On the Capacity of Special Classes of Gaussian Relay Networks with Orthogonal Components and Noncausal State Information At Source

Amir Sonee and Ghosheh Abed Hodtani
Department of Electrical Engineering, Faculty of Engineering
 Ferdowsi University of Mashhad
 Mashhad, Iran
 Email: Amir.Sonee@stu-mail.um.ac.ir, ghodtani@gmail.com

Abstract—THIS PAPER IS ELIGIBLE FOR THE STUDENT PAPER AWARD. In this paper, we study relay networks with orthogonal components in presence of noncausal channel state information (CSI) available at the source. We propose an upper bound on the capacity of the discrete memoryless model (DM) for the case in which just the source component intended for the destination is encoded against the CSI known non-causally at the source. Also, we derive capacity for two special classes of the Gaussian structure of the model. The first class is the one for which we have obtained the upper bound and the second class is the one in which all of the source components intended for the destination are encoded against the CSI available non-causally at the source. Moreover, we derive capacity for two special classes of Gaussian relay network with orthogonal components and noncausal CSI at the source.

The paper is organized as follows. In section I, the network model is introduced. In section II, Our main results including the lower and upper bounds and special capacity results for the Gaussian model is presented. Finally, the paper is concluded in section V.

II. Preliminaries

A discrete memoryless (DM) relay network with noncausal CSI at source is a channel with $N+2$ terminals consisting of a source (node 1), $N$ relays (node $i \in \mathcal{T} = \{2, \ldots, N+1\}$) and a destination (node $N+2$). The model is characterized as $(X_1 \times X_2 \times \ldots \times X_{N+1}, p(y_2, y_3, \ldots, y_{N+2}|s, x_1, x_2, \ldots, x_{N+1}), Y_2 \times Y_3 \times \ldots \times Y_{N+2})$ with $(x_1, x_2, \ldots, x_{N+1}, s, y_2, \ldots, y_{N+2}) \in (X_1 \times X_2 \times \ldots \times X_{N+1} \times S \times Y_2 \times Y_3 \times \ldots \times Y_{N+2})$ where $X_1$ and $X_i$ are the channel input alphabets at the source and relay $i$, respectively, over which the channel input random variables $X_1$ and $X_i$ take value. Also, $Y_{N+2}$ and $Y_i$ are the channel output alphabets at the destination and relay $i$, respectively, over which the channel output random variables $Y_{N+2}$ and $Y_i$ take value and these output random variables are governed by the input random variables and the channel state $S$.

A $(2^{nR}, n)$ code for the relay network with noncausal CSI at the source consists of a set of integers $\mathcal{W} = \{1, \ldots, 2^{nR}\}$, one encoder $f : S \times \mathcal{D} \rightarrow X_i^n$ and $N$ sets of functions $\{r_{p,i}\}_{i=1}^n$ at the relays, such that $x_{p,i} = r_{p,i}(y_{p,1}, y_{p,2}, \ldots, y_{p,i-1})$, $p \in \mathcal{T}$, where $x_{p,i}$ is the $i$'th component of the relay $p$ codeword $x_p = (x_{p,1}, \ldots, x_{p,n})$ and a decoder $g : Y_{N+2}^n \rightarrow \mathcal{W}$ at the destination. The average error probability for this code is defined as $P_e^n = E_{\mathcal{W}}(P[g(y_{N+2}^n) \neq W])$. A rate $R$ for the relay network with noncausal state at the source is said to be achievable if there exists a sequence of $(2^{nR}, n)$ codes such that $P_e^n < \epsilon$ for any $\epsilon > 0$ and sufficiently large $n$. Also, the capacity $C$ of this channel is defined as the supremum of the set of achievable rates.
A DM relay network with orthogonal components and noncausal CSI at the source as it is shown in Figure 1, is a DM relay network in which the source and relay i’s input alphabets $X_i$ and $X_t$ are divided into $N + 1$ and $N$ orthogonal components $\{X_{i,j}\}_{i \in T \cup \{N+2\}}$ and $\{X_{t,j}\}_{j \in T \cup \{N+2\}, j \neq i}$, respectively. In this case, the channel probability distribution function can be written as ([5])

$$p(x_j)_{j \in \{N+2\} \cup \{s\}, x_j \in \{1\} \cup \{N+2\} \cup \{1\} \cup \{N+2\} \cup \{1\}} = \prod_{i \in T} p(y_i | s, x_{ji})_{j \in \{1\} \cup \{N+2\} \cup \{1\} \cup \{N+2\} \cup \{1\}}$$

III. MAIN RESULTS

In this section, first, a lower bound on the capacity of relay networks with orthogonal components and noncausal CSI at the source is derived and then an upper bound on the capacity of a special class of this network is obtained in which just the source component intended for the destination is encoded against the noncausal CSI. Next, we establish the capacity for two special cases of the Gaussian structure of the network model.

**Proposition 1** (modified version of [5, Theorem 1]): A lower bound on the capacity of DM relay network with orthogonal components and noncausal CSI at the source is given by

$$C \geq \max_{\text{MCST}} \left\{ \sum_{i \in M} \left( R_{i,r,i}^1 + R_{i,r,i}^2 \right) + R_{1,d}^1 + R_{2,d}^1 \right\}$$

(2)

where

$$R_{i,r,i}^1 = I(U_i; Y_i | \{X_{ji}\}_{j \in \{N+2\} \cup \{s\} \cup \{N+2\} \cup \{s\}} \cup \{1\} \cup \{N+2\} \cup \{1\})$$

(3a)

$$R_{i,r,i}^2 = I(\{X_{ji}\}_{j \in \{M\}}, Y_i | \{X_{ji}\}_{j \in \{N+2\} \cup \{s\}} \cup \{1\} \cup \{N+2\} \cup \{1\})$$

(3b)

$$R_{1,d}^1 = I(\{X_{k,N+2}\}_{k \in \{M\}}, Y_{N+2} \{X_{k,N+2}\}_{k \in \{s\}}$$

(3c)

$$R_{2,d}^1 = I(U_1; Y_{N+2} \{X_{k,N+2}\}_{k \in \{T\}}) - I(U_1; S \{X_{k,N+2}\}_{k \in \{T\}})$$

(3d)

with $M^c = T \setminus M$ and the maximum is taken over joint probability distributions of the form

$$p(s) \prod_{i \in T} \prod_{j \in \{N+2\} \cup \{s\}} p(x_{ji})_{x_{ji} \in \{1\} \cup \{N+2\} \cup \{1\}}$$

(4)

**Proof:** This rate can be easily achieved if we combine the coding techniques explained in [5, Appendix 2] for a relay network with orthogonal components with the well-known Gelfand-Pinsker coding for encoding components of the source message intended for the relays and destination against the noncausal CSI at the source. The proof is omitted for brevity.

**Theorem 1:** An upper bound on the capacity of DM relay network with orthogonal components and noncausal CSI at the source is given by

$$C \leq \max_{\text{MCST}} \left\{ \sum_{i \in M} \left( R_{i,r,i}^u + R_{2,r,i}^u \right) + R_{1,d}^u + R_{2,d}^u \right\}$$

(5)

where

$$R_{i,r,i}^u = I(X_{1,i}; Y_i | S_i, \{X_{ji}\}_{j \in \{T\} \cup \{s\}}$$

(6a)

$$R_{2,r,i}^u = I(\{X_{ji}\}_{j \in \{M\}}; Y_i | \{X_{ji}\}_{j \in \{M\}}$$

(6b)

$$R_{1,d}^u = I(X_{1,N+2}; Y_{N+2} | S, \{X_{i,N+2}\}_{i \in \{T\}}$$

(6c)

$$R_{2,d}^u = I(X_{1,N+2}; Y_{N+2} | S, \{X_{i,N+2}\}_{i \in \{T\}}$$

(6d)

and maximization is over all joint probability of the form

$$p(s) \prod_{i \in T} p(x_{ji} | x_{ji} \in \{1\} \cup \{N+2\} \cup \{1\}) p(x_{ji} | x_{ji} \in \{1\} \cup \{N+2\} \cup \{1\})$$

(7)

**Proof:** Refer to Appendix.

Now, we consider a special case of Gaussian relay networks with orthogonal components and noncausal CSI at the source in which the source sends the message intended for the destination by using dirty paper coding (DPC) and the messages sent to the relays are independent of the CSI. In this model, the channel outputs for the relay $j$ and destination at time instant $i$ is described as

$$Y_{j,i} = \sum_{r \in \{1\} \cup \{N+2\}} X_{r,j,i} + Z_{j,i}, j \in \{T\}$$

(8a)

$$Y_{N+2,i} = \sum_{r \in \{1\} \cup \{N+2\}} X_{r,N+2,i} + S_i + Z_{N+2,i}$$

(8b)

where $S_i$ is the additive Gaussian noncausal CSI random variable available at the source with zero mean and variance $Q$. Also, $Z_{j,i}, j \in \{T \cup \{N+2\}$, are mutually independent additive Gaussian noise random variables with zero mean and variance $N_j$ which are independent of the $S$ and the channel inputs. Power constraints on the components of the source and relays codewords are as
\[ \sum_{i=1}^{n} \mathbb{E}[X_{rj}^2] \leq nP_{rj}, r \in \{1\} \cup T, j \in T \cup \{N+2\}, r \neq j. \tag{9} \]

Now we use Proposition 1 and Theorem 1 to exploit the capacity of the Gaussian relay network with orthogonal components in (8).

**Theorem 2**: The capacity of Gaussian relay network with orthogonal components of (8) is given by (10) at the bottom of the page where the maximization is over all \( \rho_{i} \in [-1, 0] \) and \( \rho_{ij} \in [0, 1] \) with
\[ \rho_{i}^2 + \sum_{j \in T} \rho_{ij}^2 \leq 1 \tag{11} \]

**Proof**: We use a proof which is similar to the one in [6] in some steps. **Converse part**: We first fix a distribution as in (7) for the random variables \( S, \{X_{ij}\}_{j \in \{1\} \cup T, j \neq i; \{Y_{ij}\}_{j \in \{N+2\} \cup T} \) where
\[ \mathbb{E}[X_{ij}^2] = P_{ij}, i \in \{1\} \cup T, j \in \{N+2\} \cup T, j \neq i, \]
\[ \mathbb{E}[X_{iN+2}X_{jN+2}] = \rho_{ij} \sqrt{P_{iN+2}P_{jN+2}}, i \in T, \]
\[ \mathbb{E}[X_{iN+2}S] = \rho_{i} \sqrt{P_{iN+2}Q}. \]
Now, we calculate the terms in (5),
\[ R_{1,r}^n + R_{2,r}^n = h(Y_i) - h(Y_i|X_{iM}, j \neq i) - h(Y_i|X_{iM}, j = i) \]
\[ h(Y_i|S, X_{iM}, j \neq i) - h(Y_i|S, X_{iM}, j = i) \]
\[ = h(\sum_{j \in \{1\} \cup M^c} X_{ji} + Z_i|X_{iM}, j \neq i) - h(X_{i1} + Z_i|X_{iM}, j \neq i) \]
\[ = h(\sum_{j \in \{1\} \cup M^c} X_{ji} + Z_i|X_{iM}, j \neq i) - h(X_{i1} + Z_i|X_{iM}, j \neq i) \]
\[ \leq h(\sum_{j \in \{1\} \cup M^c} X_{ji} + Z_i - h(Z_i) \]
\[ \leq \frac{1}{2} \log \left( 1 + \frac{\sum_{j \in \{1\} \cup M^c} \tilde{P}_{j|i}}{N_{i}} \right) \tag{12} \]
\[ \gamma_1 = \frac{h(S|Y_{iN+2}, \{X_{iN+2}\}_{i \in M}) - h(S|Y_{iN+2}, \{X_{iN+2}\}_{i \in T})}{\rho_{i} \sqrt{P_{iN+2}Q}} \]
\[ \gamma_r = \frac{h(S|Y_{iN+2}, \{X_{iN+2}\}_{i \in M}) - h(S|Y_{iN+2}, \{X_{iN+2}\}_{i \in T})}{\rho_{i} \sqrt{P_{iN+2}Q}} \]
\[ \gamma_r \leq \frac{1}{2} \log \left( 1 + \frac{\tilde{P}_{N+2}(1 - \rho_{i}^2 - \sum_{j \in \{1\} \cup T} \rho_{ij}^2) + N_{i+2}}{\rho_{i} \sqrt{P_{iN+2}Q}} \right) \tag{17} \]
\[ \therefore R_{1,1}^n + R_{2,1}^n \leq \frac{1}{2} \log \left( 1 + \frac{\tilde{P}_{N+2}(1 - \rho_{i}^2 - \sum_{j \in \{1\} \cup T} \rho_{ij}^2) + N_{i+2}}{\rho_{i} \sqrt{P_{iN+2}Q}} \right) \tag{17} \]
\[ C = \max_{M \in T} \left\{ \frac{1}{2} \log \left( 1 + \frac{\sum_{j \in \{1\} \cup M^c} \tilde{P}_{j|i}}{N_{i}} \right) + \frac{1}{2} \log \left( 1 + \frac{\tilde{P}_{iN+2}(1 - \rho_{i}^2 - \sum_{j \in \{1\} \cup T} \rho_{ij}^2) + N_{i+2}}{\rho_{i} \sqrt{P_{iN+2}Q}} \right) \right\} \tag{10} \]
\[ \frac{1}{2} \log \left( 1 + \frac{\sum_{j \in M^c} (\sqrt{P_{N+2}^j + \rho_{1j}} \sqrt{P_{1N+2}^j})^2}{\Delta} \right) \]
where \( \Delta = \left( 1 - \rho_{s}^2 - \sum_{j \in T} \rho_{1j}^2 \right) P_{N+2} + (\sqrt{Q} + \rho_{s} \sqrt{P_{N+2}})^2 + N_{N+2} \).
It can be readily seen that the terms obtained in (12) and (18) are increasing functions of \( P_{ji} \) for \( i \in T, \ j \in \{1\} \cup T \), and \( P_{N+2} \) for \( j \in T \), and as a result the upper bound (5) is maximized for the maximum value of these parameters i.e. \( P_{ji} = P_{ji}^* \) and \( P_{N+2} = P_{N+2}^* \). Also, it has been shown in [6] that the second term in (16) is an increasing function of \( P_{N+2} \) so, the maximum value of this parameter i.e. \( P_{N+2} \) maximizes this term. Finally, as it can be seen if we let \( \rho_{s} \in [-1, 0] \) and \( \rho_{ij} \in [0, 1] \) the two terms in (18) would take their maximum value. As a result, we showed that if \( (S, \{X_{ij}, Z_{j}\}_{i \in \{1\} \cup T, j \in \{N+2\} \cup T, j \neq i}) \) are jointly Gaussian the the upper bound (5) is maximized.

Achievability part: We use the same procedure as in Proposition 1 along with the following assumptions. We let \( X_{ij} \)'s for \( i \in \{1\} \cup T, j \in \{N+2\} \cup T, i \neq j \) to be jointly Gaussian random variables \( \mathcal{N}(0, P_{ij}) \). Also, we assume that \( X_{1j} \)'s, \( j \in T \), and \( X_{ij} \)'s, \( i \in T, j \in \{N+2\} \cup T, i \neq j \), are all independent of each other and independent of the state \( S \). The input random variable \( X_{N+2} \) is jointly Gaussian with \( S \) and is independent of \( X_{ij} \)'s for \( i \in \{1\} \cup T, j \in T, i \neq j \) with \( \mathbb{E}[X_{N+2}^2] = \rho_{s} \sqrt{Q} P_{N+2} \) and \( \mathbb{E}[X_{N+2}^2 X_{N+2}^2] = \rho_{s}^2 \sqrt{P_{N+2}^2} P_{N+2} \), \( j \in T \), where \( \rho_{s} \in [-1, 1] \) and \( \rho_{ij} \in [-1, 1] \). Next, we set \( U_j = X_{ij}, j \in T, \) and

\[ X_{N+2} + \frac{P_{N+2}^j}{P_{N+2}^j + 1} X_{N+2} + \rho_{s} \frac{P_{N+2}^j}{P_{N+2}^j + 1} S + \tilde{X}_{N+2} \]

where \( \tilde{X}_{N+2} \sim \mathcal{N}(0, (1 - \rho_{s}^2 - \sum_{j \in T} \rho_{ij}^2) P_{N+2}) \) and it can be easily verified that it is independent of \( S \) and \( \{X_{N+2} \}_{j \in \{1\} \cup T} \). So \( Y_{N+2} \) can be written as

\[ Y_{N+2} = \tilde{X}_{N+2} + \sum_{j \in T} (1 + \rho_{s} \sqrt{P_{N+2}^j}) X_{N+2}^j + \rho_{s} \frac{P_{N+2}^j}{P_{N+2}^j + 1} S + Z_{N+2} \]

With these choices of random variables, we have

\[ R_{1,d} + R_{2,d} \]

\[ I(X_{1j}; Y_{1j} | \{X_{ij}\}_{i \in \{N+2\} \cup T, j \neq i}, \{X_{ij}\}_{i \in \{1\} \cup T, j \neq i}) \]
\[ + I(X_{1j} | \{X_{ij}\}_{i \in \{1\} \cup M^c}, Y_{1j} | \{X_{ij}\}_{i \in \{N+2\} \cup T, j \neq i}, \{X_{ij}\}_{i \in \{1\} \cup T, j \neq i}) \]
\[ = I(X_{ij}, Y_{ij} | \{X_{ij}\}_{j \in M^c} | \{X_{ij}\}_{i \in \{1\} \cup M^c}, \{X_{ij}\}_{i \in \{N+2\} \cup T, j \neq i}, \{X_{ij}\}_{i \in \{1\} \cup T, j \neq i}) \]
\[ \frac{1}{2} \log \left( 1 + \frac{\sum_{j \in M^c} P_{ji}}{N_i} \right) \]

where \( (e) \) holds because \( X_{1j} \) is independent of \( S \) in the special case (8) and \( (f) \) follows from (12) with \( P_{ji} \) being replaced by \( P_{ji}^* \) and going through the same steps as in (12).

Also, for \( R_{2,d}^* \) and \( R_{1,d}^* \) we have

\[ R_{2,d}^* = \frac{1}{2} \log \left( 1 + \frac{\mathbb{E}[X_{N+2}^2]}{N_{N+2}} \right) \]
\[ = \frac{1}{2} \log \left( 1 + \frac{(1 - \rho_{s}^2 - \sum_{j \in T} \rho_{ij}^2) P_{N+2}^2}{N_{N+2}} \right) \]
\[ R_{1,d}^* = h(Y_{N+2} | \{X_{ij} \}_{j \in \{1\} \cup M^c}, Y_{N+2} | \{X_{ij} \}_{j \in \{1\} \cup M^c}) \]
\[ = h(X_{N+2} + \sum_{j \in M^c} (1 + \rho_{s} \sqrt{P_{N+2}^j}) X_{N+2}^j + (1 + \rho_{s} \sqrt{P_{N+2}^j}) S + Z_{N+2}) \]
\[ = \frac{1}{2} \log \left( 1 + \frac{\sum_{j \in M^c} (\sqrt{P_{N+2}^j} + \rho_{s} \sqrt{P_{N+2}^j}) S + Z_{N+2} \right) \]

\[ = \frac{1}{2} \log \left( 1 + \frac{\sum_{j \in M^c} P_{N+2}^j \rho_{s} \sqrt{P_{N+2}^j} + \rho_{s} \sqrt{P_{N+2}^j}) S + Z_{N+2} \right) \]

where \( \Theta = \left( 1 - \rho_{s}^2 - \sum_{j \in T} \rho_{ij}^2 \right) P_{N+2} + (\sqrt{Q} + \rho_{s} \sqrt{P_{N+2}})^2 + N_{N+2} \). This completes our proof. □

The other special case of interest for which we have derived the capacity, is the Gaussian relay network with orthogonal components and no interference at the relays and the destination and with the CSI known noncausally at the source. The channel outputs for this case can be considered as

\[ Y_{r,j,i} = X_{r,j,i} + Z_{r,j,i}, \ r \in T, \ j \neq r \]
\[ Y_{j,i} = X_{j,i} + S_i + Z_{j,i}, \ j \in \{N+2\} \cup T \]
\[ Y_{N+2,i} = X_{N+2,i} + Z_{N+2,i}, \ r \in T \]

where \( Y_{r,j,i} \) and \( Y_{N+2,i} \) are the channel outputs at the relay \( j \) and the destination due to the channel inputs \( X_{r,j} \) and \( X_{r,N+2,i} \) from relay \( r \neq j \), respectively. Also, \( Z_{j,i}^r, s, r \in T, r \neq j \), are the additive Gaussian noise random variables at node \( j, j \in \{N+2\} \cup T \), with zero means and equal variances \( N_j \) all of which are mutually independent and independent of CSI and the channel inputs \( X_1 = \{X_{ij}\}_{i \in T} \) and \( X_r = \{X_{r,j}\}_{j \in \{N+2\} \cup T} \).

**Proposition 2:** (modified version of [5, Theorem 1]) An upper bound on the capacity of DM relay network with orthogonal components and noncausal CSI at the source is given by

\[ C \leq \max \min_{M, C, T} \left\{ \sum_{i \in M} I(X_{ij} | \{X_{ij}\}_{i \in \{N+2\} \cup T, j \neq i}, \{X_{ij}\}_{i \in \{1\} \cup T, j \neq i}) \right\} \]

for some joint probability of the form

\[ p(s, \{x_{ij}\}_{i \in \{1\} \cup T, j \neq i}) \]
\[ \prod_{i \in T} p(y_{N+2} | s, \{x_{ij}\}_{j \in \{N+2\} \cup T}) \]

**Proof:** This bound has been proved in [5] for the case of no CSI \( (S = 0) \) using the cut-set upper bound
The capacity of the Gaussian relay network with orthogonal components, no interference at the relays and destination with CSI known noncausally at the source is
\[
C = \max_{\mathcal{M} \subset \mathcal{T}} \min_{i \in \mathcal{M}} \left\{ \sum_{j \in \{1\} \cup \mathcal{M}} \frac{1}{2} \log(1 + \frac{\beta_i P_j}{N_i}) + \sum_{k \in \{1\} \cup \mathcal{M}} \frac{1}{2} \log(1 + \frac{\beta_{N+2} P_k}{N_{N+2}}) \right\}
\] (27)
where the maximization is over all \( \beta_i \in [0, 1] \) subjected to \( \sum_{i \in \{1\} \cup \mathcal{M}} r_i \neq 0 \), \( k \in \{1\} \cup \mathcal{T} \).

Proof: The achievability part can be proved by setting \( U_j = X_{ij} + \frac{\beta_i P_i}{\beta_j P_j + N_j} S \) in Proposition 1 and applying the model (24) with \( X_{kj} \sim \mathcal{N}(0, \beta_j P_k) \) for \( j \in \{N+2\} \cup \mathcal{T} \), \( k \in \{1\} \cup \mathcal{T} \) which are all independent of \( S \). Also, \( X_{ij} \)'s, \( i \in \mathcal{T} \) are all independent of \( X_{ij} \)'s for \( i \in \mathcal{T} \), \( j \in \{N+2\} \cup \mathcal{T} \), \( j \neq i \). So, we have
\[
R_{1,r_i} = \frac{1}{2} \log(1 + \frac{\beta_i P_i}{N_i}), \quad R_{2,r_i} = \sum_{j \in \mathcal{M}} \frac{1}{2} \log(1 + \frac{\beta_j P_j}{N_j}),
\]
\[
R_{1,d} + R_{2,d} = \frac{1}{2} \sum_{k \in \{1\} \cup \mathcal{M}} \log(1 + \frac{\beta_{N+2} P_k}{N_{N+2}})\quad (28)
\]

For the converse part we make use of the result in Proposition 2, following the same steps as in the converse part of Theorem 2. First, we fix a distribution on \( (S, \{X_{ij}\}_{i \in \mathcal{T}, j \in \{N+2\} \cup \mathcal{T}, i \neq j}, \{Y_{ij}\}_{j \in \{N+2\} \cup \mathcal{T}} \) as
\[
p(s) \prod_{i \in \{1\} \cup \mathcal{T}, j \in \{N+2\} \cup \mathcal{T}, i \neq j} p(x_{ij}) \prod_{i \in \mathcal{T}} p(u_i | s, \{x_{ij}\}_{j \in \{1\} \cup \mathcal{T}, i \neq j})
\]
\[
p(y_{N+2}) p(s, \{x_{ij}\}_{j \in \{N+2\} \cup \mathcal{T}, i \neq j}, (29)
\]
that satisfies \( \mathbb{E}[X_{ij}^2] = \tilde{P}_{ij} \leq P_{ij}, \quad i \in \{1\} \cup \mathcal{T}, \quad j \in \{N+2\} \cup \mathcal{T} \).

Next, we calculate the first term in the relation (25). We can write
\[
I(\{X_{ij}\}_{i \in \{1\} \cup \mathcal{M}}; Y_{N+2}) S, \{X_{ij}\}_{j \in \mathcal{M}, i \neq j}) \quad (a)
\]
\[
= h(\{X_{ij} + Z_{ij}^{(r)}\}_{r \in \{1\} \cup \mathcal{M}, r \neq i}) + h(\{Z_{ij}^{(r)}\}_{r \in \mathcal{M}, r \neq i}) \quad (b)
\]
\[
\leq h(\{X_{ij} + Z_{ij}^{(r)}\}_{r \in \{1\} \cup \mathcal{M}, r \neq i}) \quad (a)
\]
\[
= h(\{X_{ij} + Z_{ij}^{(r)}\}_{r \in \{1\} \cup \mathcal{M}, r \neq i}) \quad (b)
\]
\[
\leq \sum_{r \in \{1\} \cup \mathcal{M}} \frac{1}{2} \log(1 + \frac{\tilde{P}_{ij}}{N_i})\quad (30)
\]
where \( (a) \) holds since removing condition does not reduces entropy and \( (b) \) holds because of the fact that the entropy \( h(\{X_{ij} + Z_{ij}^{(r)}\}_{r \in \{1\} \cup \mathcal{M}, r \neq i}) \) is maximized if \( \{X_{ij}\}_{r \in \{1\} \cup \mathcal{T}}, \{Z_{ij}^{(r)}\}_{r \in \{1\} \cup \mathcal{T}} \) are jointly Gaussian.

Now, the second term in (25) is calculated. We have
\[
I(\{X_{ij}\}_{i \in \{1\} \cup \mathcal{M}}; Y_{N+2}) S, \{X_{ij}\}_{i \in \{1\} \cup \mathcal{T}}) \quad (c)
\]
\[
= h(\{X_{ij} + Z_{ij}^{(r)}\}_{r \in \{1\} \cup \mathcal{M}, r \neq i}) + h(\{Z_{ij}^{(r)}\}_{r \in \mathcal{M}, r \neq i}) \quad (c)
\]
\[
= h(\{X_{ij} + Z_{ij}^{(r)}\}_{r \in \{1\} \cup \mathcal{T}}, \{Z_{ij}^{(r)}\}_{r \in \mathcal{T}}) \quad (d)
\]
\[
\leq \sum_{r \in \{1\} \cup \mathcal{M}} \frac{1}{2} \log(1 + \frac{\tilde{P}_{ij}}{N_i})\quad (31)
\]
where \( (c) \) is deduced in the same way as \( (a) \) and \( (d) \) holds since the entropy \( h(\{X_{ij} + Z_{ij}^{(r)}\}_{r \in \{1\} \cup \mathcal{T}}, \{Z_{ij}^{(r)}\}_{r \in \mathcal{T}}) \) is maximized if \( \{X_{ij}\}_{r \in \{1\} \cup \mathcal{M}}, \{Z_{ij}^{(r)}\}_{r \in \{1\} \cup \mathcal{T}} \) are jointly Gaussian.

Thus, we have
\[
C \leq \min_{\mathcal{M} \subset \mathcal{T}} \left\{ \sum_{i \in \mathcal{M}} \sum_{r \in \{1\} \cup \mathcal{M}} \frac{1}{2} \log(1 + \frac{\tilde{P}_{ij}}{N_i}) + \sum_{r \in \mathcal{T}} \frac{1}{2} \log(1 + \frac{\tilde{P}_{ij}}{N_i}) \right\}
\] (32)

Right hand side of the above relation is an increasing function of \( \tilde{P}_{ij} \) then the maximum happens for the values \( \tilde{P}_{ij} = P_{ij} \). This completes the proof of Theorem 3.

**Corollary:** Our results in Theorem 2 and 3 reduces to that of obtained in [6, Theorem 7] and [6, Corollary 2] as special cases when we assume just one relay i.e. \( \mathcal{T} = \{2\} \).

**IV. Conclusion**

In this paper, the relay network with orthogonal components and noncausal CSI at the source was investigated. For the DM model, a lower bound and an upper bound was developed which met for two special cases of the Gaussian structure of the model. The fist case was the one in which just the intended message for the destination was sent using the CSI and the other case was the one in which all of the source components were encoded against the CSI but there existed no interference at the relays and destination.

**References**

[1] G. Kramer, M. Gastpar and P. Gupta, “Cooperative strategies and capacity theorems for relay networks,” *IEEE Trans. Inf. Theory*, vol. 51, no. 9, Sep. 2005.

[2] G. A. Hodtani and M. R. Aref “New achievable rate and a certain capacity result for a stochastic two network with no interference,” *IET Commun.*, vol. 3, no. 7, pp. 1153-1162, 2009.

[3] A. El-Gamal and S. Zahedi, “Capacity of a class of relay channels with orthogonal components,” *IEEE Trans. Inf. Theory*, vol. 51, no. 5, May 2005.

[4] W. Nam and S. Y. Chung, “Relay networks with orthogonal components,” in *Proc. 46th Annual Allerton Conference on Communication, Control, and Computing*, III, USA, pp. 738-745, Sep. 2008.

[5] L. Gubelli and M. R. Aref, “Capacity of a class of symmetric relay networks with orthogonal components,” *IET Commun.*, vol. 4, no. 18, pp. 2181-2186, 2010.

[6] A. Zaidi, S. Shamai, P. Piantanida and L. Vandendorpe, “Bounds on the capacity of the relay channel with noncausal state at the source,” early access available in *IEEE Trans. Inf. Theory*, no. 99, 2012.
Proof of Theorem 1:
We prove that for any $(2^nR, n)$ code with average probability of error $P^n \to 0$ as $n \to \infty$, the rate $R$ must satisfy (5). By Fano’s inequality, we have $H(W) = H(W)\{Y^n_j\} \leq nR \leq H(W)\{Y^n_j\} \leq nR \leq H(W)\{Y^n_j\} + n\delta_n$

$$nR = H(W) = I(W; \{Y^n_j\} \in S^c \cup (N+2)) + H(W; \{Y^n_j\} \in S^c \cup (N+2))$$

\((a)\)

$$\leq I(W; \{Y^n_j\} \in S^c \cup (N+2)) + n\delta_n$$

$$= I(W, S^n; \{Y^n_j\} \in S^c \cup (N+2)) - I(S^n; \{Y^n_j\} \in S^c \cup (N+2) | W) + n\delta_n$$

$$= \sum_{i=1}^{n} (I(W, S^n; \{Y^n_j\} \in S^c \cup (N+2) | \{Y_j^{i-1}\} \in S^c \cup (N+2)) - H(S^n) + H(S^n | W, \{Y^n_j\} \in S^c \cup (N+2)) + n\delta_n$$

\((b)\)

$$\leq \sum_{i=1}^{n} \left( H(\{Y^n_j\} \in S^c \cup (N+2) | \{X_{j,k,i}\} \in S^c, k \in (N+2) | W, \{Y^n_j\} \in S^c \cup (N+2)) - H(S^n) + H(S^n | W, \{Y^n_j\} \in S^c \cup (N+2)) \right) + n\delta_n$$

\((c)\)

$$\leq \sum_{i=1}^{n} \left( H(\{Y^n_j\} \in S^c \cup (N+2) | \{X_{j,k,i}\} \in S^c, k \in (N+2) | W, \{Y^n_j\} \in S^c \cup (N+2)) - H(S^n) + H(S^n | W, \{Y^n_j\} \in S^c \cup (N+2)) \right)$$

\((d)\)

$$\leq \sum_{i=1}^{n} \left( I(\{X_{j,k,i}\} \in S^c, k \in (N+2) | W, \{Y^n_j\} \in S^c \cup (N+2)) - H(S^n) + H(S^n | W, \{Y^n_j\} \in S^c \cup (N+2)) \right)$$

\((e)\)

$$\leq \sum_{i=1}^{n} \left( I(\{X_{j,k,i}\} \in S^c, k \in (N+2) | W, \{Y^n_j\} \in S^c \cup (N+2)) - H(S^n) + H(S^n | W, \{Y^n_j\} \in S^c \cup (N+2)) \right)$$

\((f)\)

$$\leq \sum_{i=1}^{n} \left( I(\{X_{j,k,i}\} \in S^c, k \in (N+2) | W, \{Y^n_j\} \in S^c \cup (N+2)) - H(S^n) + H(S^n | W, \{Y^n_j\} \in S^c \cup (N+2)) \right)$$

\((g)\)

$$\leq \sum_{i=1}^{n} \left( I(\{X_{j,k,i}\} \in S^c, k \in (N+2) | W, \{Y^n_j\} \in S^c \cup (N+2)) - H(S^n) + H(S^n | W, \{Y^n_j\} \in S^c \cup (N+2)) \right)$$

\((h)\)

$$\leq H(\{Y^n_j\} \in S^c | \{X_{j,k,i}\} \in S^c, k \in (N+2) | W) - H(\{Y^n_j\} \in S^c | W, \{X_{j,k,i}\} \in S^c, k \in (N+2) | W)$$
\[ \begin{align*}
+ & H(Y_{N+2}|\{Y_j\}_{j \in S^c}, \{X_{jk}\}_{j \in S^c, k \in \{N+2\} \cup T}) - H(Y_{N+2}|S, \{Y_j\}_{j \in S^c}, \{X_{jk}\}_{j \in \{1\} \cup T, k \in \{N+2\} \cup T}) \\
& \quad = H(\{Y_j\}_{j \in S^c}|\{X_{jk}\}_{j \in S^c, k \in \{N+2\} \cup T}) - \sum_{j \in S^c} H(Y_j|S, \{X_{jk}\}_{k \in \{1\} \cup S}, \{X_{jk}\}_{k \in S^c}, \{X_{jk}\}_{j \in \{N+2\} \cup T, \kappa \neq j}) \\
& \quad \quad + H(Y_{N+2}|\{Y_j\}_{j \in S^c}, \{X_{jk}\}_{j \in S^c, k \in \{N+2\} \cup T}, - H(Y_{N+2}|S, \{X_{nk+2}\}_{k \in \{1\} \cup T}) \\
(\text{a}) & \leq \sum_{j \in S^c} H(Y_j|\{X_{jk}\}_{k \in S^c, \kappa \neq j}, \{X_{jk}\}_{j \in \{N+2\} \cup T, \kappa \neq j}) - H(Y_j|\{X_{jk}\}_{k \in \{1\} \cup S}, \{X_{jk}\}_{k \in S^c}, \{X_{jk}\}_{j \in \{N+2\} \cup T, \kappa \neq j}) \\
& \quad \quad + H(Y_{N+2}|\{X_{jk}\}_{j \in S^c}) - H(Y_{N+2}|S, \{X_{nk+2}\}_{k \in \{1\} \cup T}) \\
(\text{b}) & \leq \sum_{j \in S^c} I(\{X_{jk}\}_{k \in S, \kappa \neq j}, Y_j|\{X_{jk}\}_{k \in S^c, \kappa \neq j}, \{X_{jk}\}_{j \in \{N+2\} \cup T, \kappa \neq j}) + \sum_{j \in S^c} H(S|\{X_{jk}\}_{k \in \{1\} \cup S}, \{X_{jk}\}_{k \in S^c}, \{X_{jk}\}_{j \in \{N+2\} \cup T}) \\
& \quad \quad + I(\{X_{jk}\}_{j \in S; Y_{N+2}}|\{X_{jk}\}_{j \in S^c}) + I(X_{1N+2}; Y_{N+2}|S, \{X_{nk+2}\}_{k \in T}) \\
(\text{c}) & = \sum_{j \in S^c} I(\{X_{jk}\}_{k \in S, \kappa \neq j}, Y_j|\{X_{jk}\}_{k \in S^c, \kappa \neq j}, \{X_{jk}\}_{j \in \{N+2\} \cup T, \kappa \neq j}) + I(X_{1}; Y_j|S, \{X_{jk}\}_{k \in \{1\} \cup S}, \{X_{jk}\}_{k \in S^c}, \{X_{jk}\}_{j \in \{N+2\} \cup T, \kappa \neq j}) \\
& \quad \quad + I(\{X_{jk}\}_{j \in S; Y_{N+2}}|\{X_{jk}\}_{j \in S^c}) + I(X_{1N+2}; Y_{N+2}|S, \{X_{nk+2}\}_{k \in T})
\end{align*} \]

where (a) follows by Fano's inequality; (b) follows since \(S^n\) is an i.i.d process and that it is independent of message \(W\); (c) follows since \(X_j = \{X_{jk}\}_{j \in S^c, k \in \{N+2\} \cup T}\) is a deterministic function of \(Y_{j-1}\) and that \((W, S^n, Y_{j-1}) \rightarrow (S_i, \{X_{jk}\}_{j \in S^c, k \in \{N+2\}}) \rightarrow Y_j\); (d) follows since \(Y_{j-1} \rightarrow (S_i, W, \{Y_j\}_{j \in T \cup \{N+2\}}) \rightarrow S_i\); (e) follows since removing conditions does not reduce entropy; (f) and (i) follows since \(S_i\) is independent of \(\{X_{jk}\}_{j \in T, k \in \{N+2\} \cup T}\); (j) follows by the common arguments for removing \(n\) letter characterization presented in most of the upper bounds which is also explained in [6, Theorem 4]; (m) follows since removing condition does not reduce entropy; (n) follows from the distribution in (7); (q) follows since removing condition does not reduce entropy and the fact that the joint entropy of random variables is not greater than the entropy of each random variable; (r) follows in the same way as (i).