_{21}, m_{22}^{(2)}, M_{11}^{(2)}, M_{21}^{(2)}, M_{22}^{(2)}, M_{21}^{(2)}, y_2^n) \in A_n^{(e)}$

We define the following error events for the encoding section

$E_1^{enc} = \{(w^n(m_{11}, M_{11}^{(2)}), s^n) \in A_n^{(e)} \text{ for all } M_{11}^{(2)} \in [1: 2^nR_1^n]\}$

$E_2^{enc} = \{(u^n(m_{21}, m_{22}^{(2)}, M_{21}^{(2)}, M_{22}^{(2)}), s^n) \in A_n^{(e)} \text{ for all } (M_{21}^{(2)}, M_{22}^{(2)}) \in [1: 2^nR_2^n] \times [1: 2^nR_2^n]\}$

$E_k^{enc} = \{(u_k^n(m_{21}, m_{22}^{(2)}, m_{21}^{(2)}, m_{22}^{(2)}), s^n) \in A_n^{(e)} \text{ for all } m_{2k}^{(2)} \in [1: 2^nR_2^n], k = 1, 2, \}$

The decoding error events for the first receiver are defined as follows

$E_{11}^{dec} = \{(w^n(1, M_{11}^{(2)}), u^n(1, 1, M_{21}^{(2)}, M_{22}^{(2)}), u_1^n(1, M_{11}^{(2)}, M_{21}^{(2)}, M_{22}^{(2)}), y_1^n) \in A_n^{(e)}\}$

$E_{12}^{dec} = \{(w^n(m_{11}, m_{11}^{(2)}), u^n(m_{21}, m_{22}^{(2)}, m_{21}^{(2)}, M_{22}^{(2)}), u_1^n(1, M_{11}^{(2)}, m_{21}^{(2)}, m_{22}^{(2)}, m_{21}^{(2)}, M_{22}^{(2)}, y_1^n) \in A_n^{(e)} \text{ for some } m_{21}^{(2)}, m_{22}^{(2)}, m_{21}^{(2)}, m_{22}^{(2)} \neq (1, 1, 1, 1)\}$

$E_{13}^{dec} = \{(w^n(1, M_{11}^{(2)}), u^n(m_{21}, m_{22}^{(2)}, m_{21}^{(2)}, M_{22}^{(2)}), u_1^n(1, M_{21}^{(2)}, m_{21}^{(2)}, m_{22}^{(2)}, m_{21}^{(2)}, M_{22}^{(2)}, M_{22}^{(2)}, y_1^n) \in A_n^{(e)} \text{ for some } m_{21}^{(2)}, m_{22}^{(2)}, m_{21}^{(2)}, m_{22}^{(2)} \neq (1, 1, 1, 1)\}$

$E_{14}^{dec} = \{(w^n(m_{11}, m_{11}^{(2)}), u^n(1, 1, M_{11}^{(2)}, M_{22}^{(2)}), u_1^n(1, 1, M_{11}^{(2)}, M_{21}^{(2)}, M_{22}^{(2)}, M_{21}^{(2)}, M_{22}^{(2)}, y_1^n) \in A_n^{(e)} \text{ for some } m_{21}^{(2)} \neq 1 \text{ and some } m_{22}^{(2)} \neq M_{22}^{(2)}\}$

$E_{15}^{dec} = \{(w^n(1, M_{11}^{(2)}), u^n(1, 1, M_{21}^{(2)}, M_{22}^{(2)}), u_1^n(1, 1, M_{21}^{(2)}, M_{21}^{(2)}, M_{22}^{(2)}, M_{22}^{(2)}, y_1^n) \in A_n^{(e)} \text{ for some } m_{21}^{(2)} \neq 1 \text{ and some } m_{22}^{(2)} \neq M_{22}^{(2)}\}$

The decoding error events for the second receiver are as follows

$E_{21}^{dec} = \{(u^n(1, 1, M_{11}^{(2)}, M_{22}^{(2)}), u_2^n(1, 1, M_{11}^{(2)}, M_{21}^{(2)}, M_{22}^{(2)}, y_2^n) \in A_n^{(e)}\}$

$E_{22}^{dec} = \{(u^n(m_{21}, m_{22}^{(2)}, M_{21}^{(2)}, M_{22}^{(2)}), u_2^n(1, M_{21}^{(2)}, m_{21}^{(2)}, m_{22}^{(2)}, m_{22}^{(2)}, y_2^n) \in A_n^{(e)} \text{ for some } m_{21}^{(2)} \neq 1 \text{ and some } m_{22}^{(2)} \neq M_{22}^{(2)}\}$

$E_{23}^{dec} = \{(u^n(1, 1, M_{21}^{(2)}, M_{22}^{(2)}), u_2^n(1, 1, M_{21}^{(2)}, M_{22}^{(2)}, M_{22}^{(2)}, y_2^n) \in A_n^{(e)} \text{ for some } m_{21}^{(2)} \neq 1 \text{ and some } m_{22}^{(2)} \neq M_{22}^{(2)}\}$

Now we bound the probability of encoding error

$p(E_1^{enc}) = p\left(\{(w^n(m_{11}, m_{11}^{(2)}), s^n) \in A_n^{(e)} \text{ for all } m_{11} \in [1: 2^nR_1^n]\}\right) = \prod_{m_{11}=1}^{2^nR_1^n} p\left((w^n(m_{11}, m_{11}^{(2)}), s^n) \in A_n^{(e)}\right) = 1 - p\left((w^n(m_{11}, m_{11}^{(2)}), s^n) \in A_n^{(e)}\right) \leq (1 - 2^{-n(R_1 + R_3 + \delta_1)})^{2^nR_1^n}$
\[ 1 - 2^{-n(I(W;S) + \delta_1(\epsilon)) - \kappa_1} \leq \frac{\epsilon}{12} \]

provided that
\[ \bar{R}_{11} \geq I(W;S) + \delta_1(\epsilon) \] (4)

In the same way we can prove that
\[ p(E_{11}^{enc}) + p(E_{12}^{enc}) + p(E_{21}^{enc}) \leq \frac{\epsilon}{4} \]

provided that
\[ \bar{R}_{12} + \bar{R}_{22} \geq I(U;S) + \delta_2(\epsilon) \]
\[ \bar{R}_{21} \geq I(U_1;S|U) + \delta_3(\epsilon) \]
\[ \bar{R}_{21} \geq I(U_2;S|U) + \delta_4(\epsilon) \] (5) (6) (7)

The probability of error at the first receiver is bound as follows
\[ p(E_{11}^{dec}) \leq \frac{\epsilon}{12} \]

by the weak law of large numbers.

\[ p(E_{11}^{dec}) = \sum_{(n_1,n_2,n_3,n_4) \in \{(1,1,1,1),(1,1,1,2),(1,1,2,1),(1,2,1,1),(2,1,1,1)\}} \sum \left( w_u w_v \gamma_1 \gamma_2 \gamma_3 \gamma_4 \right) \frac{p(u^n|w^n)}{p(v^n|w^n)} p(u^n|w^n u^n) p(u^n|w^n u^n) \]

\[ \leq 2^{|R_{11} + R_{12} + R_{22} + R_{21} + R_{21}^p + R_{22}^p + R_{21}^p + R_{22}^p + R_{21}^p + R_{21}^p|} 2^n(h(WU_{11}Y_1) + \epsilon), \]

\[ 2^{-n(H(W)-\epsilon)} 2^{-n(H(W)|W)-2\epsilon} 2^{-n(H(U)|W)-2\epsilon} 2^{-n(H(Y)|Y)-\epsilon} \]

\[ = 2^{-n(I(WUU_1Y_1) + I(W;U) + I(W;U_1|U)) - \sum R_i - 7\epsilon} \leq \frac{\epsilon}{12} \]

provided that
\[ R_{11} + R_{12} + R_{22} + R_{21} + R_{22}^p + R_{21}^p + R_{21}^p \leq I(WUU_1Y_1) + I(W;U) + I(W;U_1|U) - 7\epsilon \] (8)

where

\[ \sum R_i = R_{11} + R_{11} + R_{22} + R_{22} + R_{22} + R_{22} + R_{21} + R_{21} + R_{21}^p \]

In just the same way one can prove that
\[ \sum_{k=3}^{6} p(E_{11}^{dec}) \leq \frac{\epsilon}{3} \]

provided that
\[ R_{21} + R_{22} + R_{22}^p + R_{21}^p + R_{21}^p \leq I(UUU_1Y_1|W) + I(W;U) + I(W;U_1|U) \]
\[ + I(W;U) + I(W;U_1|U) \] (9)

\[ R_{11} + R_{11} + R_{21}^p + R_{21}^p \leq I(WU_1Y_1|U) \]
\[ + I(W;U) + I(W;U_1|U) \] (10)

\[ R_{21} + R_{21}^p \leq I(U_1Y_1|U) \]
\[ + I(W;U) + I(W;U_1|U) \] (11)

\[ R_{11} + R_{11} \leq I(W;Y_1|U_1) + I(W;U) + I(W;U_1|U) \] (12)

For the second receiver, the analysis of error events imply that
\[ p(E_{21}^{dec}) + p(E_{22}^{dec}) \leq \frac{\epsilon}{6} \]

provided that
\[ R_{21}^p + R_{22}^p + R_{22} + R_{22} + R_{22}^p \leq I(UU_1;Y_1) \]
\[ R_{21}^p + R_{22}^p \leq I(U;Y_1;S|U) \] (13) (14)

Therefore, the probability of error can be bound as
\[ p_e^{(n)} = P\left( \bigcup_{k=1}^{2} E_{k}^{enc} \cup E_{1k}^{dec} \cup E_{2k}^{dec} \right) \leq \frac{\epsilon}{3} + \frac{\epsilon}{2} + \frac{\epsilon}{6} = \epsilon \]

Now combining (8) – (14) with (4) – (7), we obtain the following expressions
\[ R_{11} + R_{12} + R_{12} + R_{21} + R_{22} + R_{21} + R_{22} + R_{21} + R_{21} + R_{21}^p \leq I(WUU_1Y_1) - I(WUU_1;S) \] (15)
\[ R_{21} + R_{22} + R_{21} + R_{22} \leq I(UUU_1;Y_1|W) - I(UUU_1;S|W) \] (16)
\[ R_{11} + R_{21} \leq I(UU_1;Y_1|U) - I(UU_1;S|U) \] (17)
\[ R_{21} + R_{22} + R_{22}^p \leq I(UUU_2Y_2) - I(UUU_2;S) \] (18)
\[ R_{21}^p \leq I(U_1;Y_1;S|U) - I(U_1;S|U) \] (19)
\[ R_{22}^p \leq I(U_1;Y_1|U) - I(U_1;S|U) \] (20)
\[ R_{21} + R_{22} + R_{21} + R_{22} + R_{22}^p \leq I(UUU_1Y_1) - I(UUU_1;S) \] (21)

Now applying the Fourier-Motzkin elimination schemes to (15) – (21), with the constraints \( R_{22} = R_{22}^p + R_{22} \) and \( R_{21} = R_{21}^p + R_{21} \), we obtain the inequalities in (3).

IV. DEGRADED “Z” CHANNEL WITH CHANNEL STATE INFORMATION

Here we determine an inner bound and an outer bound on the capacity region of degraded discrete memoryless Z channel with channel state information. We will see that the derived outer bound coincides with the inner bound on the rates of the second transmitter to the second receiver, i.e. \( R_{21} + R_{22} \). Then we prove that using dirty paper coding, we can remove the effect of interference from the direction of one transmitter-receiver pair inside the inner bound provided.

A. An achievable rate region for the degraded discrete memoryless Z channel with channel state information

Definition 1: A “Z” channel is said to be degraded if the following Markov chain holds conditioned on every \( s \in \mathcal{S} \)
\[ X_2 \rightarrow (X_1, S, Y_2) \rightarrow Y_1 \]

i.e. the distribution in (2) is limited to the following
\[ p(y_1, y_2|x_1, x_2, s) = p(y_2|x_2, s)p(y_1|x_1, y_2, s) \]

Theorem 2: The achievable rate region for the degraded “Z” channel is the closure of the set of all rate triples \((R_{11}, R_{21}, R_{22})\) such that
\[ R_{11} + R_{21} \leq I(UW;Y_1) - I(UW;S) \] (27)
\[ R_{21} \leq I(U;Y_1|W) - I(U;S|W) \] (28)
\[ R_{11} \leq I(W;Y_1|U) - I(W;S|U) \] (29)
\[
R_{21} + R_{22} \leq I(UU_2;Y_2) - I(UU_2;S) \quad (30)
\]
\[
R_{22} \leq I(U_2;Y_2|U) - I(U_2;S|U) \quad (31)
\]

**Proof:** Set \( R_{21}^p = 0, R_{22}^p = 0, R_{21} = R_{21}, R_{22} = R_{22}, \) and \( U_1 = U \) in (15) – (21).

**Remark 2.1:** Notice that if the receivers be also aware of the channel state, then using Definition 1, one can show that (30) is redundant provided that the deterministic function \( x_1 \) be a one-to-one function.

**B. An outer bound on the capacity region of the degraded discrete memoryless Z channel with channel state information**

**Theorem 3:** The set of all rate triples \((R_1, R_{21}, R_{22})\) satisfying

\[
R_{11} + R_{21} \leq I(UW; Y_1) - I(W; S) \quad (32)
\]
\[
R_{21} \leq I(U; Y_1|WS) \quad (33)
\]
\[
R_{21} + R_{22} \leq I(UU_2; Y_2) - I(UU_2; S) \quad (34)
\]
\[
R_{22} \leq I(U_2; Y_2|U) - I(U_2; S|U) \quad (35)
\]

for some distribution of the form \( p(s,w,x_1,u,u_2,x_2) = p(s)p(w|s)p(x_1|w,s)p(u|s)p(u_2|u,s)p(x_2|u,u_2,s) \)

constitutes an outer bound on the capacity region of the degraded discrete memoryless Z channel with side information.

**Proof:** See Appendix II.

**Remark 3.1:** Notice that achievable rates of the second transmitter coincide with their counterparts in the outer bound and therefore, the second transmitter can communicate optimally with the receivers.

**C. Achievable rate for the degraded Gaussian Z channel with interference**

Now we study the Gaussian version of the Z channel with channel state information. First we define the Gaussian Z channel model with interference. Then we extend the achievable rate found for the discrete memoryless degraded Z channel with side information to the Gaussian case and use dirty paper coding to remove the negative effect of the interference in the channel associated with the first transmitter-receiver pair. The general model of the Gaussian Z channel is depicted in Fig. 3. The outputs are:

\[
Y_1^* = a_1 X_1^* + a_{21} X_2^* + (a_{11} + a_{21}) S + Z_1^* \quad (36)
\]
\[
Y_2^* = a_2 X_2^* + a_2 S + Z_2^* \quad (37)
\]

where for \( k = 1,2 \)

\[
\frac{1}{n} \sum_{i=1}^{n} E(X_k)^2 \leq P_k^*
\]
\[
Z_k^* \sim N(0,N_k)
\]
\[
S \sim N(0,Q)
\]

Now using the scaling transformation described in [10], we normalize each channel output to its respective noise power. The transformed model known as the standard Gaussian model is depicted in Fig. 4. The channel outputs are:

\[
Y_1 = X_1 + a X_2 + a_1 S + Z_1 \quad (38)
\]
\[
Y_2 = X_2 + a_2 S + Z_2 \quad (39)
\]

where for \( k = 1,2 \)

\[
\frac{1}{n} \sum_{i=1}^{n} E(X_k)^2 \leq P_k^*
\]
\[
Z_k^* \sim N(0,1)
\]
\[
a \triangleq \frac{a_{21}}{a_{22}} \sqrt{\frac{N_2}{N_1}}
\]
\[
a_1 \triangleq \frac{a_{11} + a_{21}}{\sqrt{N_1}}
\]
\[
a_2 \triangleq \frac{a_{22}}{\sqrt{N_2}}
\]

Now we use the dirty paper coding presented in [11] to derive the Gaussian version of the achievable rate presented in Theorem 2. We first present a Lemma to prove that the Gaussian version of Theorem 2, i.e. Theorem 4, is indeed the non-interfered achievable rate region in the direction of one transmitter-receiver pair, namely while one transmitter can successfully remove the negative effect of the interference, the other transmitter struggles to remove the interference inside its part of achievable region. In this paper, we optimize the Costa-coefficients so that the first transmitter achieves its own edge of the non-interfered achievable rate.
Let $\overline{U}, \overline{U}_2, \overline{W}$ be three pair-wise independent Gaussian random variables with zero mean and unit variance. Notice that $\overline{U}, \overline{U}_2, \overline{W}$ are also assumed to be independent of the noise and interference. We also assume that $0 \leq \xi \leq 1$ is an arbitrary real number with $\xi = 1 - \xi$. Also define

$U \triangleq \overline{U} + aS$

$W \triangleq \overline{W} + bS$

$U_2 \triangleq \overline{U}_2 + \gamma S$

$X_1 \triangleq \sqrt{P_1} \overline{W}$

$X_2 \triangleq \sqrt{P_2} \overline{U} + \sqrt{\xi P_1 \overline{U}_2}$

**Lemma 1:** Using (38), (39) and the definitions in (40), we have

$I(UW; Y_1) = I(UW_1; Y_1, S)$

$I(U; Y_1 | W) = I(U; Y_1, S | W)$

$I(W; Y_1 | U) = I(W; Y_1, S | U)$

Provided that

$\alpha = \frac{a \sqrt{P_2}}{1 + P_1 + \alpha^2 P_2}$

$\beta = \frac{a \sqrt{P_1}}{1 + P_1 + \alpha^2 P_2}$

**Proof:** See Appendix 1.

**Theorem 4:** The closure of the convex hull of the set of all triples $(R_{11}, R_{21}, R_{22})$ satisfying (61) – (65) is an achievable rate region for the degraded Gaussian $Z$ channel for any $0 \leq \xi \leq 1$.

**Proof:** Using the definitions in (40), the equalities in (41), and the inequalities (27) – (31), we derive (61) – (65).

**Remark 4.1:** Notice that with the Costa-coefficients of (42) and (43), we can remove the effect of the interference from the first three inequalities. In fact, it can be shown that (42) changes if one wants to remove the effect of interference in the rates of the second transmitter.

**Remark 4.2:** Notice that inequalities (61) – (65), are like those found in Corollary 1 of [3] where there is no interference.

**V. LATTICE STRATEGIES FOR THE GAUSSIAN DEGRADED “Z” CHANNEL WITH ADDITIVE INTERFERENCE**

Now we propose an achievable rate region using lattice based coding for the Gaussian degraded “Z” channel with additive interference non-causally available at both of the encoders under the high-SNR and strong interference regime utilizing the standard notation of [13], [14], and [15]. Our method is based on lattice transmission scheme, jointly decoding at the first decoder and successive decoding at the second decoder. Exploiting such coding scheme we remove the effect of the interference completely. The model of the Gaussian degraded $Z$ channel with the additive interference that we use in this section is depicted in Fig. 5.

The outputs are:

$Y_1 = X_1 + \alpha X_2 + (1 + \alpha)S + Z_1$

$Y_2 = X_2 + S + Z_2$

where $\alpha$ is a real number, $X_i$, $i = 1, 2$, is the channel input transmitted by user $i$ which is subject to the power constraint $P_i$, $Z_i$ is an AWGN with zero mean and variance $N_i$ ($Z_i \sim \mathcal{N}(0, N_i)$), and the interference signal $S$ is assumed to be i.i.d. Gaussian with variance $Q$, i.e. $S \sim \mathcal{N}(0, Q)$, independent of everything else and known non-causally at both encoders.

**Theorem 5:** An achievable region rate for the Gaussian degraded “Z” channel with side information non-causally available at both transmitters, denoted by $\mathcal{R}$, is given by

$$\mathcal{R} = \bigcup_{\rho \in [0, 1], \alpha \in [0, 1]} \mathcal{R}(\rho, \alpha)$$

$$\mathcal{R}(\rho, \alpha) = \begin{cases} R_{21} \leq \min(A_1, A_2) \\ R_{11} \leq \frac{1}{2} \log \left( 1 + \frac{P_1}{\alpha^2 \rho P_2 + N_1} \right) \\ R_{22} \leq \frac{1}{2} \log \left( \frac{\rho P_2}{\alpha^2 \rho P_2 + N_2} \right) \end{cases}$$

where
\[
A_1 = \frac{1}{2} \log \left( \frac{a^2 \rho P_2}{\tilde{\alpha}_0^2 a^2 \rho P_2 + \tilde{\alpha}_0^2 (a^2 \tilde{\rho} P_2 + N_1)} \right)
\]

\[
A_2 = \frac{1}{2} \log \left( \frac{\rho P_2}{\tilde{\alpha}_0^2 \rho P_2 + \tilde{\alpha}_0^2 (\tilde{\rho} P_2 + N_2)} \right)
\]

**Proof:** Our method is based on lattice transmission scheme, jointly decoding at the first decoder and successive decoding at the second decoder.

**Encoding:** Consider three lattices \( \Lambda_i \), \( i = 0,1,2 \) with fundamental Voronoi regions \( \mathcal{V}_i \), and the second moments \( \sigma_{\Lambda_0}^2 = \rho P_2, \sigma_{\Lambda_1}^2 = \rho P_2, \sigma_{\Lambda_2}^2 = \rho P_2 \), respectively. Encoder 1 wants to send message \( m_1 \) to receiver 1 while encoder 2 wants to send message \( m_2 \) to both receivers and message \( m_2 \) to receiver 2. We let \( R_0 = R_{21} \), \( R_1 = R_{11} \) and \( R_2 = R_{22} \). The messages \( m_i \) are carried by vector \( V_i \), where \( V_1 \) is uniformly distributed over \( \mathcal{V}_1 \), and \( V_0 \), \( V_1 \), and \( V_2 \), are pair-wise independent. Also, let \( D_i \sim \text{Unif}(\mathcal{V}_i) \), \( i = 0,1,2 \), be three dither signals which are uniformly distributed over \( \mathcal{V}_i \), and independent of each other. In our encoding structure senders 1 and 2 send \( X_1 = W \) and \( X_2 = U + U_2 \), respectively, where \( W, U \) and \( U_2 \) are generated as:

\[
U = [V_0 - \alpha_0 S + D_0] \mod \Lambda_0
\]

\[
W = [V_1 - \alpha_1 \tilde{\alpha}_0 S + D_1] \mod \Lambda_1
\]

\[
U_2 = [V_2 - \alpha_2 \tilde{\alpha}_0 S + D_2] \mod \Lambda_2
\]

where \( \tilde{\alpha}_0 = 1 - \alpha_0 \) and the MMSE criterion is used to determine \( \alpha_i \), \( i = 0,1,2 \). Note that using this encoding structure we have: \( \frac{1}{n} E \|X_i\|^2 = P_i \) and \( \frac{1}{n} E \|X_2\|^2 = P_2 \).

**Decoding:** To decode \( \{V_0, V_2\} \) at decoder 2, we use a successive decoding scheme in which decoder 2 first decodes \( V_0 \) and then decodes \( V_2 \). Therefore, decoder 2, after receiving \( Y_2 \) and using the lattice \( \Lambda_0 \), computes

\[
Y_2^{(2)} = [\alpha_2 Y_2 - D_0] \mod \Lambda_0
\]

\[
= [\alpha_0 (U + U_2 + S + D_0) - D_0] \mod \Lambda_0
\]

\[
= [V_0 - \tilde{\alpha}_0 U + \alpha_0 (U_2 + Z_2)] \mod \Lambda_0
\]

\[
= [V_0 + Z_{02}] \mod \Lambda_0
\]

where \( Z_{02} = - \tilde{\alpha}_0 U + \alpha_0 (U_2 + Z_2) \). Therefore, we have:

\[
R_0 = \frac{1}{n} I(V_0; Y_2^{(2)}) = \frac{1}{n} \left\{ h(Y_2^{(2)}) - h(Y_2^{(2)}|V_1) \right\}
\]

\[
= \frac{1}{n} \left\{ h(Y_2^{(2)}) - h([-\tilde{\alpha}_0 U + \alpha_0 (U_2 + Z_2)] \mod \Lambda_0) \right\}
\]

\[
\geq \frac{1}{2} \log \left( \frac{\rho P_2}{\tilde{\alpha}_0^2 \rho P_2 + \tilde{\alpha}_0^2 (\tilde{\rho} P_2 + N_2)} \right) - \frac{1}{2} \log \left( 2 \pi e \left( \tilde{\alpha}_0^2 \rho P_2 + \tilde{\alpha}_0^2 (\tilde{\rho} P_2 + N_2) \right) \right)
\]

where (51) is due to the fact that (i) \( Y_0^{(2)} \) is uniformly distributed over \( Y_0 \), (ii) for fixed second moment, Gaussian distribution maximizes the entropy, and (iii) modulo operation reduces the second moment. Hence, as long as \( \Lambda_0 \) is a good lattice for quantization, we have:

\[
R_0 \leq \frac{1}{2} \log \left( \frac{\rho P_2}{\tilde{\alpha}_0^2 \rho P_2 + \tilde{\alpha}_0^2 (\tilde{\rho} P_2 + N_2)} \right) = A_2 \quad (52)
\]

Note that the optimal \( \alpha_0 \) for sender 2 from the second receiver standpoint is \( \alpha^{\text{opt}}_0 = \frac{\rho P_2}{\rho P_2 + \tilde{\rho} P_2 + N_2} = \frac{\rho P_2}{\rho P_2 + N_2} \), and by substituting this \( \alpha^{\text{opt}}_0 \) into inequalities (44), we obtain:

\[
R_0 \leq \frac{1}{2} \log \left( 1 + \frac{\rho P_2}{\rho P_2 + N_2} \right) \quad (53)
\]

Also, note that this \( \alpha^{\text{opt}}_0 \) is non-optimal from the first receiver standpoint.

Now, decoder 2 using lattice \( \Lambda_2 \) computes

\[
Y_2^{(2)} = [\alpha_2 (\tilde{\alpha}_0 Y_2 + Z_{02}) - D_2] \mod \Lambda_2
\]

\[
= [\alpha_2 (U_2 + Z_2) + \alpha_2 \tilde{\alpha}_0 S - D_2] \mod \Lambda_2
\]

\[
= [V_2 - \tilde{\alpha}_2 U_2 + \alpha_2 Z_2] \mod \Lambda_2
\]

Therefore, we have:

\[
R_2 = \frac{1}{n} I(V_2; Y_2^{(2)}) = \frac{1}{n} \left\{ h(Y_2^{(2)}) - h(Y_2^{(2)}|V_2) \right\}
\]

\[
= \frac{1}{n} \left\{ h(Y_2^{(2)}) - h([-\tilde{\alpha}_2 U_2 + \alpha_2 Z_2] \mod \Lambda_2) \right\}
\]

\[
\geq \frac{1}{2} \log \left( \frac{\tilde{\rho} P_2}{\tilde{\alpha}_2^2 \tilde{\rho} P_2 + \alpha_2^2 N_2} \right) - \frac{1}{2} \log \left( 2 \pi e (\tilde{\alpha}_2^2 \tilde{\rho} P_2 + \alpha_2^2 N_2) \right) \quad (54)
\]

and for good lattice for quantization, the achievable rate is given by

\[
R_2 \leq \frac{1}{2} \log \left( \frac{\tilde{\rho} P_2}{\tilde{\alpha}_2^2 \tilde{\rho} P_2 + \alpha_2^2 N_2} \right) \quad (55)
\]

Note that the optimal \( \alpha_2 \) is \( \alpha^{\text{opt}}_2 = \frac{\tilde{\rho} P_2}{\rho P_2 + N_2} \), and by substituting this \( \alpha^{\text{opt}}_2 \) into (55), we obtain:

\[
R_2 \leq \frac{1}{2} \log \left( 1 + \tilde{\rho} P_2 \right) \quad (56)
\]

To decode \( \{V_0, V_1\} \) at decoder 1, we use a similar method as [9] for MAC. Therefore, as long as \( \Lambda_0 \) is a good lattice for quantization, we have:

\[
R_0 \leq \frac{1}{2} \log \left( \frac{a^2 \rho P_2}{\tilde{\alpha}_0^2 a^2 \rho P_2 + \tilde{\alpha}_0^2 (a^2 \tilde{\rho} P_2 + N_1)} \right) = A_1 \quad (57)
\]

Note that the optimal \( \alpha_0 \) for sender 2 from the first receiver standpoint is \( \alpha^{\text{opt}}_0 = \frac{a^2 \rho P_2}{a^2 \rho P_2 + a^2 \tilde{\rho} P_2 + N_1} = \frac{a^2 \rho P_2}{a^2 \tilde{\rho} P_2 + N_1} \), and by substituting this \( \alpha^{\text{opt}}_0 \) into (57), we obtain:

\[
R_0 \leq \frac{1}{2} \log \left( 1 + \frac{a^2 \rho P_2}{a^2 \tilde{\rho} P_2 + N_1} \right) \quad (58)
\]
Also, note that this $\alpha_0^{\text{opt}}$ is non-optimal from the second receiver standpoint. Similarly, for good lattice for quantization we have:

$$R_1 \leq \frac{1}{2} \log \left( \frac{P_1}{\bar{\alpha}_1 P_1 + \alpha_1^2 (a^2 \bar{\rho} P_2 + N_1)} \right)$$  \(59\)

Meanwhile, the optimal $\alpha_1$ is $\alpha_1^{\text{opt}} = \frac{P_1}{P_1 + a^2 \bar{\rho} P_2 + N_1}$, and by substituting this $\alpha_1^{\text{opt}}$ into (59), we obtain:

$$R_1 \leq \frac{1}{2} \log \left( 1 + \frac{P_1}{a^2 \bar{\rho} P_2 + N_1} \right)$$  \(60\)

VI. CONCLUSION

In this paper, we derived an achievable rate region for the general "Z" channel with side information non-causally available at the transmitters using Han-Kobayashi, rate splitting, Gelf’and-Pinsker coding, and CMGE jointly decoding. We also showed that our rate region subsumes the achievable rate region of the multiple access channel with side information and degraded broadcast channel with side information. We then derived the achievable rate region of a special case of degraded Z channels with side information. We also saw that using the Costa dirty paper coding for the degraded Gaussian Z channel, we can remove the effect of interference in the direction of one of the transmitter-receiver pairs. We saw that using dirty paper coding, we can remove the effect of the interference completely.

REFERENCES

\[ R_{11} + R_{21} \leq \frac{1}{2} \log \left( 1 + \frac{P_1 + a^2 \xi P_2}{a^2 \xi P_2 + 1} \right) \] (61)

\[ R_{21} \leq \frac{1}{2} \log \left( 1 + \frac{a^2 \xi P_2}{a^2 \xi P_2 + 1} \right) \] (62)

\[ R_{11} \leq \frac{1}{2} \log \left( 1 + \frac{P_1}{a^2 \xi P_2 + 1} \right) \] (63)

\[ R_{21} + R_{22} \leq \frac{1}{2} \log \left( \frac{P_2 + a^2 Q + 1}{\det \begin{pmatrix} P_2 + a^2 Q + 1 & \sqrt{\xi P_2 + \alpha_2 Q} & \sqrt{\xi P_2 + \alpha_2 Q} \\ \sqrt{\xi P_2 + \alpha_2 Q} & 1 + a^2 Q & \alpha Q \\ \sqrt{\xi P_2 + \alpha_2 Q} & \alpha Q & 1 + \gamma^2 Q \end{pmatrix}} \right) \] (64)

\[ R_{22} \leq \frac{1}{2} \log \left( \frac{\det \begin{pmatrix} P_2 + a^2 Q + 1 & \sqrt{\xi P_2 + \alpha_2 Q} \\ \sqrt{\xi P_2 + \alpha_2 Q} & 1 + a^2 Q \end{pmatrix}}{\det \begin{pmatrix} P_2 + a^2 Q + 1 & \sqrt{\xi P_2 + \alpha_2 Q} \\ \sqrt{\xi P_2 + \alpha_2 Q} & \alpha Q \end{pmatrix}} \right) \] (65)

\section*{APPENDIX I}

\textbf{Proof of Lemma 1:} First we prove that

\[ I(UW; Y_1) = I(UW; Y_1, S) \]

In fact, it suffices to show that

\[ h(UW | Y_1) = h(UW | Y_1, S). \]

We have

\[ h(UW | Y_1) = h(\tilde{U} + aS, \tilde{W} + \beta S | X_1 + aX_2 + a_1 S + Z_1) = h \left( \tilde{U} + aS, \tilde{W} + \beta S \Big| \sqrt{P_1} \tilde{W} + a \sqrt{\xi P_2} \tilde{U} + a \sqrt{\xi P_2} \tilde{U}_2 + a_1 S + Z_1 \right) = \]

\[ = h \left( \tilde{U} - \frac{\alpha}{a_1} \left( \sqrt{P_1} \tilde{W} + a \sqrt{\xi P_2} \tilde{U} + a \sqrt{\xi P_2} \tilde{U}_2 + Z_1 \right), \tilde{W} - \frac{\beta}{a_1} \left( \sqrt{P_1} \tilde{W} + a \sqrt{\xi P_2} \tilde{U} + a \sqrt{\xi P_2} \tilde{U}_2 + Z_1 \right) | X_1 + aX_2 + a_1 S + Z_1 \right) = \]

\[ = h \left( \psi_u, \psi_w \bigg| \sqrt{P_1} \tilde{W} + a \sqrt{\xi P_2} \tilde{U} + a \sqrt{\xi P_2} \tilde{U}_2 + a_1 S + Z_1 \right) \]

where we define \( \psi_u \equiv \tilde{U} - \frac{\alpha}{a_1} \left( \sqrt{P_1} \tilde{W} + a \sqrt{\xi P_2} \tilde{U} + a \sqrt{\xi P_2} \tilde{U}_2 + Z_1 \right) \) and \( \psi_w \equiv \tilde{W} - \frac{\beta}{a_1} \left( \sqrt{P_1} \tilde{W} + a \sqrt{\xi P_2} \tilde{U} + a \sqrt{\xi P_2} \tilde{U}_2 + Z_1 \right) \).

It is easily seen that

\[ E \left[ \psi_u \times \left( \sqrt{P_1} \tilde{W} + a \sqrt{\xi P_2} \tilde{U} + a \sqrt{\xi P_2} \tilde{U}_2 + Z_1 \right) \right] = E \left[ \psi_w \times \left( \sqrt{P_1} \tilde{W} + a \sqrt{\xi P_2} \tilde{U} + a \sqrt{\xi P_2} \tilde{U}_2 + Z_1 \right) \right] = 0 \]

provided that

\[ \alpha = \frac{aa_1 \sqrt{\xi P_2}}{1 + P_1 + a^2 P_2} \]
\[ \beta = \frac{a_1 \sqrt{P_1}}{1 + P_1 + a^2 P_2} \]

and therefore \( \psi_u \) and \( \psi_w \) are both independent of \( \left( \sqrt{P_1} W + a \sqrt{\xi} P_2 U + a \sqrt{\xi} P_2 U_2 + Z_1 \right) \). We also know that \( \psi_u \) and \( \psi_w \) are both independent of \( S \). Therefore, we have

\[ h(\psi_u, \psi_w | Y_2) = h(\psi_u, \psi_w) \]

we can also easily prove that

\[ h(U, W | Y_2, S) = h(\psi_u, \psi_w) \]

The other inequalities of (37) are proved in just the same way.

**APPENDIX II**

**Proof of Theorem 3**: Suppose that \( (2^n R_1, 2^n R_2, n, \epsilon) \) is a code for the degraded discrete memoryless Z channel. We define the auxiliary random variables as follows

\[
W_i \triangleq (M_{1i}, Y_{1i}^{i-1}, S_i^{(n)}_{i+1}) \\
U_i \triangleq (M_{2i}, Y_{2i+1}^{i-1}, S_i^{(n)}) \\
U_{2i} \triangleq (M_{2i}, M_{2i+1}, Y_{2i+1}^{i-1}, S_i^{(n)})
\]

First we prove the bound on \( R_{11} + R_{21} \). We have

\[
n(R_{11} + R_{21})^{(a)} = H(M_{11}, M_{21}|Y^n) + I(M_{11}, M_{21}; Y^n) - I(M_{11}, M_{21}; S^n)
\]

\[
\leq n\epsilon_{11n} + \sum_{i=1}^{n} I(M_{1i}, M_{2i}; Y_{1i}^{i-1}) - I(M_{1i}, M_{2i}; S_i^{(n)}_{i+1})
\]

\[
\leq n\epsilon_{11n} + \sum_{i=1}^{n} I(M_{1i}, M_{2i}, Y_{1i}^{i-1}; Y_{1i}) - I(M_{1i}, M_{2i}, S_i^{(n)}_{i+1}; S_i)
\]

\[
= n\epsilon_{11n} + \sum_{i=1}^{n} I(M_{1i}, M_{2i}, Y_{1i}^{i-1}, S_i^{(n)}_{i+1}; Y_{1i}) - I(S_i^{(n)}_{i+1}; Y_{1i}^{i-1}; M_{1i}, M_{2i}, Y_{1i}^{i-1})
\]

\[
- I(M_{1i}, M_{2i}, S_i^{(n)}_{i+1}, Y_{1i}^{i-1}, S_i) + I(Y_{1i}^{i-1}, S_i | M_{1i}, M_{2i}, S_i^{(n)}_{i+1})
\]

\[
\leq n\epsilon_{11n} + \sum_{i=1}^{n} I(M_{1i}, M_{2i}, Y_{1i}^{i-1}, S_i^{(n)}_{i+1}; Y_{1i}) - I(M_{1i}, M_{2i}, S_i^{(n)}_{i+1}, Y_{1i}^{i-1}, S_i)
\]

\[
\leq n\epsilon_{11n} + \sum_{i=1}^{n} I(U_i, W_i; Y_{1i}) - I(W_i; S_i).
\]

where (a) follows from the independence of the messages from the state of the channel, (b) follows from Fano’s inequality and the chain rule for mutual information, (c) follows from the non-negativity of mutual information and from the fact that the channel state elements are i.i.d., and (d) follows from Csiszar-Körner identity [12].

Next, we prove the bound on \( R_{21} \). We have

\[
nR_{21} = H(M_{21}|M_{11}, S^n, Y^n) + I(M_{21}; Y^n|M_{11}, S^n)
\]

\[
\leq n\epsilon_{12n} + \sum_{i=1}^{n} I(M_{2i}; Y_{1i}|M_{1i}, S_i^{(n)}_{i+1}, Y_{1i}^{i-1})
\]
Now we prove the bound on $R_{21} + R_{22}$. We have

$$n(R_{21} + R_{22}) = H(M_{21}, M_{22} | Y^n_2) + I(M_{21}, M_{22} | Y^n_2) - I(M_{21}, M_{22}; S^n)$$

$$\leq n\varepsilon_{21n} + \sum_{i=1}^{n} I(M_{21}, M_{22}, Y^n_{2,i+1}; Y_{2i}) - I(M_{21}, M_{22}, S^{i-1}; S_i)$$

$$= n\varepsilon_{21n} + \sum_{i=1}^{n} I(M_{21}, M_{22}, Y^n_{2,i+1}, S^{i-1}; Y_{2i}) - I(S^{i-1}, Y_{2i} | M_{21}, M_{22}, Y^n_{2,i+1})$$

$$- I(M_{21}, M_{22}, S^{i-1}, Y^n_{2,i+1}; S_i) + I(Y^n_{2,i+1}; S_i | M_{21}, M_{22}, S^{i-1})$$

$$= n\varepsilon_{21n} + \sum_{i=1}^{n} I(M_{21}, M_{22}, Y^n_{2,i+1}, S^{i-1}; Y_{2i}) - I(M_{21}, M_{22}, S^{i-1}, Y^n_{2,i+1}; S_i)$$

$$= n\varepsilon_{21n} + \sum_{i=1}^{n} I(U_i, U_{2i}; Y_{2i}) - I(U_i, U_{2i}; S_i).$$

where (a) follows from Csiszar-Körner identity [12].

Finally, we prove the bound on $R_{22}$. We have

$$nR_{22} = H(M_{22} | M_{21}, Y^n_2) + I(M_{22}; Y^n_2 | M_{21}) - I(M_{22}; S^n | M_{21})$$

$$\leq n\varepsilon_{22n} + \sum_{i=1}^{n} I(M_{22}; Y_{2i} | M_{21}, Y^n_{2,i+1}) - I(M_{22}; S_i | M_{21}, S^{i-1})$$

$$= n\varepsilon_{22n} + \sum_{i=1}^{n} I(M_{22}, S^{i-1}, Y_{2i} | M_{21}, Y^n_{2,i+1}) - I(S^{i-1}, Y_{2i} | M_{21}, M_{22}, Y^n_{2,i+1})$$

$$- I(M_{22}, Y^n_{2,i+1}; S_i | M_{21}, S^{i-1}) + I(Y^n_{2,i+1}; S_i | M_{21}, M_{22}, S^{i-1})$$

$$= n\varepsilon_{22n} + \sum_{i=1}^{n} I(M_{21}, M_{22}, S^{i-1}, Y^n_{2,i+1}, Y_{2i} | M_{21}, Y^n_{2,i+1}) - I(M_{21}, M_{22}, Y^n_{2,i+1}, S^{i-1}; S_i | M_{21}, S^{i-1})$$

$$= n\varepsilon_{22n} + \sum_{i=1}^{n} I(S^{i-1}, Y_{2i} | M_{21}, Y^n_{2,i+1}) + I(M_{21}, M_{22}, Y^n_{2,i+1}, Y_{2i} | M_{21}, Y^n_{2,i+1}, S^{i-1})$$

$$- I(Y^n_{2,i+1}; S_i | M_{21}, S^{i-1}) - I(M_{21}, M_{22}, S^{i-1}, S_i | M_{21}, S^{i-1}, Y^n_{2,i+1})$$

$$= n\varepsilon_{22n} + \sum_{i=1}^{n} I(M_{21}, M_{22}, Y^n_{2,i+1}, Y_{2i} | M_{21}, Y^n_{2,i+1}, S^{i-1}) - I(M_{21}, M_{22}, S^{i-1}; S_i | M_{21}, S^{i-1}, Y^n_{2,i+1})$$

$$= n\varepsilon_{22n} + \sum_{i=1}^{n} I(M_{21}, M_{22}, Y^n_{2,i+1}, Y_{2i} | M_{21}, Y^n_{2,i+1}, S^{i-1}) - I(M_{21}, M_{22}, S^{i-1}; S_i | M_{21}, S^{i-1}, Y^n_{2,i+1}).$$

(b)
\[
\sum_{i=1}^{n} I(M_{21}, M_{22}, S_{i-1}^{(n)}, Y_{2, i+1}^{(n)}; Y_{2i} | M_{21}, Y_{2, i+1}^{(n)}; X_{i-1}), D(M_{21}, M_{22}, S_{i-1}^{(n)}, Y_{2, i+1}^{(n)}; D_{i-1}, Y_{2, i+1}^{(n)})
\]

where (a) and (b) both follow from Csiszar-Körner identity [12].