Achievable Rates and Schedules for Half Duplex Phase Fading Multiple-Relay Networks

Lawrence Ong, Wei Wang and Mehul Motani

Department of Electrical and Computer Engineering
National University of Singapore
E-mail: lawrence.ong@cantab.net, {wang.wei, motani}@nus.edu.sg

Abstract

We present part of our work on the capacity upper bound, achievable rates, and scheduling for the half duplex multiple-relay channel (HD MRC) where every node can either transmit or listen, but not both, at any time. We derive a capacity upper bound based on the cut-set argument, and achievable rates based on the decode-forward coding strategy (DF). We discover that the upper bound and achievable rates are functions of the transmit state vector (a description of which nodes transmit and which receive). More precisely, they are functions of the time fraction of different transmit state vectors, which we term a schedule. We formulate the optimal scheduling problem to find the best schedule, one that maximizes the DF rate. For the phase fading HD MRC, surprisingly, we show that the expressions for the capacity upper bound and for DF rate can be transformed into linear programming problems.

1. Introduction

In the information theory literature, multiple-terminal networks are often studied assuming full duplex (FD) nodes, i.e., the nodes can listen and transmit at the same time on the same frequency band. Most radio frequency transceivers available today, however, either operate in half duplex (HD) mode or in FD mode with the transmitter and the receiver operating on separate frequency bands (the latter can be modeled as orthogonal channels [1]). In view of this hardware limitation, we investigate a more realistic network: the HD network, where every node can either listen or transmit but not both at any time. In this paper, we focus on the half duplex multiple-relay channel (HD MRC) that consists a source, a destination, and multiple relays, which enable us to illustrate concepts in the HD networks that do not appear in the FD counterpart [2, 3, 4].

We first introduce the idea of *transmit state vector* to describe which nodes are transmitting and which nodes are receiving at any time. Take a D-node MRC for example, where the source is always transmitting and the destination is always listening. The total number of transmit state vectors is 2^{D-2}, capturing if each of the $D - 2$ relays is transmitting or listening.

We then derive an upper bound on the capacity based on the cut-set argument and achievable rates based on the decode-forward coding strategy (DF) [5, 2, 3]. We show that the capacity upper bound and achievable DF rates for the HD MRC depend on the time fractions of the transmit state vectors, but not on the sequence or the order. This is counter-intuitive as we do not need to coordinate the transmit-listen sequence among the nodes to maximize the achievable DF rate or to find the capacity upper bound. We view the combination of time fractions of different transmit state vectors as a *schedule* for the network.

We formulate the optimal scheduling problem for the HD MRC to find the schedule that maximizes the DF rate. The optimal scheduling problem turns out to be a max-min problem, which is not easily solved in general. For the phase fading HD MRC, we show that the optimization can be simplified to a linear programming problem. For the general HD MRC, we propose [6] a method for solving the optimal scheduling problem based on minimax hypothesis testing [7, Ch. II.C]. Closed form solutions can be obtained for certain channel conditions.

2. Channel Model

Though the HD single-relay channel (SRC) has been studied [8, 9, 10, 11], the HD MRC has not been studied except for the case where the relays only receive signals from the source and the destination only receives signals from the relays [12, 13]. Here, we investigate the general HD MRC where all nodes can potentially hear all other nodes.

Consider a D-node HD MRC with nodes $\{1, 2, \ldots, D\} \triangleq \mathcal{D}$. Node 1 is the source, node D is the destination, and nodes 2 to $(D - 1)$ are relays.

We define $\mathcal{R}$ as the set of all relays. We assume that the source always transmits and the destination always listens. As not all relays are always needed, we consider an *active relay set* $\mathcal{A} \subseteq \mathcal{R}$ that consists of relays that help the source-destination pair in data transmission, and the set of *unused relays* $\mathcal{U} \triangleq \mathcal{R} \setminus \mathcal{A}$. We define transmit state vector to describe which nodes (in $\mathcal{A}$) are transmitting and which are listening.

Definition 1 *Consider the multiple-relay network and an active relay set* $\mathcal{A} = \{a_1, a_2, \ldots, a_{|\mathcal{A}|}\}$. *The transmit state vector can be expressed as* $\mathbf{s} = (s_1, s_2, \ldots, s_{|\mathcal{A}|}) \in \{L, T\}^{|\mathcal{A}|}$, *where*

$$s_i = \begin{cases} T, & \text{if node } a_i \text{ transmits} \\ L, & \text{if node } a_i \text{ listens} \end{cases}. \quad (1)$$

For $a_i, a_j \in \mathcal{A}, a_i \neq a_j$, *we assume* $a_i > a_j$ *if* $i > j$.

Definition 2 *Consider an active relay set* $\mathcal{A}$. *We define* $\mathcal{T}(\mathbf{s})$ *as the set of all* active relays *that are transmitting, i.e.,* $\{a_i : \text{ all } a_i \in \mathcal{A} \text{ where } s_i = T\}$; *and* $\mathcal{L}(\mathbf{s})$ *as the set of all active relays that are listening, i.e.,* $\{a_j : \text{ all } a_j \in \mathcal{A} \text{ where } s_j = L\}$. *Note that* $\mathcal{T}(\mathbf{s}) \cup \mathcal{L}(\mathbf{s}) = \mathcal{A}$.

Let x_i be the channel input from node i and y_j be the channel output to node j. The channel model of the HD MRC follows that of the FD MRC except for the following. We set $\tilde{x}_i \in \mathcal{X}_i$ to be the "transmit" message of node i when it is in the listening state ($i \in \mathcal{L}(\mathbf{s})$) and set $\tilde{y}_j \in \mathcal{Y}_j$ to be the "received" signal of node j when it is in the transmitting state ($j \in \mathcal{T}(\mathbf{s})$). For unused relays ($i \in \mathcal{U}$), we set them to the listening mode, i.e., $x_i = \tilde{x}_i$.

For the phase fading HD MRC, we set $\tilde{x}_i = 0$ and $\tilde{y}_t = 0, \forall i, t$. So, during transmit state $\mathbf{s}$, node t receives

$$Y_t = \begin{cases} \sum_{i \in (\{1\} \cup \mathcal{T}(\mathbf{s})) \setminus \{t\}} \sqrt{\lambda_{i,t}} e^{j\theta(i,t)} X_i + Z_t \\ \qquad\qquad , \text{ if } t \in \mathcal{L}(\mathbf{s}) \cup \mathcal{U} \cup \{D\} \\ 0 \qquad\qquad , \text{ otherwise} \end{cases}, \quad (2)$$

where X_i is a zero-mean, complex random variable. Z_t, the receiver noise at node t, is an i.i.d., zero-mean, complex, Gaussian random variable with variance $E[Z_t Z_t^\dagger] = N_t$. $Z_t^\dagger$ is the complex-conjugate transpose of Z_t. $\lambda_{i,t} = \kappa$ if $d_{i,t} < 1$, and $\lambda_{i,j} = \kappa d_{i,t}^{-\eta}$ otherwise. Here, $d_{i,t} \geq 0$ is the distance between nodes i and t, $\eta \geq 2$ the attenuation exponent, and κ a positive constant. $e^{j\theta(i,t)}$ is the phase fading random variable, where $\theta(i,t)$ is uniformly distributed over $[0, 2\pi)$. $\theta(i,t)$ for all i and t are jointly independent of each other.

We consider individual-node *per-symbol* power constraint.

$$E[X_i X_i^\dagger] \leq \begin{cases} P_i & , \text{ if } i \in \mathcal{T}(\mathbf{s}) \cup \{1\} \\ 0 & , \text{ otherwise} \end{cases}. \quad (3)$$

We assume that all nodes know $\tilde{x}_i$, $\tilde{y}_t$, κ, $d_{i,t}$, and $\mathbf{s}$, $\forall i, t$, *a priori*. We also assume that node t only knows $\theta(i,t), \forall i$, and does not know any $\theta(i,j)$ for $j \neq t$. Hence, the transmitted signals of node i cannot be chosen as a function of $\theta(i,t)$ for any t.

In this paper, we follow the definitions for block codes, achievable rates, and capacity in [3].

3. Capacity Upper Bound

Modifying the cut-set upper bound for the FD network [14, Ch. 14.10], we can show that an upper bound on the capacity of the phase fading HD MRC is given by the following theorem.

Theorem 1 *The capacity of the D-node phase fading HD MRC is upper bounded by*

$$R_{\text{ub}} = \max_{p(\mathbf{s})} \sum_{\mathbf{s} \in \{L, T\}^{D-2}} \Bigg[p(\mathbf{s})$$
$$\times L \Bigg(\sum_{j \in (\mathcal{B}^c \cap \mathcal{L}(\mathbf{s})) \cup \{D\}} \frac{\sum_{i \in \{1\} \cup (\mathcal{B} \cap \mathcal{T}(\mathbf{s}))} \lambda_{i,j} P_i}{N_j} \Bigg) \Bigg], \quad (4)$$

for all $\mathcal{B} \subseteq \mathcal{R}$ *and* $\mathcal{B}^c = \mathcal{R} \setminus \mathcal{B}$. $\mathcal{R}$ *is the set of all relays,* $\mathcal{A} = \mathcal{R}$, *and* $L(x) = \log(1 + x)$.

Proof 1 (Proof of Theorem 1) *Refer to [6].*

4. Achievability

Using the concept of decode-forward [5, 2, 3], we can show that a lower bound on the capacity of the D-node phase fading HD MRC is given by the following theorem.

Theorem 2 *Consider a D-node phase fading HD MRC, rates up to* R_{DF} *are achievable.*

$$R_{\text{DF}} = \max_{\mathcal{M} \in \Pi(D)} \max_{p(\mathbf{s})} \min_{m_t \in \mathcal{M} \setminus \{1\}} \sum_{\substack{\mathbf{s} \in \{L, T\}^{|\mathcal{M}|-2} \\ m_t \in \mathcal{L}(\mathbf{s}) \cup \{D\}}} \Bigg[p(\mathbf{s})$$
$$\times L \Bigg(\frac{\sum_{i \in \{1\} \cup (\{m_2, \ldots, m_{t-1}\} \cap \mathcal{T}(\mathbf{s}))} \lambda_{i,m_t} P_i}{N_{m_t}} \Bigg) \Bigg]. \quad (5)$$

Here, $\mathcal{M} = \{m_1 = 1, m_2, \ldots, m_{|\mathcal{M}|} = D\}$ is the route [15, 16], i.e., an ordered set of nodes from the source to the destination. $\Pi(\mathcal{D})$ is the set of all possible routes from the source to the destination. We assume that nodes not in route $\mathcal{M}$ do not transmit. So, $\mathcal{A} = \{m_2, m_3, \ldots, m_{|\mathcal{M}|-1}\}$.

Remark 1 *The first maximization is over all possible route selections. The second maximization is over all possible schedule $p(\mathbf{s})$. The minimization is over all relay nodes in the route, since using DF, each of them needs to fully decode the source messages.*

Proof 2 (Sketch of Proof for Theorem 2) *Refer to Appendix A.*

5. Optimal DF Scheduling Problem

We note that the capacity upper bound and DF rates are function of the time fraction of the transmit state vectors. Hence, we define a schedule of the HD network as follow.

Definition 3 *A schedule for the HD network is the probability mass function of the transmit state vector.*

We formulate the optimal DF scheduling problem, to find a schedule that maximizes the DF rate. Consider a D-node phase fading HD MRC. On some route $\mathcal{M} \in \Pi(\mathcal{D})$, an optimal DF schedule is a probability mass function $p^*(\mathbf{s})$ on $\{L,T\}^{|\mathcal{M}|-2}$ such that

$$p^*(\mathbf{s}) \in \underset{p(\mathbf{s})}{\mathrm{argmax}} \min_{m_t \in \mathcal{M}\setminus\{1\}} \sum_{\substack{\mathbf{s}\in\{L,T\}^{|\mathcal{M}|-2}\\ m_t\in\mathcal{L}(\mathbf{s})\cup\{D\}}}$$

$$\left[p(\mathbf{s})L\left(\frac{\sum_{i\in\{1\}\cup(\{m_2,\dots,m_{t-1}\}\cap\mathcal{T}(\mathbf{s}))} \lambda_{i,m_t}P_i}{N_{m_t}} \right) \right]. \quad (6)$$

We note that a valid schedule must satisfy the following conditions. $p(\mathbf{s}) \geq 0$ for all $\mathbf{s}$, and $\sum p(\mathbf{s}) = 1$.

Remark 2 *In the optimal scheduling problem here, we assume that a route already is chosen. For route selection, the reader can refer to [15, 16].*

6. Example: Four-Node HD MRC

In this section, we find the capacity upper bound, the optimal DF rate and schedule for the four-node phase fading HD MRC.

6.1. Linear Programming

For the phase fading channel, we can convert the optimal scheduling problem to a linear programming problem. We use the example of $D = 4$ and route $\mathcal{M} = \{1,2,3,4\}$. For this route, $\mathcal{A} = \{2,3\}$ and there are four transmit state vectors with the following probability: $p(L,L) = p_0$, $p(L,T) = p_1$, $p(T,L) = p_2$, and $p(T,T) = p_3$. (L,T) means that node 2 is listening and node 3

is transmitting, and so on. We introduce an auxiliary variable u. The DF rate can be written as

$$R_{\mathrm{DF}} = \max u, \quad (7)$$

subject to

$$p_0 L(\lambda_{1,2}P_1/N_2) + p_1 L(\lambda_{1,2}P_1/N_2) - u \geq 0,$$
$$p_0 L(\lambda_{1,3}P_1/N_3) + p_2 L((\lambda_{1,3}P_1 + \lambda_{2,3}P_2)/N_3) - u \geq 0,$$
$$p_0 L(\lambda_{1,4}P_1/N_4) + p_1 L((\lambda_{1,4}P_1 + \lambda_{3,4}P_3)/N_4)$$
$$\quad + p_2 L((\lambda_{1,4}P_1 + \lambda_{2,4}P_2)/N_4)$$
$$\quad + p_3 L((\lambda_{1,4}P_1 + \lambda_{2,4}P_2 + \lambda_{3,4}P_3)/N_4) - u \geq 0,$$
$$p_0 + p_1 + p_2 + p_3 = 1,$$
$$p_0 \geq 0, \quad p_1 \geq 0, \quad p_2 \geq 0, \quad p_3 \geq 0, \quad u \geq 0.$$

We can solve the above linear programming optimization problem using the simplex method [17, Ch. 3]. Let the solution be $(p_0^*, p_1^*, p_2^*, p_3^*, u^*)$. $(p_0^*, p_1^*, p_2^*, p_3^*)$ is an optimal schedule and u^* is the DF rate.

Similarly, we can convert the capacity upper bound (4) to a linear programming problem. By converting these optimizations, (4) and (5), to linear programming problems, we are able to compute upper bounds, DF rates, and optimal DF schedules for large networks with more than ten relays.

Remark 3 *For channels of which the optimal scheduling problem cannot be converted to linear programming problem, we proposed [6] a method to find optimal schedules based on minimax hypothesis testing. Closed form solutions can be obtained using this method, for certain channel topologies.*

6.2. Computation Results

Now, we show computation results for the four-node HD MRC in the line topology, with coordinates: node 1 $(0,0)$, node 2 $(0,y_2)$, node 3 $(0,y_3)$, and node 4 $(0,100)$. We assume that $\kappa = 1$, $\eta = 2$, $P_i = 10, \forall i$, and $N_j = 0.01, \forall j$.

Fig. 1 shows DF rates (R_{DF}) for the four-node phase fading HD MRC with varying relay positions. Fig. 2 shows capacity upper bounds (R_{ub}) and the difference between upper bounds and DF rates ($R_{\mathrm{ub}} - R_{\mathrm{DF}}$).

The highest DF rate is 0.5956 bits/channel use, when $y_2 = 34, y_3 = 56$. The optimal schedule for this topology is $p_0 = 0.0038, p_1 = 0.6585, p_2 = 0.3377, p_3 = 0$. The worst DF rate is 0.1375 bits/channel use, when $y_2 = 100$ (relay 2 at the destination). The optimal schedule for this topology is $p_0 = 1, p_1 = p_2 = p_3 = 0$, i.e., nodes 2 and 3 only listen. For the capacity upper bound, the highest rate is 0.7080 bits/channel use,
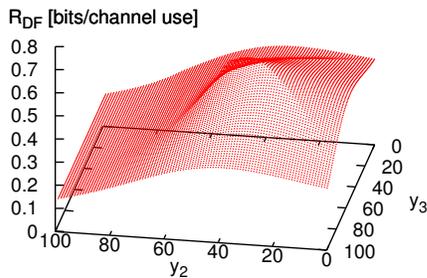
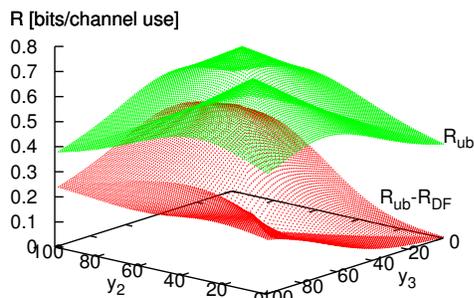Figure 1: DF rates for the four-node MRC with varying relay positions

Figure 2: Cut-set bounds and the gap between upper bounds and DF rates.

when $y_2 = 35, y_3 = 65$ or $y_2 = 65, y_3 = 35$ (note the symmetry). The lowest rate is 0.3757 bits/channel use, when $y_2 = 0, y_3 = 0$ or $y_2 = 100, y_3 = 100$, i.e., when both relays are at the source or at the destination.

Unlike the FD case where DF achieves the cut-set upper bound and hence achieves the capacity when the relays are within certain distances from the source [3], DF in the HD case does not achieve (though close to) the cut-set bound if a relay is ϵ away from the source, for any arbitrarily small distance $\epsilon > 0$.

References

[1] A. A. El Gamal and S. Zahedi, "Capacity of a class of relay channels with orthogonal components," *IEEE Trans. Inf. Theory*, vol. 51, no. 5, pp. 1815–1817, May 2005.

[2] L. Xie and P. R. Kumar, "An achievable rate for the multiple level relay channel," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1348–1358, Apr. 2005.

[3] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3037–3063, Sep. 2005.

[4] L. Ong and M. Motani, "Myopic coding in multiterminal networks," *IEEE Trans. Inf. Theory*, vol. 54, no. 7, pp. 3295–3314, Jul. 2008.

[5] T. M. Cover and A. A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. IT-25, no. 5, pp. 572–584, Sep. 1979.

[6] L. Ong, W. Wang, and M. Motani, "Achievable rates and optimal schedules for half duplex multiple-relay networks," Proc. 46th Allerton Conf. on Commun., Control, and Comput., Monticello, IL, Sep. 23-26 2008.

[7] H. V. Poor, *An introduction to signal detection and estimation.* New York: Springer-Verlag, 1994.

[8] A. Host-Madsen, "On the capacity of wireless relaying," in *Proc. of IEEE Semiannual Vehicular Technology Conference (VTC)*, Birmingham, AL, May 6-9 2002, pp. 1333–1337.

[9] M. A. Khojastepour, A. Sabharwal, and B. Aazhang, "On the capacity of 'cheap' relay networks," in *Proc. Conf. on Inf. Science and Syst. (CISS)*, Baltimore, MD, Mar. 12-14 2003.

[10] G. Kramer, "Models and theory for relay channels with receive constraints," in *Proc. 42nd Allerton Conf. on Commun., Control, and Comput.*, Monticello, IL, Sep. 29-Oct. 1 2004.

[11] A. Host-Madsen and J. Zhang, "Capacity bounds and power allocation for the wireless relay channel," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 2020–2040, Jun. 2005.

[12] M. Gastpar and M. Vetterli, "On the capacity of large gaussian relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 3, pp. 765–, Mar. 2005.

[13] F. Xue and S. Sandhu, "Cooperation in a half-duplex gaussian diamond relay chanenl," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3806–3814, Oct. 2007.

[14] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* John Wiley and Sons, 1991.

[15] L. Ong and M. Motani, "Optimal routing for the decode-and-forward strategy in the Gaussian multiple relay channel," in *Proc. IEEE Int. Symposium on Inf. Theory (ISIT)*, Nice, France, Jun. 24-29 2007, pp. 1061–1065.

[16] ——, "Optimal routing for decode-and-forward based cooperation in wireless network," in *Proc. 4th Annu. IEEE Conf. on Sensor and Ad Hoc Commun. and Networks (SECON)*, San Diego, CA, Jun. 18-21 2007, pp. 334–343.

[17] D. G. Luenberger, *Linear and Nonlinear Programming.* Addison-Wesley, 1984.

A. Proof of Theorem 2

We will prove achievable rates for the HD MRC for $D = 4$ for the route $\mathcal{M} = \{1, 2, 3, 4\}$. The generalization to Theorem 2, i.e., any D and any route, is straight forward but tedious, and hence omitted in this paper. We first prove the following lemma.

Lemma 1 *Consider a four-node HD MRC. On route $\{1, 2, 3, 4\}$, rates up to the following are achievable.*

$$R_{\text{DF}} = \max_{p(\mathbf{s})} \max_{p(x_1, x_2, x_3 | \mathbf{s})} \min \left\{ R_2, R_3, R_4 \right\}, \quad (9)$$

where

$$R_2 = p(L, L) I(X_1; Y_2 | X_2, X_3, \mathbf{S} = (L, L))$$
$$+ p(L, T) I(X_1; Y_2 | X_2, X_3, \mathbf{S} = (L, T)) \quad (10a)$$

$$R_3 = p(L, L) I(X_1; Y_3 | X_2, X_3, \mathbf{S} = (L, L))$$
$$+ p(T, L) I(X_1, X_2; Y_3 | X_3, \mathbf{S} = (T, L)) \quad (10b)$$

$$R_4 = p(L, L) I(X_1; Y_4 | X_2, X_3, \mathbf{S} = (L, L))$$
$$+ p(L, T) I(X_1, X_3; Y_4 | X_2, \mathbf{S} = (L, T))$$
$$+ p(T, L) I(X_1, X_2; Y_4 | X_3, \mathbf{S} = (T, L))$$
$$+ p(T, T) I(X_1, X_2, X_3; Y_4 | \mathbf{S} = (T, T)). \quad (10c)$$

Proof 3 (Proof of Lemma 1) *We consider block coding and let n be the length of a block. We define $n_0 = np(L, L)$, $n_1 = np(L, T)$, $n_2 = np(T, L)$, and $n_3 = np(T, T)$. We assume that n_i, for $i = 0, 1, 2, 3$, are positive integers. Now, we describe the processes of codebook generation, encoding, and decoding. These steps are similar to that for the FD MRC [2, 3]. Here, we highlight the differences.*

Generating Random Codes:

1. *Fix $\tilde{x}_1, \tilde{x}_2, \tilde{x}_3$, and the probability distribution function $p(x_1, x_2, x_3 | \mathbf{s}) = p(x_3 | \mathbf{s}) p(x_2 | x_3, \mathbf{s}) p(x_1 | x_2, x_3, \mathbf{s})$ with transmitting constraints for each $\mathbf{s}$.*

2. *(For node 3) For $\mathbf{s} = (L, T)$, randomly generate 2^{nR} independent and identically distributed (i.i.d.) n_1-sequences in $\mathcal{X}_3^{n_1}$, each drawn according to $p(x_3^{n_0+1:n_0+n_1} | \mathbf{s} = (L, T)) = \prod_{i=n_0+1}^{n_0+n_1} p(x_{3,i} | \mathbf{s} = (L, T))$. Index the sub-codewords as $x_3^{n_0+1:n_0+n_1}(z_3), z_3 \in \{1, 2, \ldots, 2^{nR}\}$. We use the notation $x_t^{i:j} = x_{t,i}, x_{t,i+1}, \ldots, x_{t,j}$ for $i \leq j$.*

3. *(For node 3) For $\mathbf{s} = (T, T)$, randomly generate 2^{nR} independent and identically distributed (i.i.d.) n_3-sequences in $\mathcal{X}_3^{n_3}$, each drawn according to $p(x_3^{n_0+n_1+n_2+1:n} | \mathbf{s} = (T, T)) = \prod_{i=n_0+n_1+n_2+1}^{n} p(x_{3,i} | \mathbf{s} = (T, T))$. Index the sub-codewords as $x_3^{n_0+n_1+n_2+1:n}(z_3), z_3 \in \{1, 2, \ldots, 2^{nR}\}$.*

4. *(For node 3) Now, fill in the rest of the indices in the codewords by setting $x_{3,j} = \tilde{x}_3$ for $j = 1, 2, \ldots, n_0, n_0 + n_1 + 1, n_0 + n_1 + 2, \ldots, n_0 + n_1 + n_2$. Index the codewords as $\mathbf{x}_3(z_3), z_3 \in \{1, 2, \ldots, 2^{nR}\}$. See Tab. 1 for the codewords.*

5. *(For node 2) For $\mathbf{s} = (T, L)$ randomly generate 2^{nR} independent and identically distributed (i.i.d.) n_2-sequences in $\mathcal{X}_2^{n_2}$, each drawn according to $p(x_2^{n_0+n_1+1:n_0+n_1+n_2} | \tilde{x}_3^{n_2}, \mathbf{s} = (T, L)) = \prod_{i=n_0+n_1+1}^{n_0+n_1+n_2} p(x_{2,i} | \tilde{x}_3, \mathbf{s} = (T, L))$. Index the sub-codewords as $x_2^{n_0+n_1+1:n_0+n_1+n_2}(z_2), z_2 \in \{1, 2, \ldots, 2^{nR}\}$.*

6. *(For node 2) For $\mathbf{s} = (T, T)$ and for each $x_3^{n_0+n_1+n_2+1:n}(z_3)$, randomly generate 2^{nR} independent and identically distributed (i.i.d.) n_3-sequences in $\mathcal{X}_2^{n_3}$, each drawn according to $p(x_2^{n_0+n_1+n_2+1:n} | x_3^{n_0+n_1+n_2+1:n}(z_3), \mathbf{s} = (T, T)) = \prod_{i=n_0+n_1+n_2+1}^{n} p(x_{2,i} | x_{3,i}(z_3), \mathbf{s} = (T, T))$. Index the sub-codewords as $x_2^{n_0+n_1+n_2:n}(z_2 | z_3), z_2 \in \{1, 2, \ldots, 2^{nR}\}$.*

7. *(For node 2) Now, fill in the rest of the indices in the codewords by setting $x_{2,j} = \tilde{x}_2$ for $j = 1, 2, \ldots, n_0 + n_1$. Index the codewords as $\mathbf{x}_2(z_2 | z_3), z_2 \in \{1, 2, \ldots, 2^{nR}\}$. Refer to Tab. 1 for the codewords.*

8. *(For node 1) The similar technique is used to create codewords $\mathbf{x}_1(z_1 | z_2, z_3), z_1 \in \{1, 2, \ldots, 2^{nR}\}$, by combining sub-codewords from the four transmit state vectors (refer to Tab. 1).*

Encoding:

1. *Assume that at the beginning of block t, node $i, i = 1, 2, 3, 4$, has correctly decoded messages $w^{t-(i-1)} = (w_1, w_2, \ldots, w_{t-(i-1)})$.*

2. *As shown in Tab. 1, in block t, node 1 transmits $\mathbf{x}_1(w_t | w_{t-1}, w_{t-2})$, node 2 $\mathbf{x}_2(w_{t-1} | w_{t-2})$, and node 3 $\mathbf{x}_3(w_{t-2})$.*

Decoding and Achievable Rates:

1. *Recall that we assume that node $i, i = 1, 2, 3, 4$, has correctly decoded messages $w^{t-(i-1)} = (w_1, w_2, \ldots, w_{t-(i-1)})$ at the beginning of block t.*

2. *In block t, node 2 receives $\mathbf{y}_2$ where $y_{2,i} = \tilde{y}_2$ for $i = n_0 + n_1 + 1, \ldots, n$. These are the "received" signals when node 2 is in the transmitting state. Hence, it uses the first $n_0 + n_1$ received signals for decoding.*
   For $1 \leq i \leq n_0$, node 2 knows that $x_{2,i} = \tilde{x}_2$ and $x_{3,i} = \tilde{x}_3$. It finds $\mathcal{G}_{2,1}$, the set of all w_t such

Table 1: A transmission scheme for the four-node HD MRC.

Codeword	State (L,L)			State (L,T)			State (T,L)			State (T,T)		
Block t												
$\mathbf{x}_1(w_t\|w_{t-1},w_{t-2})$	$x_{1,1}$	$\cdots$	x_{1,n_0}	x_{1,n_0+1}	$\cdots$	x_{1,n_0+n_1}	x_{1,n_0+n_1+1}	$\cdots$	$x_{1,n_0+n_1+n_2}$	$x_{1,n_0+n_1+n_2+1}$	$\cdots$	$x_{1,n}$
$\mathbf{x}_2(w_{t-1}\|w_{t-2})$	$\tilde{x}_2$	$\cdots$	$\tilde{x}_2$	$\tilde{x}_2$	$\cdots$	$\tilde{x}_2$	x_{2,n_0+n_1+1}	$\cdots$	$x_{2,n_0+n_1+n_2}$	$x_{2,n_0+n_1+n_2+1}$	$\cdots$	$x_{2,n}$
$\mathbf{x}_3(w_{t-2})$	$\tilde{x}_3$	$\cdots$	$\tilde{x}_3$	x_{3,n_0+1}	$\cdots$	x_{3,n_0+n_1}	$\tilde{x}_3$	$\cdots$	$\tilde{x}_3$	$x_{3,n_0+n_1+n_2+1}$	$\cdots$	$x_{3,n}$
Block $t+1$												
$\mathbf{x}_1(w_{t+1}\|w_t,w_{t-1})$	$x_{1,1}$	$\cdots$	x_{1,n_0}	x_{1,n_0+1}	$\cdots$	x_{1,n_0+n_1}	x_{1,n_0+n_1+1}	$\cdots$	$x_{1,n_0+n_1+n_2}$	$x_{1,n_0+n_1+n_2+1}$	$\cdots$	$x_{1,n}$
$\mathbf{x}_2(w_t\|w_{t-1})$	$\tilde{x}_2$	$\cdots$	$\tilde{x}_2$	$\tilde{x}_2$	$\cdots$	$\tilde{x}_2$	x_{2,n_0+n_1+1}	$\cdots$	$x_{2,n_0+n_1+n_2}$	$x_{2,n_0+n_1+n_2+1}$	$\cdots$	$x_{2,n}$
$\mathbf{x}_3(w_{t-1})$	$\tilde{x}_3$	$\cdots$	$\tilde{x}_3$	x_{3,n_0+1}	$\cdots$	x_{3,n_0+n_1}	$\tilde{x}_3$	$\cdots$	$\tilde{x}_3$	$x_{3,n_0+n_1+n_2+1}$	$\cdots$	$x_{3,n}$
Block $t+2$												
$\mathbf{x}_1(w_{t+2}\|w_{t+1},w_t)$	$x_{1,1}$	$\cdots$	x_{1,n_0}	x_{1,n_0+1}	$\cdots$	x_{1,n_0+n_1}	x_{1,n_0+n_1+1}	$\cdots$	$x_{1,n_0+n_1+n_2}$	$x_{1,n_0+n_1+n_2+1}$	$\cdots$	$x_{1,n}$
$\mathbf{x}_2(w_{t+1}\|w_t)$	$\tilde{x}_2$	$\cdots$	$\tilde{x}_2$	$\tilde{x}_2$	$\cdots$	$\tilde{x}_2$	x_{2,n_0+n_1+1}	$\cdots$	$x_{2,n_0+n_1+n_2}$	$x_{2,n_0+n_1+n_2+1}$	$\cdots$	$x_{2,n}$
$\mathbf{x}_3(w_t)$	$\tilde{x}_3$	$\cdots$	$\tilde{x}_3$	x_{3,n_0+1}	$\cdots$	x_{3,n_0+n_1}	$\tilde{x}_3$	$\cdots$	$\tilde{x}_3$	$x_{3,n_0+n_1+n_2+1}$	$\cdots$	$x_{3,n}$

that $(x_1^{1:n_0}(w_t|w_{t-1},w_{t-2}),x_2^{1:n_0},x_3^{1:n_0},y_2^{1:n_0})$ are jointly ϵ-typical [14, p. 384].

For $n_0+1 \leq i \leq n_0+n_1$, node 2 knows that $x_{2,i} = \tilde{x}_2$ and $x_3^{n_0+1:n_0+n_1}(w_{t-2})$ (since node 2 has already decoded w_{t-2} and it knows node 3's codebooks). It finds $\mathcal{G}_{2,2}$, the set of all w_t such that $(x_1^{n_0+1:n_0+n_1}(w_t|w_{t-1},w_{t-2}),x_2^{n_0+1:n_0+n_1}, x_3^{n_0+1:n_0+n_1}(w_{t-2}),y_2^{n_0+1:n_0+n_1})$ are jointly ϵ-typical.

Node 2 then estimates $\hat{w}_t$ by finding a unique $w \in \mathcal{G}_{2,1} \cap \mathcal{G}_{2,2}$. It can be shown [5, 4] that if

$$nR \leq I(X_1^{1:n_0};Y_2^{1:n_0}|X_2^{1:n_0},X_3^{1:n_0},\mathbf{S}=(L,L))$$
$$+ I(X_1^{n_0+1:n_0+n_1};Y_2^{n_0+1:n_0+n_1}|$$
$$X_2^{n_0+1:n_0+n_1},X_3^{n_0+1:n_0+n_1},\mathbf{S}=(L,T)) \tag{11a}$$

$$= n_0 I(X_1,Y_2|X_2,X_3,\mathbf{S}=(L,L))$$
$$+ n_1 I(X_1,Y_2|X_2,X_3,\mathbf{S}=(L,T)) \tag{11b}$$

$$R \leq p_0 I(X_1,Y_2|X_2,X_3,\mathbf{S}=(L,L))$$
$$+ p_1 I(X_1,Y_2|X_2,X_3,\mathbf{S}=(L,T)), \tag{11c}$$

$\hat{w}_t = w_t$ with arbitrarily small error-probability.

3. *Node 3 listens during (L,L) and (T,L). It uses two blocks of received signals, i.e., blocks t and $t+1$, to decodes w_t at the end of block $t+1$, whereby node 3 has decoded messages up to w^{t-1}. For $1 \leq i \leq n_0$ in block t, node 3 knows that $x_{2,i} = \tilde{x}_2$ and $x_{3,i} = \tilde{x}_3$. It finds $\mathcal{G}_{3,1}$, the set of all w_t such that $(x_1^{1:n_0}(w_t|w_{t-1},w_{t-2}),x_2^{1:n_0},x_3^{1:n_0},y_3^{1:n_0})$ are jointly ϵ-typical.*

For $n_0+n_1+1 \leq i \leq n_0+n_1+n_2$ in block t, node 3 knows $x_{3,i} = \tilde{x}_3$ and $x_2^{n_0+n_1+1:n_0+n_1+n_2}$. It finds $\mathcal{G}_{3,2}$, the set of all w_t such that $(x_1^{n_0+n_1+1:n_0+n_1+n_2}(w_t|w_{t-1}),x_2^{n_0+n_1+1:n_0+n_1+n_2}(w_{t-1}), x_3^{n_0+n_1+1:n_0+n_1+n_2},y_3^{n_0+n_1+1:n_0+n_1+n_2})$ are jointly ϵ-typical.

For $n_0+n_1+n_2+1 \leq i \leq n$ in block $t+1$, node 3 knows $x_{3,i} = \tilde{x}_3$. It finds $\mathcal{G}_{3,3}$, the set of all w_t such that $(x_2^{n_0+n_1+1:n_0+n_1+n_2}(w_t),$

$x_3^{n_0+n_1+1:n_0+n_1+n_2},y_3^{n_0+n_1+1:n_0+n_1+n_2})$ are jointly ϵ-typical.

Node 3 then estimates $\hat{w}_t$ by finding a unique $w \in \mathcal{G}_{3,1} \cap \mathcal{G}_{3,2} \cap \mathcal{G}_{3,3}$. If

$$nR$$
$$\leq I(X_1^{1:n_0};Y_3^{1:n_0}|X_2^{1:n_0},X_3^{1:n_0},\mathbf{S}=(L,L))$$
$$+ I(X_1^{n_0+n_1+1:n_0+n_1+n_2};Y_3^{n_0+n_1+1:n_0+n_1+n_2}|$$
$$X_2^{n_0+n_1+1:n_0+n_1+n_2},X_3^{n_0+n_1+1:n_0+n_1+n_2},$$
$$\mathbf{S}=(T,L))$$
$$+ I(X_2^{n_0+n_1+1:n_0+n_1+n_2};Y_3^{n_0+n_1+1:n_0+n_1+n_2}|$$
$$X_3^{n_0+n_1+1:n_0+n_1+n_2},\mathbf{S}=(T,L)) \tag{12a}$$
$$= n_0 I(X_1,Y_3|X_2,X_3,\mathbf{S}=(L,L))$$
$$+ n_2 I(X_1,X_2;Y_3|X_3,\mathbf{S}=(T,L)) \tag{12b}$$
$$R \leq p_0 I(X_1,Y_3|X_2,X_3,\mathbf{S}=(L,L))$$
$$+ p_2 I(X_1,X_2;Y_3|X_3,\mathbf{S}=(T,L)), \tag{12c}$$

$\hat{w}_t = w_t$ with arbitrarily small error-probability.

4. *By a similar decoding method, we can show that node 4 can decodes $\hat{w}_t = w_t$ with arbitrarily small error-probability if (10c) is satisfied.*

Combining these inequalities, we have Lemma 1. ∎

The generalization of Lemma 1 to Theorem 2 is straightforward. We modify the DF rate derivation for the D-node FD MRC [2, 3] using the same technique in this section to obtain the DF rate for the D-node HD MRC on route $\{1,2,\ldots,D\}$. Next, for route $\mathcal{M}$ on the D-node HD MRC, where $|\mathcal{M}| < D$, the DF rate is the same as that of the $|\mathcal{M}'|$-node HD MRC on route $\mathcal{M}'$, where $|\mathcal{M}'| = |\mathcal{M}|$, with node re-indexing and the assumption that all nodes not in the route do not transmit. This gives us the DF rate for any route on the D-node HD MRC. Since the maximum DF rate in the phase fading channel can be attained by independent Gaussian inputs [3], we get Theorem 2.