Optimality of Energy-Efficient Scheduling and Relaying for Half-Duplex Relay Networks

Sarthak Jain, Martina Cardone, Soheil Mohajer

University of Minnesota, Minneapolis, MN 55455, USA,
Email: {jain0122, cardo089, soheil}@umn.edu

***Abstract*—This paper considers a single-source single-destination half-duplex n-relay network with arbitrary topology, where the source communicates with the destination through a direct link and with the help of n half-duplex relays. The focus is on the linear deterministic approximation of the Gaussian noise network model. First, sufficient conditions under which operating the network in the $n+1$ *energy-efficient* states (out of the 2^n possible states) is sufficient to achieve the approximate capacity (that is, an additive gap approximation of the Shannon capacity) are characterized. Specifically, these $n+1$ energy-efficient states are those in which at most one relay is in transmit mode while the rest of the relays are in receive mode. Under such sufficient network conditions, closed-form expressions for the scheduling and the approximate capacity are provided. Then, a time-block relaying scheme, where at each point in time at most one relay is in transmit mode, is designed. In particular, the designed relaying scheme leverages information flow preservation at each relay to explicitly provide the information that each relay is exclusively responsible to store and forward to the destination. Furthermore, the destination can decode the information bits sent by the source in block B by the end of block $B+1$, and the proposed scheme is shown to achieve the approximate capacity whenever the sufficient conditions are satisfied. Such features make the designed scheme relevant for practical use.**

I. Introduction

Today, a massive number of critical services, such as healthcare and education, heavily relies on the use of the wireless medium. Such services, together with the staggering amount of data traffic generated every day, are currently fueling a transformative wireless revolution. *Relaying* is foreseen to be integrated in several technology components of the next generation IoT networks and evolving 5G architecture, hence playing a vital role in this wireless revolution. For instance, relaying promises performance enhancement of device-to-device communication [3], [4], millimeter wave communication [5], [6], vehicular communication [7], [8], and unmanned aerial vehicles communication [9], [10], which are all key components of the 5G architecture.

Depending on their mode of operation, relays can be classified into two main categories: (1) *full-duplex* when a relay can transmit and receive over the same time/frequency bands; and (2) *half-duplex* when a relay must use different times/bands for transmitting and receiving. Despite the better performance (e.g., throughput) promised by full-duplex compared to half-duplex, several practical restrictions arise when a node operates in full-duplex, among all how to properly cancel the self-interference. Although in the past few years several Self-Interference Cancellation (SIC) techniques have been developed (e.g., antenna separation, analog cancellation and digital cancellation) [11]–[13], their physical layer robustness is yet to be demonstrated in many operating scenarios. Examples of other practical restrictions of the full-duplex technology are given by currently available prototypes, which are large and complicated [14], [15], and by the severe energy consumption required by SIC techniques, which represents an implementation burden, especially in scenarios where low-cost communication modules are needed and nodes have limited power supply. Given such considerations, half-duplex is expected to continue to play a fundamental role in next generation wireless networks [16].

In this work, we investigate the optimality of an *energy-efficient schedule* (which, as described later in details, at each point on time schedules at most one relay for transmission) for half-duplex relay networks with arbitrary topology, that is, all the network nodes may be connected between one another. In particular, we derive sufficient conditions under which such energy-efficient scheduling suffices to operate the network close to its Shannon capacity (i.e., the supremum rate at which information can be transmitted), and we design a provable optimal time-block relaying scheme when such an energy-efficient schedule is used.

A. Related Work

Characterizing the Shannon capacity of wireless relay networks is a long-standing open problem, even for networks with a single relay. The cut-set bound has been shown to offer a constant (i.e., which only depends on the number of relays n) additive gap approximation of the Shannon capacity for Gaussian relay networks [17]–[21].

In half-duplex, such an approximation (referred to as *approximate capacity* throughout the paper) can be computed by solving a linear program that involves 2^n cut constraints and 2^n state variables corresponding to the receive/transmit configurations of the n half-duplex relays[1]. It has been surprisingly shown that it suffices to operate the network in only $n+1$ states out of the 2^n possible ones to achieve the approximate

This research was supported in part by NSF under Award #1907785. Part of this work was presented at the 2020 IEEE Information Theory Workshop [1] and at the 2021 IEEE International Symposium on Information Theory [2].

[1]It has been shown that *randomly* switching between the 2^n states can offer an increase in the information rate that can be transmitted over the network [22]–[24]. However, with n half-duplex relays, such as an increase is at most equal to n and hence, it only contributes to the additive gap.

capacity [25]. This result is promising as it implies that rates close to the Shannon capacity can be achieved with a *linear* (instead of exponential) number of states. However, *can a set of (at most) $n+1$ states sufficient to characterize the approximate capacity be found in polynomial time in n?* To the best of our knowledge, this question is open for half-duplex relay networks with arbitrary topology. In particular, such a set of *critical* states and the approximate capacity can be computed in polynomial time only for networks with a special topology, such as line networks [26], a special class of layered networks [27], and diamond networks (where information is hopped though one layer of n non-communicating relays) either with $n = 2$ relays [28] or with an arbitrary n under certain network conditions expressed in form of matrix determinants [29]. In this work, we consider unicast (i.e., single-source, single-destination) half-duplex relay networks with an *arbitrary* topology, and we derive network conditions under which it suffices to operate the network in the $n+1$ states in which *at most* one relay transmits (and all the remaining $n-1$ relays receive) to achieve the network approximate capacity. To the best of our knowledge, this is the first work that provides network conditions under which a specific set of at most $n+1$ states suffices to characterize the approximate capacity for half-duplex relay networks with arbitrary topology.

B. Contributions

In this work, we consider unicast half-duplex relay networks with *arbitrary* topology, where a source communicates with a destination with the help of n relays that operate in half-duplex. In particular, we analyze the *linear deterministic approximation* of the Gaussian noise channel, a.k.a. ADT model [17], which captures - in a simple deterministic way - the interaction between interfering signals and neglects the noise. The significance of this model stems mainly from the facts that: (1) its approximate capacity offers a constant gap approximation for the capacity of the original noisy Gaussian network for several relevant topologies, such as line and diamond networks; and (2) its study has been shown to offer suitable guidelines on how to operate the original noisy Gaussian network so as to achieve rates close to the Shannon capacity. The main contribution of our work is two-fold:

- We characterize sufficient network conditions under which *at* most one relay is required to transmit at any given time to achieve the approximate capacity. Under such conditions, we provide closed-form expressions for the approximate capacity and for the network schedule, that is the fraction of time each relay should transmit. Such a schedule leads to a significant reduction in the average power consumption at the relays, compared to a random network with identical n (where potentially at each point in time more than one relay is transmitting) and hence, the proposed scheduling is *energy-efficient*. Another advantage of a schedule with at most one relay in transmit mode at each point in time is that it simplifies the synchronization problem at the destination. Moreover, the scheduling and the approximate capacity under these conditions can be computed by solving a system of $O(n)$ equations in $O(n)$ variables, rather than solving a linear program with $O(2^n)$ variables and constraints. This can be advantageous even in scenarios with a relatively small number of relays n, e.g., in the case of fast fading channels, where the scheduling and the approximate capacity have to be updated in short intervals of time as the channel gains vary rapidly. However, such a complexity reduction becomes particularly appealing when n is relatively large. This is expected to be the case in a number of next generation networks that rely on using high-frequency bands. In such networks, in fact, multi-hop relaying stands as a viable solution to counter the hostile propagation quality of the high-frequency spectrum, such as high path loss, atmospheric and rain absorption, and low penetration through objects. Examples of such networks include satellite communication and mega-constellations [30] [31], and device-to-device millimeter wave infrastructure networks, such as Terragraph [32]. It is worth noting that our result can also be readily translated to obtain sufficient conditions under which operating the network only in states with at most one relay in receive mode is sufficient to achieve the approximate capacity.
- We design a time-block relaying scheme, where in each block the $n+1$ energy-efficient states are operated in a fixed order. Our designed scheme has the following appealing features: (i) each relay simply stores and forwards some selected information bits that it receives, without any further processing or coding; (ii) it leverages *information flow preservation* at the relays to explicitly provide the information that each relay is exclusively responsible to store and forward to the destination; (iii) it guarantees that the destination is capable of successfully decoding all the information bits sent by the source in block B by the end of block $B+1$; and (iv) it achieves a rate that is equal to the approximate capacity when the derived sufficient network conditions are satisfied.

Finally, we highlight that, although the derived results are for the linear deterministic approximation of the Gaussian noise channel, they represent a first fundamental step towards identifying a set of provable optimal set of $n+1$ states and corresponding relaying scheme for half-duplex relay networks with arbitrary topology. In particular, the derived sufficient conditions for the optimality of an energy-efficient scheduling and the designed relaying scheme can be readily translated to obtain similar results for the original noisy Gaussian network for relevant topologies, such as line and diamond networks, and in general they can be leveraged as guidelines when the Gaussian network has an arbitrary topology.

C. Paper Organization

Section II introduces the notation, it describes the Gaussian and the linear deterministic half-duplex networks with n relays and summarizes known capacity results. Section III presents the first main result of the paper, the proof of which is in Section IV. Specifically, Section III characterizes sufficient conditions under which the set of (at most) $n+1$ network states in which at most one relay is transmitting (and the set with at most $n+1$ states with at most one relay receiving) suffices to characterize the approximate capacity of the binary-valued linear deterministic approximation of the Gaussian noise channel. Section V presents the second main result of the paper, namely

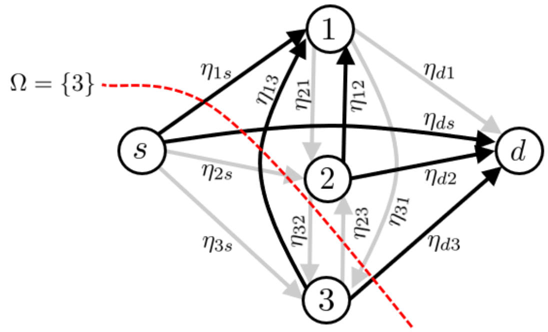

Fig. 1: $n{=}3$ relay network with $\Omega{=}\{3\}$ and $\mathcal{S}{=}\{2,3\}$.

a time-block relaying scheme which achieves a rate equal to the network approximate capacity whenever the derived sufficient conditions are satisfied. Section VI concludes the paper. Some proofs can be found in the appendix.

II. Notation and System Model

Notation: We denote the set of integers $\{i,\dots,m\}$ by $[i:m]$, and $\{1,\dots,m\}$ by $[m]$; note that $[i:m]=\varnothing$ if $i>m$. For a variable θ and a set $\mathcal{X}$, $\theta_{\mathcal{X}}=\{\theta_x : x\in\mathcal{X}\}$. We use boldface letters to refer to matrices. For a matrix $\mathbf{M}$, $\det(\mathbf{M})$ is the determinant of $\mathbf{M}$, $\mathbf{M}^T$ is the matrix transpose of $\mathbf{M}$ and $\mathbf{M}[\mathcal{A},\mathcal{B}]$ is the submatrix of $\mathbf{M}$ obtained by retaining all the rows indexed by the set $\mathcal{A}$ and all the columns indexed by the set $\mathcal{B}$. $\lfloor\cdot\rfloor$ and $\lceil\cdot\rceil$ are the floor and ceiling operations, respectively, and $[a]^+=\max\{a,0\}$. $\mathbf{0}_{p\times q}$ is the zero matrix of dimension $p\times q$; $\mathbf{I}_p$ is the $p\times p$ identity matrix.

The Gaussian half-duplex network with n relays consists of a source (node s) that communicates with a destination (node d) through a direct link (from s to d) and through an arbitrary network of n relays that operate in half-duplex. At each time t, the input/output relationship of this network is described as

$$\begin{aligned} Y_d^t &= \sum_{i=1}^{n} S_i^t h_{di} X_i^t + h_{ds} X_s^t + Z_d^t, \\ Y_i^t &= \left(1-S_i^t\right)\left(h_{is}X_s^t + \sum_{j\in[n]} S_j^t h_{ij} X_j^t + Z_i^t\right), \end{aligned} \tag{1}$$

for $i\in[n]$. Note that, at each time instant t: (i) S_i^t is a binary random variable that indicates the state of relay $i\in[n]$, with $S_i^t=0$ (respectively, $S_i^t=1$) indicating that relay i is receiving (respectively, transmitting); (ii) the source and the destination are always transmitting and receiving, respectively; (iii) X_i^t is the channel input at node i that satisfies the unit average power constraint $\mathbb{E}[|X_i^t|^2]\le 1$ for $i\in\{s\}\cup[n]$; (iv) h_{ij} with $i\in[n]\cup\{d\}$ and $j\in\{s\}\cup[n]$ is the *time-invariant* complex channel gain from node j to node i; note that, without loss of generality, we let $h_{ii}=0$; (v) $Z_i^t\sim\mathcal{CN}(0,1)$ is the complex additive white Gaussian noise at node $i\in\{d\}\cup[n]$; and finally, (vi) Y_i^t is the received signal at node $i\in\{d\}\cup[n]$.

The Shannon capacity $\mathsf{C}^{\rm G}$ of the network in (1) is not known in general. However, $\mathsf{C}^{\rm G}$ can be approximated within an additive gap by using the cut-set upper bound together with relaying schemes such as quantize-map-and-forward [17] and noisy network coding [19]. Here the gap only depends on the number of nodes in the network, but it is independent from the channel gains and operating SNR. In particular, for some relevant topologies, such as line and diamond networks, we can focus on the binary *linear deterministic* approximation of the Gaussian noise network model [17], for which the approximate capacity is known and provides an approximation for $\mathsf{C}^{\rm G}$. The linear deterministic model (a.k.a. ADT model [17] based on the first letter of the names of the authors Avestimehr, Diggavi and Tse) corresponding to the Gaussian noise network in (1) has an input-output relationship given by

$$\begin{aligned} Y_d^t &= \sum_{i=1}^{n} S_i^t \mathbf{D}^{\eta-\eta_{d,i}} X_i^t + \mathbf{D}^{\eta-\eta_{d,s}} X_s^t, \\ Y_i^t &= \left(1-S_i^t\right)\left(\mathbf{D}^{\eta-\eta_{i,s}} X_s^t + \sum_{j\in[n]} S_j^t \mathbf{D}^{\eta-\eta_{i,j}} X_j^t\right), \end{aligned} \tag{2}$$

for $i\in[n]$, where

$$\mathbf{D}^{\eta-m} = \left[\begin{array}{c|c} \mathbf{0}_{(\eta-m)\times m} & \mathbf{0}_{(\eta-m)\times(\eta-m)} \\ \hline \mathbf{I}_m & \mathbf{0}_{m\times(\eta-m)} \end{array}\right],$$

and $\eta_{i,j}=\left\lceil \log|h_{ij}|^2\right\rceil^+,\ i\in[n]\cup\{d\}, j\in\{s\}\cup[n], i\neq j$. Here, the vectors X_s^t, X_i^t, Y_d^t, and Y_i^t with $i\in[n]$ are binary of length $\eta=\max\eta_{i,j}$, where the maximization is taken over all channels $\eta_{i,j}$'s in the network; $\mathbf{D}$ is the so-called $\eta\times\eta$ shift matrix, and $S_i^t, i\in[n]$ is the ith relay binary-valued state random variable.

Example 1. Consider the $n=3$ relay half-duplex network in Fig. 1. For the cut $\Omega=\{3\}$ and state $\mathcal{S}=\{2,3\}$, we have $\{s\}\cup(\Omega\cap\mathcal{S})=\{s,3\}$ and $\{d\}\cup(\Omega^c\cap\mathcal{S}^c)=\{d,1\}$. The input-output relationship for this cut and state is given by

$$\begin{bmatrix} Y_d \\ Y_1 \end{bmatrix} = \begin{bmatrix} \mathbf{D}^{\eta-\eta_{d,s}} & \mathbf{D}^{\eta-\eta_{d,3}} \\ \mathbf{D}^{\eta-\eta_{1,s}} & \mathbf{D}^{\eta-\eta_{1,3}} \end{bmatrix} \begin{bmatrix} X_s \\ X_3 \end{bmatrix}.$$

The approximate capacity of the linear deterministic model in (2) is given by the solution of [17]

$$\begin{aligned} \mathsf{C}^{\rm LD} &= \max_{\boldsymbol{\lambda}} t \\ &\quad \text{s.t. } t \le g_\Omega \triangleq \sum_{\mathcal{S}\subseteq[n]} \lambda_{\mathcal{S}} f_{\mathcal{S}}^{\Omega}, \qquad \forall\Omega\subseteq[n], \\ &\qquad g_p \triangleq \sum_{\mathcal{S}\subseteq[n]} \lambda_{\mathcal{S}} \le 1, \quad \lambda_{\mathcal{S}}\ge 0, \quad \forall\mathcal{S}\subseteq[n], \end{aligned} \tag{3}$$

where: (i) $\mathcal{S}=\{i\in[n]:S_i=1\}$ is the set of relay nodes in transmit mode; (ii) $\lambda_{\mathcal{S}}\ge 0$ is the fraction of time that the network operates in state $\mathcal{S}$ and hence, $\sum_{\mathcal{S}\subseteq[n]}\lambda_{\mathcal{S}}{\le}1$; (iii) $\boldsymbol{\lambda}$ is referred to as a network *schedule* and is a vector obtained by stacking together $\lambda_{\mathcal{S}}$ for all $\mathcal{S}\subseteq[n]$; (iv) $\Omega\subseteq[n]$ denotes a partition of the relays in the 'side of s', i.e., $\{s\}\cup\Omega$ is a network cut; similarly, $\Omega^c=[n]\setminus\Omega$ is a partition of the relays in the 'side of d'. Moreover, we define

$$f_{\mathcal{S}}^{\Omega} \triangleq I\left(X_s, X_{\Omega\cap\mathcal{S}}; Y_d, Y_{\Omega^c\cap\mathcal{S}^c} | X_{\Omega^c\cap\mathcal{S}}, \mathcal{S}\right) = \text{rank}\left(\mathbf{F}_{\mathcal{S}}^{\Omega}\right), \tag{4}$$

where $\mathbf{F}_{\mathcal{S}}^{\Omega}$ is the transfer matrix from $X_{\{s\}\cup(\Omega\cap\mathcal{S})}$ to $Y_{\{d\}\cup(\Omega^c\cap\mathcal{S}^c)}$, corresponding to the ADT model [17].

Therefore, we have

$$f_{\{2,3\}}^{\{3\}} = \text{rank}\left(\mathbf{F}_{\{2,3\}}^{\{3\}}\right) = \text{rank}\begin{bmatrix} \mathbf{D}^{\eta-\eta_{d,s}} & \mathbf{D}^{\eta-\eta_{d,3}} \\ \mathbf{D}^{\eta-\eta_{1,s}} & \mathbf{D}^{\eta-\eta_{1,3}} \end{bmatrix}. \diamond$$

$$\max_{\boldsymbol{\lambda}} t$$

$$\begin{aligned}
&\text{For } \Omega = \varnothing: && t \leq \sum_{\mathcal{S}\subseteq[2]} f_{\mathcal{S}}^{\varnothing}\lambda_{\mathcal{S}} = \max\{\eta_{1,s},\eta_{2,s},\eta_{d,s}\}\lambda_{\varnothing} + \max\{\eta_{2,s},\eta_{d,s}\}\lambda_{\{1\}} + \max\{\eta_{1,s},\eta_{d,s}\}\lambda_{\{2\}} + \eta_{d,s}\lambda_{\{1,2\}},\\
&\text{For } \Omega = \{1\}: && t \leq \sum_{\mathcal{S}\subseteq[2]} f_{\mathcal{S}}^{\{1\}}\lambda_{\mathcal{S}} = \max\{\eta_{2,s},\eta_{d,s}\}\lambda_{\varnothing} + f_{\{1\}}^{\{1\}}\lambda_{\{1\}} + \eta_{d,s}\lambda_{\{2\}} + \max\{\eta_{d,1},\eta_{d,s}\}\lambda_{\{1,2\}},\\
&\text{For } \Omega = \{2\}: && t \leq \sum_{\mathcal{S}\subseteq[2]} f_{\mathcal{S}}^{\{2\}}\lambda_{\mathcal{S}} = \max\{\eta_{1,s},\eta_{d,s}\}\lambda_{\varnothing} + \eta_{d,s}\lambda_{\{1\}} + f_{\{2\}}^{\{2\}}\lambda_{\{2\}} + \max\{\eta_{d,2},\eta_{d,s}\}\lambda_{\{1,2\}},\\
&\text{For } \Omega = \{1,2\}: && t \leq \sum_{\mathcal{S}\subseteq[2]} f_{\mathcal{S}}^{\{1,2\}}\lambda_{\mathcal{S}} = \eta_{d,s}\lambda_{\varnothing} + \max\{\eta_{d,1},\eta_{d,s}\}\lambda_{\{1\}} + \max\{\eta_{d,2},\eta_{d,s}\}\lambda_{\{2\}} + \max\{\eta_{d,1},\eta_{d,2},\eta_{d,s}\}\lambda_{\{1,2\}},\\
&\text{Sum of } \boldsymbol{\lambda}: && 1 = \lambda_{\varnothing} + \lambda_{\{1\}} + \lambda_{\{2\}} + \lambda_{\{1,2\}},\\
&\text{Feasibility}: && \boldsymbol{\lambda} \geq 0,
\end{aligned}$$

where $\boldsymbol{\lambda} = \{\lambda_{\mathcal{S}} : \mathcal{S} \subseteq [2]\}$, and $f_{\{1\}}^{\{1\}} = \text{rank}\begin{bmatrix}\mathbf{D}^{\eta-\eta_{d,s}} & \mathbf{D}^{\eta-\eta_{d,1}}\\ \mathbf{D}^{\eta-\eta_{2,s}} & \mathbf{D}^{\eta-\eta_{2,1}}\end{bmatrix}$, $f_{\{2\}}^{\{2\}} = \text{rank}\begin{bmatrix}\mathbf{D}^{\eta-\eta_{d,s}} & \mathbf{D}^{\eta-\eta_{d,2}}\\ \mathbf{D}^{\eta-\eta_{1,s}} & \mathbf{D}^{\eta-\eta_{1,2}}\end{bmatrix}$. ◇

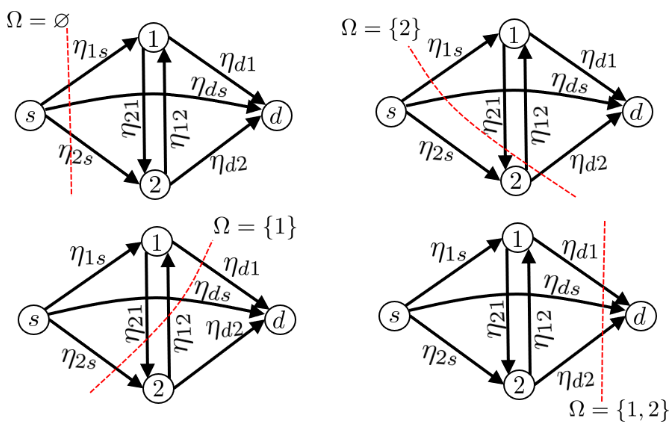

Fig. 2: The 2^n cuts of an $n = 2$ relay network.

Example 2. We use the optimization problem in (3) to illustrate how to compute the approximate capacity of an $n = 2$ relay half-duplex network. For the $n = 2$ relay network, we have a cut constraint corresponding to each of the $2^2 = 4$ possible cuts shown in Fig. 2. The information flow through each cut is the sum of information flows through that cut in all the 2^2 states. Thus, for this $2-$relay network, the optimization problem at the top of this page gives the approximate capacity.

In this work, we seek to identify *sufficient* network conditions which allow to determine a set of $n+1$ states (out of the 2^n possible ones) that suffice to achieve the approximate capacity in (3) and can be readily translated into a similar result for the original noisy Gaussian channel model in (1).

III. Conditions for Optimality of States with at Most One Relay Transmitting

Without loss of generality, we assume that the relay nodes are arranged in increasing order of their left link capacities, that is, $\eta_{1,s} \leq \eta_{2,s} \leq \cdots \leq \eta_{n,s}$. We define $\mathbf{P}$ to be an $(n+2)\times(n+2)$ matrix, the rows and columns of which are indexed by $[0:n+1]$, and

$$\mathbf{P}_{i,j} = \begin{cases} -f_{\{j\}}^{[i:n]}, & (i,j) \in [n+1]^2,\\ 0, & (i,j) = (0,0),\\ 1, & \text{otherwise,}\end{cases} \tag{5}$$

where we define $f_{\{n+1\}}^{\Omega} = f_{\varnothing}^{\Omega}$, for consistency. Moreover, for $i \in [0:n+1]$ we use $\mathsf{P}_{(i)}$ to denote the *minor* of $\mathbf{P}$ associated with the row 0 and column i of the matrix $\mathbf{P}$, that is

$$\mathsf{P}_{(i)} \triangleq \det\left(\mathbf{P}[[n+1],[0:n+1]\setminus\{i\}]\right).$$

Finally, we define $\mathbb{S} \triangleq \{\{1\},\{2\},\ldots,\{n\},\varnothing\}$ to be the set of the $n+1$ states, where at most one relay is transmitting in each state.

The main result of this paper is presented in Theorem 1, which characterizes sufficient network conditions for the optimality of operating the network only in states $\mathcal{S} \in \mathbb{S}$.

Theorem 1. *Consider a half-duplex relay network with $\eta_{1,s} \leq \eta_{2,s} \leq \cdots \leq \eta_{n,s}$. Whenever $\det(\mathbf{P}) \neq 0$ and $\frac{(-1)^{n+1}\mathsf{P}_{(n+1)}}{\det(\mathbf{P})} \geq 0$, then it is optimal to operate the network in states in $\mathbb{S}$ to achieve C^{LD} in (3).*

Example 3. Consider the relay network in Fig. 1 with link capacities given by $\eta_{d,s} = 1, \eta_{1,s} = 2,\ \eta_{2,s} = 3,\ \eta_{3,s} = 5,\ \eta_{d,1} = 6,\ \eta_{d,2} = 5,\ \eta_{d,3} = 3,\ \eta_{1,2} = 3,\ \eta_{2,1} = 4,\ \eta_{3,2} = 5,\ \eta_{2,3} = 3,\ \eta_{3,1} = 2$ and $\eta_{1,3} = 4$. For this network, the matrix $\mathbf{P}$ is given by

$$\mathbf{P} = \begin{bmatrix} 0 & 1 & 1 & 1 & 1\\ 1 & -6 & -5 & -3 & -1\\ 1 & -1 & -7 & -4 & -2\\ 1 & -3 & -2 & -7 & -3\\ 1 & -5 & -5 & -3 & -5\end{bmatrix}. \tag{6}$$

For matrix $\mathbf{P}$, $\det(\mathbf{P}) = 198 \neq 0$ and $\frac{(-1)^4\mathsf{P}_{(4)}}{\det(\mathbf{P})} = \frac{11}{198} \geq 0$, i.e., the conditions in Theorem 1 are satisfied for this $n = 3$ relay network. Thus, operating this network in states $\mathbb{S} = \{\{1\},\{2\},\{3\},\varnothing\}$ achieves C^{LD} in (3). ◇

Remark 1. *Note that $\mathbb{S}$ consists of all the states where at most one relay transmits, while the rest of the relays receive. A similar condition can be obtained for the optimality of the states*

$$\mathbb{S}' = \{[n],[n]\setminus\{1\},[n]\setminus\{2\},\ldots,[n]\setminus\{n\}\},$$

where at most one relay is in receive mode.

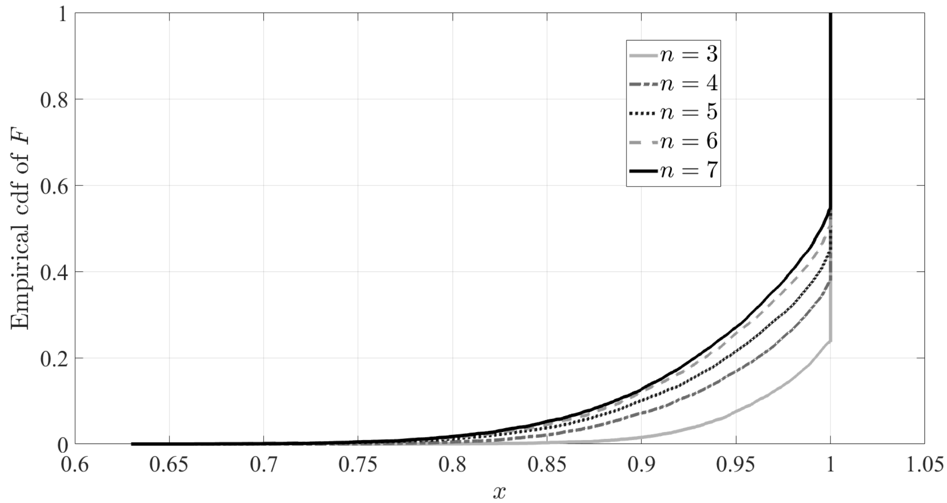

Fig. 3: Empirical cdf of $F = \mathsf{R}^{\mathbb{S}}/\mathsf{C}^{\mathrm{LD}}$ for $n = [3:7]$ relay networks.

Remark 2. *The conditions in Theorem 1 are a consequence of the relaying scheme proposed in Section V used to operate the network in states $\mathbb{S}$. This scheme is based on information flow preservation at each relay, i.e., the amount of unique linearly independent bits of information that each relay is responsible to receive is equal to the amount of unique linearly independent bits of information that the relay transmits. The conditions in Theorem 1 ensure the feasibility and optimality of this scheme as we show in Section V.*

Remark 3. *An important motivation behind the study of the energy-efficient states $\mathbb{S}$ is that even when the conditions in Theorem 1 are not satisfied, the network can still achieve a significant fraction of the approximate capacity by operating in only these $n+1$ states. This is illustrated via numerical simulations in Fig. 3. For these simulations, we generated random $n = [3 : 7]$ relay fully-connected networks, where the capacity $\eta_{i,j}$ for $i \in [1:n] \cup \{d\}, j \in [1:n] \cup \{s\}, i \neq j$ of each network link is a random integer in $[1 : 10]$. For each randomly generated network, we computed $F = \mathsf{R}^{\mathbb{S}}/\mathsf{C}^{LD}$, where $\mathsf{R}^{\mathbb{S}}$ is the rate achieved by operating the network in the energy-efficient states $\mathbb{S}$, and C^{LD} is the approximate capacity computed by solving the linear program in (3). For each of the five values of n, we generated 10^4 random networks. In Fig. 3, we plot the empirical cumulative distribution function (cdf) of F for $n = [3 : 7]$. These numerical results show that, even when it is not optimal to operate the network in the energy-efficient states, we can still achieve a significant fraction (e.g., 80%, 85%, 90% and 95% achieved by 98%, 94%, 87%, 72% of the generated networks, respectively) of the approximate capacity by restricting the network to operate in these states. This makes the study of these energy-efficient states important.*

In the remainder of this section, we analyze the variables $f_{\mathcal{S}}^{\Omega}$ and present some of their properties, which play an important role in the proof of Theorem 1, presented in Section IV.

A. Properties of $f_{\mathcal{S}}^{\Omega}$

We present two properties of $f_{\mathcal{S}}^{\Omega} = \mathrm{rank}(\mathbf{F}_{\mathcal{S}}^{\Omega})$ that we will use in the proof of Theorem 1.

Lemma 1. *For all $\Omega \subseteq [n]$ and $\mathcal{S} \subseteq [n]$, we have that*

$$f_{\mathcal{S}}^{\Omega} \geq \max\left\{ \max_{i \in \Omega^c \cap \mathcal{S}^c} \eta_{i,s}, \max_{j \in \Omega \cap \mathcal{S}} \eta_{d,j}, \eta_{d,s} \right\}, \tag{7}$$

with equality if $\Omega^c \cap \mathcal{S}^c = \varnothing$ or $\Omega \cap \mathcal{S} = \varnothing$.

Proof. Let

$$i^\star = \arg \max_{i \in \Omega^c \cap \mathcal{S}^c} \eta_{i,s}, \quad \text{and} \quad j^\star = \arg \max_{j \in \Omega \cap \mathcal{S}} \eta_{d,j}.$$

Then, the submatrix of $\mathbf{F}_{\mathcal{S}}^{\Omega}$ induced by row blocks $\{d, i^\star\}$ and column blocks $\{s, j^\star\}$ is

$$\left[\begin{array}{c|c} \mathbf{D}^{\eta-\eta_{d,s}} & \mathbf{D}^{\eta-\eta_{d,j^\star}} \\ \hline \mathbf{D}^{\eta-\eta_{i^\star,s}} & \mathbf{D}^{\eta-\eta_{i^\star,j^\star}} \end{array}\right],$$

the rank of which is at least $\max\{\max\{\eta_{i^\star,s}, \eta_{d,s}\}, \max\{\eta_{d,j^\star}, \eta_{d,s}\}\}$ (from the first column and first row of the above matrix). This provides a lower bound on the rank of $\mathbf{F}_{\mathcal{S}}^{\Omega}$. Moreover, if $\Omega^c \cap \mathcal{S}^c = \varnothing$, then $\mathbf{F}_{\mathcal{S}}^{\Omega}$ only consists of one row induced by $\{d\}$ and columns induced by $\{\eta_{d,j} : j \in (\Omega \cap \mathcal{S}) \cup \{s\}\}$ and hence, $\mathrm{rank}(\mathbf{F}_{\mathcal{S}}^{\Omega}) = \max\{\eta_{d,s}, \max_{j \in \Omega \cap \mathcal{S}} \eta_{d,j}\}$. Similarly, if $\Omega \cap \mathcal{S} = \varnothing$, then $\mathbf{F}_{\mathcal{S}}^{\Omega}$ has only one column and hence, $\mathrm{rank}(\mathbf{F}_{\mathcal{S}}^{\Omega}) = \max\{\eta_{d,s}, \max_{i \in \Omega^c \cap \mathcal{S}^c} \eta_{i,s}\}$. If both $\Omega \cap \mathcal{S} = \varnothing$ and $\Omega^c \cap \mathcal{S}^c = \varnothing$, then $\mathbf{F}_{\mathcal{S}}^{\Omega} = \mathbf{D}^{\eta-\eta_{d,s}}$, the rank of which is $\eta_{d,s}$. This concludes the proof of Lemma 1. □

Lemma 2. *For a given state $\mathcal{S} \subseteq [n]$, $f_{\mathcal{S}}^{\Omega}$ is submodular in Ω, that is,*

$$f_{\mathcal{S}}^{\Omega_1} + f_{\mathcal{S}}^{\Omega_2} \geq f_{\mathcal{S}}^{\Omega_1 \cap \Omega_2} + f_{\mathcal{S}}^{\Omega_1 \cup \Omega_2},$$

for any subsets $\Omega_1, \Omega_2 \subseteq [n]$. Similarly, for a given cut $\Omega \subseteq [n]$, $f_{\mathcal{S}}^{\Omega}$ is submodular in $\mathcal{S}$, that is,

$$f_{\mathcal{S}_1}^{\Omega} + f_{\mathcal{S}_2}^{\Omega} \geq f_{\mathcal{S}_1 \cup \mathcal{S}_2}^{\Omega} + f_{\mathcal{S}_1 \cap \mathcal{S}_2}^{\Omega},$$

for any subsets $\mathcal{S}_1, \mathcal{S}_2 \subseteq [n]$.

Proof. The proof is given in [25, Appendix A]. □

IV. Proof of Theorem 1

In this section, we present the proof of Theorem 1, which consists of three main steps. We first introduce an auxiliary optimization problem in (8) in Section IV-A, which is a relaxed version of the problem in (3) and hence, its solution $\mathsf{C}^{\rm U}$ provides an upper bound on the optimum value of (3), i.e., $\mathsf{C}^{\rm U} \geq \mathsf{C}^{\rm LD}$. In Section IV-B, we propose a solution $(\boldsymbol{\lambda}^\star, t^\star)$ for the (relaxed) optimization problem in (8), where $\lambda^\star_{\mathcal{S}} = 0$ for all $\mathcal{S} \notin \mathbb{S}$. We show that, under the conditions of Theorem 1, $(\boldsymbol{\lambda}^\star, t^\star)$ is feasible and optimal, which leads to $\mathsf{C}^{\rm U} = t^\star$. Finally, in Section IV-C we show that the proposed solution is feasible for the original problem in (3), implying that $\mathsf{C}^{\rm LD} \geq t^\star$. Therefore, putting the three results together, we get $t^\star = \mathsf{C}^{\rm U} \geq \mathsf{C}^{\rm LD} \geq t^\star$. This shows that $\mathsf{C}^{\rm LD} = t^\star$ can be attained by $\boldsymbol{\lambda}^\star$ that satisfies the claim of Theorem 1. This concludes the proof of the theorem.

A. An Upper Bound for the Approximate Capacity

The optimization problem in (3) consists of 2^n cut constraints. We can relax these constraints except for $n+1$ of them. More formally, we define

$$
\begin{aligned}
\mathsf{C}^{\rm U} = \max_{\boldsymbol{\lambda}} \ & t \\
\text{s.t. } & t \leq g_{[i:n]} = \sum_{\mathcal{S}\subseteq[n]} \lambda_{\mathcal{S}} f_{\mathcal{S}}^{[i:n]}, \quad i \in [n+1], \\
& g_p \triangleq \sum_{\mathcal{S}\subseteq[n]} \lambda_{\mathcal{S}} \leq 1, \quad \lambda_{\mathcal{S}} \geq 0, \quad \mathcal{S} \subseteq [n],
\end{aligned} \tag{8}
$$

where $[n+1:n] = \varnothing$. Note that the optimization problem in (8) is less constrained compared to (3). Hence, the problem in (8) has a wider feasible set than (3), and its maximum objective function cannot be smaller than that of the problem in (3), i.e., $\mathsf{C}^{\rm U} \geq \mathsf{C}^{\rm LD}$.

B. An Optimal Solution for the Relaxed Optimization Problem

We show that under the conditions of Theorem 1, the states in $\mathbb{S}$ are sufficient to achieve $\mathsf{C}^{\rm U}$ in (8), that is, there exists an optimal solution with $(t^\star, \lambda^\star_{\mathbb{S}}) \geq 0$ and $\lambda^\star_{\mathcal{S}} = 0$, for all $\mathcal{S} \notin \mathbb{S}$. In particular, we start by the following proposition, which is proved in Appendix A.

Proposition 1. *Assume* $\det(\mathbf{P}) \neq 0$ *and* $\frac{(-1)^{n+1}\mathsf{P}_{(n+1)}}{\det(\mathbf{P})} \geq 0$. *Then, the variables*

$$
\begin{aligned}
\lambda^\star_{\varnothing} &= \lambda^\star_{\{n+1\}} := (-1)^{n+1}\frac{\mathsf{P}_{(n+1)}}{\det(\mathbf{P})}, \\
\lambda^\star_{\{i\}} &:= (-1)^{i}\frac{\mathsf{P}_{(i)}}{\det(\mathbf{P})}, \qquad i \in [n], \\
\lambda^\star_{\mathcal{S}} &:= 0, \qquad \mathcal{S} \notin \mathbb{S},
\end{aligned} \tag{9}
$$

are non-negative.

We now leverage the Karush–Kuhn–Tucker (KKT) conditions to prove the proposition below.

Proposition 2. *Assume* $\det(\mathbf{P}) \neq 0$ *and* $\frac{(-1)^{n+1}\mathsf{P}_{(n+1)}}{\det(\mathbf{P})} \geq 0$. *Then,* $\boldsymbol{\lambda}^\star = \{\lambda_{\mathcal{S}} : \mathcal{S} \subseteq [n]\}$ *defined in (9) is an optimal solution for the problem in (8). Consequently, we have*

$$
\mathsf{C}^{\rm U} = t^\star = g^\star_{[i:n]} = \sum_{\mathcal{S}\in\mathbb{S}} \lambda^\star_{\mathcal{S}} f_{\mathcal{S}}^{[i:n]} = \frac{\mathsf{P}_{(0)}}{\det(\mathbf{P})}, \tag{10}
$$

for every $i \in [n+1]$.

Proof of Proposition 2. The proof leverages the KKT conditions for the optimality of the proposed solution. For the KKT multipliers $\boldsymbol{\mu} = (\mu_p, \mu_1, \ldots, \mu_{n+1})$ and $\boldsymbol{\sigma} = (\sigma_{\mathcal{S}} : \mathcal{S} \subseteq [n])$, the Lagrangian $\mathcal{L}(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\lambda}, t)$ for the optimization problem in (8) is given by

$$
-t + \sum_{i\in[n+1]} \mu_i (t - g_{[i:n]}) + \mu_p (g_p - 1) - \sum_{\mathcal{S}\subseteq[n]} \sigma_{\mathcal{S}} \lambda_{\mathcal{S}}. \tag{11}
$$

In the following, we proceed with a choice of $(\boldsymbol{\mu}, \boldsymbol{\sigma})$ where $\boldsymbol{\mu}$ is the solution of

$$
\begin{bmatrix} \mu_p & \mu_1 & \ldots & \mu_n & \mu_{n+1} \end{bmatrix} \mathbf{P} = \begin{bmatrix} 1 & 0 & \ldots & 0 & 0 \end{bmatrix}, \tag{12}
$$

and

$$
\sigma_{\mathcal{S}} = \mu_p - \sum_{i=1}^{n+1} \mu_i f_{\mathcal{S}}^{[i:n]}, \tag{13}
$$

for every $\mathcal{S} \subseteq [n]$. We next prove the optimality of $(\boldsymbol{\lambda}^\star, t^\star)$ by showing that the set of KKT multipliers $(\boldsymbol{\mu}, \boldsymbol{\sigma})$ defined in (12) and (13) together with $(\boldsymbol{\lambda}^\star, t^\star)$ defined in (9) and (10) satisfy the following four groups of conditions.

• *Primal Feasibility.* First, note that Proposition 1 guarantees that the constraint $\lambda^\star_{\mathcal{S}} \geq 0$ is satisfied for every $\mathcal{S} \subseteq [n]$. In order to show the feasibility of the solution, it remains to show that $t^\star \leq g_{[i:n]}$ for every $i \in [n+1]$ and $\sum_{\mathcal{S}\subseteq[n]} \lambda^\star_{\mathcal{S}} \leq 1$. Note that by forcing these inequalities to hold with equality, and setting $\lambda_{\mathcal{S}} = 0$, for all $\mathcal{S} \subseteq [n]$ with $\mathcal{S} \notin \mathbb{S}$ in (8), we obtain a system of $(n+2)$ linear equations in $(n+2)$ variables (namely t and $\lambda_{\mathcal{S}}$ for $\mathcal{S} \in \mathbb{S}$), given by

$$
\mathbf{P} \begin{bmatrix} t & \lambda_{\{1\}} & \ldots & \lambda_{\{n\}} & \lambda_{\varnothing} \end{bmatrix}^T = \begin{bmatrix} 1 & 0 & \ldots & 0 & 0 \end{bmatrix}^T, \tag{14}
$$

where $\mathbf{P}$ is the matrix defined in (5) – see also Appendix A. The solution of (14) is indeed given in (9) and (10). Therefore, the solution $(\boldsymbol{\lambda}^\star, t^\star)$ is feasible for the optimization problem in (8). In particular, it is worth noting that $t^\star = \sum_{\mathcal{S}\in\mathbb{S}} \lambda^\star_{\mathcal{S}} f_{\mathcal{S}}^{[i:n]}$ is guaranteed by the ith row of the matrix identity above, and hence (10) holds for all values of $i \in [n+1]$. Furthermore, we can further simplify $\sum_{\mathcal{S}\in\mathbb{S}} \lambda^\star_{\mathcal{S}} f_{\mathcal{S}}^{[i:n]}$ as follows

$$
\begin{aligned}
\sum_{\mathcal{S}\in\mathbb{S}} \lambda^\star_{\mathcal{S}} f_{\mathcal{S}}^{[i:n]} &\overset{\rm (a)}{=} \sum_{j=1}^{n+1} (-1)^j \frac{\mathsf{P}_{(j)}}{\det(\mathbf{P})} f_{\{j\}}^{[i:n]} \\
&= \frac{1}{\det(\mathbf{P})} \sum_{j=1}^{n+1} (-1)^j \mathsf{P}_{(j)} f_{\{j\}}^{[i:n]} \\
&\overset{\rm (b)}{=} \frac{1}{\det(\mathbf{P})} \left(\mathsf{P}_{(0)} - \det(\mathbf{P}_{i\to 0})\right) \overset{\rm (c)}{=} \frac{\mathsf{P}_{(0)}}{\det(\mathbf{P})},
\end{aligned}
$$

where we used the values of $\lambda^\star$s given in (9) in (a), step (b) follows from the Laplace expansion $\det(\mathbf{P}_{i\to 0}) = \mathsf{P}_{(0)} + \sum_{j=1}^{n+1} (-1)^j \mathsf{P}_{(j)} \left(-f_{\{j\}}^{[i:n]}\right)$ with $\mathbf{P}_{i\to 0}$

being the matrix $\mathbf{P}$ with its row 0 replaced by row i, and (c) follows from $\det(\mathbf{P}_{i\to 0}) = 0$ since it has identical rows.

• *Complementary Slackness.* We need to show that $(\boldsymbol{\mu}, \boldsymbol{\sigma})$ and the solution $(\boldsymbol{\lambda}^\star, t^\star)$ given in (9) and (10) satisfy (i) $\mu_i\left(t^\star - g^\star_{[i:n]}\right) = 0$ for all $i \in [n+1]$, (ii) $\mu_p(g^\star_p - 1) = 0$, and (iii) $\sigma_\mathcal{S}\lambda^\star_\mathcal{S} = 0$ for every $\mathcal{S} \subseteq [n]$. Conditions (i) and (ii) are readily implied by (10) and (9), respectively, for any choice of $\boldsymbol{\mu}$. Moreover, (9) guarantees the third condition for $\mathcal{S} \notin \mathbb{S}$, with $\lambda^\star_\mathcal{S} = 0$ whenever $\mathcal{S} \notin \mathbb{S}$. Finally, consider some $\mathcal{S} \in \mathbb{S}$, say $\mathcal{S} = \{j\}$ where $j \in [n+1]$ (and $j = n+1$ if $\mathcal{S} = \varnothing$). Then, the definition of $\sigma_{\{j\}}$ in (13) and the jth column of the matrix identity in (12) imply that

$$\sigma_{\{j\}} = \mu_p - \sum_{i=1}^{n+1} \mu_i f^{[i:n]}_{\{j\}} = \boldsymbol{\mu} \cdot \mathbf{P}[[0:n+1], j] = 0.$$

This ensures that $\sigma_\mathcal{S}\lambda^\star_\mathcal{S} = 0$, for all $\mathcal{S} \in \mathbb{S}$.

• *Stationarity.* We aim to prove that, when evaluated in $\boldsymbol{\mu}$ as in (12) and $\sigma_\mathcal{S}$ in (13), the derivatives of the Lagrangian in (11) with respect to t and $\lambda_\mathcal{S}, \mathcal{S} \subseteq [n]$, are zero. By taking the derivative of $\mathcal{L}(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\lambda}^\star, t^\star)$ with respect to t we get

$$\begin{aligned}\frac{\partial}{\partial t}\mathcal{L}(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\lambda}^\star, t^\star) &= -1 + \sum_{i=1}^{n+1} \mu_i \\ &= -1 + \boldsymbol{\mu} \cdot \mathbf{P}[[0:n+1], 0] = -1 + 1 = 0.\end{aligned}$$

Similarly, by taking the derivative with respect to $\lambda_\mathcal{S}$ we get

$$\frac{\partial}{\partial \lambda_\mathcal{S}}\mathcal{L}(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\lambda}^\star, t^\star) = \mu_p - \sum_{i=1}^{n+1} \mu_i f^{[i:n]}_\mathcal{S} - \sigma_\mathcal{S} = \sigma_\mathcal{S} - \sigma_\mathcal{S} = 0,$$

in which we used the definition of $\sigma_\mathcal{S}$ in (13).

• *Dual Feasibility.* Lastly, we need to prove that the KKT multipliers in (12), (13) are non-negative. To this end, we present the following proposition, which is proved in Appendix B.

Proposition 3. *All the entries of the vector $\boldsymbol{\mu}$ obtained from (12) are non-negative.*

Next, we focus on the KKT multipliers $\sigma_\mathcal{S}, \mathcal{S} \subseteq [n]$ in (13). For any state $\mathcal{S} = \{a_1, a_2, \ldots, a_k\}$, we can write

$$\begin{aligned}\sum_{i=1}^{n+1} \mu_i f^{[i:n]}_\mathcal{S} &= \sum_{i=1}^{n+1} \mu_i f^{[i:n]}_{\{a_1,\ldots,a_k\}} \\ &= \sum_{i=1}^{n+1} \mu_i \left[\sum_{j=1}^{k-1}\left(f^{[i:n]}_{\{a_1,\ldots,a_{j+1}\}} - f^{[i:n]}_{\{a_1,\ldots,a_j\}}\right) + f^{[i:n]}_{\{a_1\}}\right] \\ &\le \sum_{i=1}^{n+1} \mu_i \left[\sum_{j=1}^{k-1}\left(f^{[i:n]}_{\{a_{j+1}\}} - f^{[i:n]}_{\varnothing}\right) + f^{[i:n]}_{\{a_1\}}\right] \\ &= \sum_{i=1}^{n+1} \mu_i \left[\left(\sum_{j=1}^{k} f^{[i:n]}_{\{a_j\}}\right) - (k-1) f^{[i:n]}_{\varnothing}\right] \\ &= \sum_{j=1}^{k}\sum_{i=1}^{n+1} \mu_i f^{[i:n]}_{\{a_j\}} - (k-1)\sum_{i=1}^{n+1} \mu_i f^{[i:n]}_{\varnothing} \\ &= \sum_{j=1}^{k} \mu_p - (k-1)\mu_p = \mu_p,\end{aligned}$$

where the inequality follows from Lemma 2, i.e., $f^\Omega_\mathcal{S}$ is submodular in $\mathcal{S}$. Therefore, we get $\sigma_\mathcal{S} = \mu_p - \sum_{i=1}^{n+1} \mu_i f^{[i:n]}_\mathcal{S} \ge 0$. This concludes the proof of Proposition 2. □

C. Feasibility of $(\lambda^\star, t^\star)$ for C^{LD}

In Section IV-B, we have shown that the solution $(\boldsymbol{\lambda}^\star, t^\star)$ given in (9) and (10) is optimal for the optimization problem in (8). This implies that $t^\star = \mathsf{C}^{\mathrm{U}} \ge \mathsf{C}^{\mathrm{LD}}$ where C^{LD} is the approximate capacity of the network, obtained by solving the problem in (3). In the following, we aim to prove that $(\boldsymbol{\lambda}^\star, t^\star)$ in (9) and (10) is a feasible solution for the optimization problem in (3), which in turn implies $\mathsf{C}^{\mathrm{LD}} \ge t^\star$, and concludes the proof of Theorem 1. Towards this end, it suffices to show the feasibility of such a solution for (3), as stated in the following proposition.

Proposition 4. *The solution $(\boldsymbol{\lambda}^\star, t^\star)$ given in (9) is feasible for (3) and thus, $\mathsf{C}^{\mathrm{LD}} \ge t^\star$.*

Proof. Note that the two optimization problems in (3) and (8) have identical objectives and similar constraints. More precisely, the constraints in (8) are a subset of those in (3) and hence, they are clearly satisfied for (3) also because $(\boldsymbol{\lambda}^\star, t^\star)$ is an optimum solution for (8). Thus, we only need to focus on the constraints of the form $t^\star \le g^\star_\Omega = \sum_{\mathcal{S}\subseteq[n]} \lambda^\star_\mathcal{S} f^\Omega_\mathcal{S}$, which exclusively appear in (3).

In order to prove Proposition 4, let us consider an arbitrary cut $\Omega = [a_1 : b_1] \cup [a_2 : b_2] \cup \ldots \cup [a_k : b_k] \subseteq [n]$, where $a_1 \le b_1 \le a_2 \le b_2 \le \ldots \le a_k \le b_k$. We also define $a_{k+1} = n+1$. Recall from Lemma 2 that, for a given state $\mathcal{S}$, the function $f^\Omega_\mathcal{S}$ is submodular in Ω. Then, for $\Omega_1 = \Omega \cup [a_{i+1} : n]$ and $\Omega_2 = [b_i + 1 : n]$ with $i \in [k]$, we have $\Omega_1 \cup \Omega_2 = \Omega \cup [a_i : n]$ and $\Omega_1 \cap \Omega_2 = [a_{i+1} : n]$. Thus,

$$f^{\Omega\cup[a_{i+1}:n]}_\mathcal{S} + f^{[b_i+1:n]}_\mathcal{S} \ge f^{\Omega\cup[a_i:n]}_\mathcal{S} + f^{[a_{i+1}:n]}_\mathcal{S}, \tag{15}$$

for all $i \in [k]$. Moreover, since $a_{k+1} = n+1$, we have $[a_{k+1} : n] = \varnothing$, and $\Omega \subseteq [a_1 : n]$. Thus

$$\begin{aligned}&\sum_{\mathcal{S}\subseteq[n]} \lambda^\star_\mathcal{S} f^\Omega_\mathcal{S} \\ &= \sum_{\mathcal{S}\subseteq[n]} \lambda^\star_\mathcal{S}\left[\sum_{i=1}^{k}\left(f^{\Omega\cup[a_{i+1}:n]}_\mathcal{S} - f^{\Omega\cup[a_i:n]}_\mathcal{S}\right) + f^{\Omega\cup[a_1:n]}_\mathcal{S}\right] \\ &\ge \sum_{\mathcal{S}\subseteq[n]} \lambda^\star_\mathcal{S}\left[\sum_{i=1}^{k}\left(f^{[a_{i+1}:n]}_\mathcal{S} - f^{[b_i+1:n]}_\mathcal{S}\right) + f^{[a_1:n]}_\mathcal{S}\right] \\ &= \sum_{i=1}^{k+1}\sum_{\mathcal{S}\subseteq[n]} \lambda^\star_\mathcal{S} f^{[a_i:n]}_\mathcal{S} - \sum_{i=1}^{k}\sum_{\mathcal{S}\subseteq[n]} \lambda^\star_\mathcal{S} f^{[b_i+1:n]}_\mathcal{S} \\ &= \sum_{i=1}^{k+1} g^\star_{[a_i:n]} - \sum_{i=1}^{k} g^\star_{[b_i+1:n]} = (k+1)t^\star - kt^\star = t^\star,\end{aligned}$$

where the inequality follows from (15). This implies that $t^\star \le \sum_{\mathcal{S}\subseteq[n]} \lambda^\star_\mathcal{S} f^\Omega_\mathcal{S}$ for any $\Omega \subseteq [n]$. This concludes the claim of Proposition 4 and completes the proof of Theorem 1. □

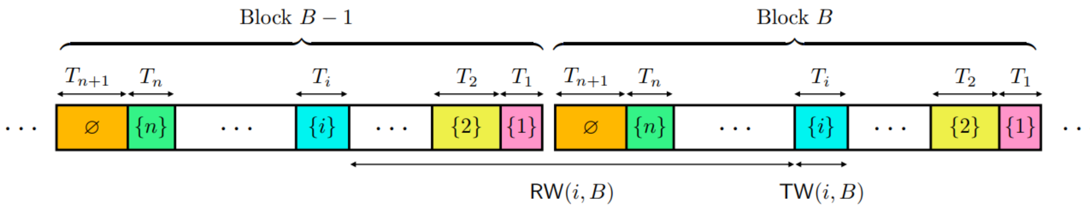

Fig. 4: The order of states in each block, and the transmit and reception windows.

V. Optimal Relaying Scheme with at Most One Relay Transmitting

In this section, we present a relaying scheme that only utilizes the states in $\mathbb{S}$. Before describing the scheme, we present some notation and preliminary facts in Section V-A and Section V-B, respectively. Next, we present the proposed relaying scheme in Section V-C, and show how the decoding is performed at the destination in Section V-D. Finally, in Section V-E we show that the proposed relaying scheme is optimal, i.e., it achieves the approximate capacity in (10).

A. Notation and Definitions

Without loss of generality, we may assume that $\{\lambda_{\mathcal{S}} : \mathcal{S} \in \mathbb{S}\}$ given by Proposition 1 are all rational numbers, otherwise we can approximate them with rational numbers with an arbitrarily small approximation error. With slight abuse of the notation, we may use state i to simply refer to state $\{i\}$. We also use $n+1$ to refer to state $\varnothing$. Let T be a sufficiently large integer such that $T_i := T\lambda_i$ is an integer for every $i \in [n+1]$. We consider blocks of communication of length $T = \sum_{i\in[n+1]} T_i$, which are divided into $n+1$ sub-blocks, each corresponding to one state in $\mathbb{S}$. The states in each block are operated in the following specific order (see also Fig. 4):

$$(\varnothing, \{n\}, \{n-1\}, \dots, \{1\}) = (n+1, n, \dots, 1)\,.$$

Thus, in block B relay i only transmits in time slot t, belonging to its *transmit window* $\mathsf{TW}(i,B)$, that is given by

$$\left[(B-1)T + \sum_{\mathcal{S}=i+1}^{n+1} T_{\mathcal{S}} + 1 : (B-1)T + \sum_{\mathcal{S}=i}^{n+1} T_{\mathcal{S}}\right]. \quad (16)$$

Similarly, we define $\mathsf{TW}(n+1,B) = [(B-1)T+1 : (B-1)T + T_{n+1}]$ to be the set of time slots in which the network operates in state $\varnothing$, i.e., only the source transmits.

For each relay $i \in [n]$, we also consider a *reception window* of length $T - T_i$, which includes the time slots between the end of the transmit window of relay i in block $B-1$ and the beginning of the transmit window of relay i in block B. More precisely, the reception window $\mathsf{RW}(i,B)$ of relay $i \in [n]$ in block B is defined as

$$\left[(B-2)T + \sum_{\mathcal{S}=i}^{n+1} T_{\mathcal{S}} + 1 : (B-1)T + \sum_{\mathcal{S}=i+1}^{n+1} T_{\mathcal{S}}\right]. \quad (17)$$

Fig. 4 shows $\mathsf{TW}(i,B)$ and $\mathsf{RW}(i,B)$ for relay i. Note that the reception window $\mathsf{RW}(i,B)$ of relay i in block B spans over blocks B and $B-1$. Moreover, $\mathsf{RW}(i,B)$ is further divided as

$$\mathsf{RW}(i,B) = \bigcup_{j\in[n+1]\setminus\{i\}} \mathsf{RW}(i,B,j),$$

where $\mathsf{RW}(i,B,j)$ is the set of time slots in the reception window $\mathsf{RW}(i,B)$ when the network operates in state $j \in [n+1]\setminus\{i\}$. It is easy to verify that (see Fig. 4)

$$\mathsf{RW}(i,B,j) = \begin{cases} \mathsf{TW}(j,B-1) & \text{if } j < i, \\ \mathsf{TW}(j,B) & \text{if } j > i. \end{cases} \quad (18)$$

B. Some Preliminary Results

The following proposition and its corollary play a central role in the proof of the feasibility of the proposed relaying scheme, and decodability of the information bits at the destination. We present the proofs of the proposition and corollary in Appendix C and Appendix D, respectively.

Proposition 5. *Consider the transfer matrix $\mathbf{F}_{\{i\}}^{[i:n]}$, and let c_j and r_j denote its jth column and jth row, respectively. Then, we have the following two properties:*
(i) If for some $j = k\eta + p$ (with some $k \in \{0,1\}$ and $0 < p < \eta$), column c_j is a linear combination of the columns in $\{c_\ell : \ell \in [j-1]\}$, then column c_{j+1} is also a linear combination of the columns in $\{c_\ell : \ell \in [j-1]\}$;
(ii) Similarly, if for some $j = k\eta + p$ (for some $k \in [0 : i-1]$ and $1 < p \le \eta$), row r_j is a linear combination of the rows in $\{r_\ell : \ell \in [j-1]\}$, then row r_{j-1} is also a linear combination of the rows in $\{r_\ell : \ell \in [j-2]\}$.

Corollary 1. *For the transfer matrix $\mathbf{F}_{\{i\}}^{[i:n]}$, consider the collection of columns*

$$\mathfrak{C}_i := \left[f_{\{i\}}^{[i+1:n]}\right] \cup \left[\eta + 1 : \eta + f_{\{i\}}^{[i:n]} - f_{\{i\}}^{[i+1:n]}\right], \quad (19)$$

and the collection of rows

$$\mathfrak{R}_i := \bigcup_{k=1}^{i} \left[k\eta - \left(f_{\{i\}}^{[k:n]} - f_{\{i\}}^{[k-1:n]}\right) + 1 : k\eta\right], \quad (20)$$

where $f_{\{i\}}^{[0:n]} = 0$. Then, the matrix $\mathbf{F}_{\{i\}}^{[i:n]}[\mathfrak{R}_i, \mathfrak{C}_i]$ is full-rank.

C. Relaying Scheme

The source has a set of independent information bits that it wishes to transmit to the destination. We denote the set of information bits to be sent in block B by $\mathcal{A}(B) := \bigcup_{i\in[n+1]} \mathcal{A}(i,B)$, where

$$\mathcal{A}(i,B) := \left\{a_{t,B}^{i}(\ell) : t \in \mathsf{TW}(i,B), \ell \in \left[f_{\{i\}}^{\varnothing}\right]\right\}, \quad (22)$$

and we define $f_{\{n+1\}}^{\Omega} = f_{\varnothing}^{\Omega}$, for consistency. In state $\varnothing$, when all the relays are in receive mode, the source sends all the bits in $\mathcal{A}(n+1,B)$ uncoded, i.e., for $t \in \mathsf{TW}(n+1,B)$ we have

$$X_s^t(\ell) = \begin{cases} a_{t,B}^{n+1}(\ell) & \text{if } \ell \in \left[f_{\{n+1\}}^{\varnothing}\right], \\ 0 & \ell \in \left[f_{\{n+1\}}^{\varnothing} + 1 : \eta\right]. \end{cases}$$

$$X_s^t(\ell) = \begin{cases} a_{t,B}^i(\ell) & \text{if } \ell \in \left[f_{\{i\}}^{[i+1:n]}\right], \\ a_{t,B}^i(\ell) - X_i^t(\ell + \eta_{j,i} - \eta_{j,s}) & \text{if } \ell \in \left[f_{\{i\}}^{[j:n]} + 1 : f_{\{i\}}^{[j+1:n]}\right] \text{ and } j \in [i+1:n], \\ 0 & \text{if } \ell \in \left[f_{\{i\}}^{[n+1:n]} + 1 : \eta\right], \end{cases} \tag{21}$$

The transmission during the other states is more sophisticated, and it is described next. In state $\mathcal{S} = \{i\}$, i.e., during the transmit window $\mathsf{TW}(i, B)$ both the source and relay i are transmitting. The source aims at sending the information bits in $\mathcal{A}(i, B)$. Towards this end, the source adopts two different coding schemes for its transmission. In particular, we further split this set into $\mathcal{A}(i, B) = \overline{\mathcal{A}}(i, B) \cup \underline{\mathcal{A}}(i, B)$, where

$$\overline{\mathcal{A}}(i, B) := \left\{a_{t,B}^i(\ell) \in \mathcal{A}(i, B) : 1 \le \ell \le f_{\{i\}}^{[i+1:n]}\right\},$$
$$\underline{\mathcal{A}}(i, B) := \left\{a_{t,B}^i(\ell) \in \mathcal{A}(i, B) : f_{\{i\}}^{[i+1:n]} < \ell \le f_{\{i\}}^{\varnothing}\right\},$$

denote the set of information bits to be sent on the top $f_{\{i\}}^{[i+1:n]}$ levels and those to be sent on the next $f_{\{i\}}^{[\varnothing]} - f_{\{i\}}^{[i+1:n]}$ levels, respectively.

Let $X_i^t(\ell)$ denote the bit sent by relay i on its level ℓ at some $t \in \mathsf{TW}(i, B)$ (which will be determined later). It is important to note that the source is aware of each information bit sent by the relays. In particular, the source transmits $X_s^t(\ell)$ for $\ell \in [\eta]$, given in (21) at the top of this page, in a time slot $t \in \mathsf{TW}(i, B)$. Note that (21) implies that the source sends its information bits uncoded over its top $\left[f_{\{i\}}^{[i+1:n]}\right]$ levels, but it applies a (dirty paper) pre-coding on its remaining bits, which are sent over levels $\left[f_{\{i\}}^{[i+1:n]} + 1 : f_{\{i\}}^{\varnothing}\right]$. We will justify this pre-coding when we show decodability of the pre-coded information bits in Lemma 8.

Next, we identify the set of bits sent by relay i during state $\{i\}$, which is a *subset* of the bits that it has received during its reception window $\mathsf{RW}(i, B)$. This selection is determined by

$$\mathcal{B}(i, B) = \bigcup_{j \in [n+1] \setminus \{i\}} \mathcal{B}(i, B, j), \tag{23}$$

where[2]

$$\mathcal{B}(i, B, j) := \left\{Y_i^t(\eta - k) : t \in \mathsf{RW}(i, B, j),\right.$$
$$\left. k \in \left[0 : \left(f_{\{j\}}^{[i+1:n]} - f_{\{j\}}^{[i:n]}\right) - 1\right]\right\}. \tag{24}$$

In other words, during each reception window $\mathsf{RW}(i, B, j)$, among all the bits received, relay i only selects a subset $\mathcal{B}(i, B, j)$ of them, to forward during its transmit window $\mathsf{TW}(i, B)$, and it discards all the other received bits. It is worth emphasizing that in the proposed scheme, each relay i only *forwards* the selected bits in $\mathcal{B}(i, B)$, without any further processing or coding.

During its transmit window in block B, i.e., $\mathsf{TW}(i, B)$, relay i only transmits over its top $f_{\{i\}}^{[i:n]} - f_{\{i\}}^{[i+1:n]}$ levels. The proposed relaying scheme relies on having each relay $i \in [n]$ forward all the selected bits $\mathcal{B}(i, B)$. More precisely, we have

$$\left\{X_i^t(\ell) : t \in \mathsf{TW}(i, B), \ell \in \left[f_{\{i\}}^{[i:n]} - f_{\{i\}}^{[i+1:n]}\right]\right\} = \mathcal{B}(i, B). \tag{25}$$

[2]Note that Lemma 1 and Lemma 3 imply that $f_{\{j\}}^{[i+1:n]} \ge f_{\{j\}}^{[i:n]}$, for every $j \neq i$.

Note that relay i sends zero on its remaining levels, i.e., $X_i^t(\ell) = 0$ for all $\ell > f_{\{i\}}^{[i:n]} - f_{\{i\}}^{[i+1:n]}$.

The following proposition shows that the assignment in (25) is feasible, i.e., the number of bits that relay i can forward during its transmit window over its top $f_{\{i\}}^{[i:n]} - f_{\{i\}}^{[i+1:n]}$ levels matches with the number of bits in $\mathcal{B}(i, B)$ that it has selected during its reception window. We present the proof of the proposition in Appendix E.

Proposition 6 (Information Flow Preservation)**.** *For each relay i and each block B, the number of selected bits in $\mathsf{RW}(i, B)$ equals the number of transmitted bits by relay i during block B, i.e., $|\mathcal{B}(i, B)| = |\mathsf{TW}(i, B)| \cdot \left(f_{\{i\}}^{[i:n]} - f_{\{i\}}^{[i+1:n]}\right)$.*

D. Decoding the Information Bits at the Destination

We here show that all the information bits $\mathcal{A}(B)$ sent by the source during block B can be decoded at the destination by the end of block $B + 1$. More precisely, we show that the information bits in $\overline{\mathcal{A}}(B) := \bigcup_{i \in [n+1]} \overline{\mathcal{A}}(i, B)$ are decoded by the end of block B, but we have to wait until the end of block $B + 1$ to decode those in $\underline{\mathcal{A}}(B) := \bigcup_{i \in [n+1]} \underline{\mathcal{A}}(i, B)$.

The following proposition guarantees that all the bits sent by relay i as well as the *uncoded* information bits sent by the source during $\mathsf{TW}(i, B)$ can be decoded by the destination at the end of the block B. We refer to Appendix F for the proof of the proposition.

Proposition 7. *For every block B and every relay i, all the bits sent by relay i during block B and all the bits sent by the source on its top $f_{\{i\}}^{[i+1:n]}$ levels are decoded by the destination by the end of block B, that is,*

$$H\left(\overline{\mathcal{A}}(i, B), \mathcal{B}(i, B) \middle| \{Y_d^t(\ell) :\right.$$
$$\left. t \in [(B-1)T + 1 : BT], \ell \in [\eta]\}\right) = 0.$$

The next proposition shows that the remaining information bits sent by the source in block B, i.e., $\underline{\mathcal{A}}(B)$ will be uncodedly forwarded by the relays during block $B + 1$. The proof of this proposition is presented in Appendix G.

Proposition 8. *All the information bits in $\underline{\mathcal{A}}(B)$ will be re-transmitted by one of the relays in block $B + 1$, i.e., $\underline{\mathcal{A}}(B) \subseteq \bigcup_{j \in [n]} \mathcal{B}(j, B + 1)$.*

Now, we are ready to prove that all the information bits sent by s in block B can be decoded at d by the end of block $B + 1$, i.e., $H\left(\mathcal{A}(B) | \{Y_d^t(\ell) : t \in [(B-1)T + 1 : (B+1)T], \ell \in [\eta]\}\right) = 0$. Indeed, this is an immediate consequence of Proposition 7 and Proposition 8, and is shown by the chain of inequalities in (26), at the top of the next page. This proves that, given the received bits by the destination in blocks B and $B+1$, all the information bits in $\mathcal{A}(B)$ can be decoded at the destination.

$$
\begin{aligned}
&H\left(\mathcal{A}(B)\middle|\left\{Y_d^t(\ell): t\in[(B-1)T+1:(B+1)T], \ell\in[\eta]\right\}\right)\\
&= H\left(\underline{\mathcal{A}}(B), \overline{\mathcal{A}}(B)\middle|\left\{Y_d^t(\ell): t\in[(B-1)T+1:(B+1)T], \ell\in[\eta]\right\}\right)\\
&= H\left(\underline{\mathcal{A}}(B), \bigcup_{j\in[n+1]}\overline{\mathcal{A}}(j,B)\middle|\left\{Y_d^t(\ell): t\in[(B-1)T+1:(B+1)T], \ell\in[\eta]\right\}\right)\\
&\leq H\left(\underline{\mathcal{A}}(B), \bigcup_{j\in[n+1]}\overline{\mathcal{A}}(j,B), \bigcup_{j\in[n+1]}\mathcal{B}(j,B+1)\middle|\left\{Y_d^t(\ell): t\in[(B-1)T+1:(B+1)T], \ell\in[\eta]\right\}\right)\\
&\leq \sum_{j\in[n+1]} H\left(\bigcup_{j\in[n+1]}\overline{\mathcal{A}}(j,B)\middle|\left\{Y_d^t(\ell): t\in[(B-1)T+1:BT], \ell\in[\eta]\right\}\right)\\
&\quad+\sum_{j\in[n+1]} H\left(\bigcup_{j\in[n+1]}\mathcal{B}(j,B+1)\middle|\left\{Y_d^t(\ell): t\in[BT+1:(B+1)T], \ell\in[\eta]\right\}\right)+H\left(\underline{\mathcal{A}}(B)\middle|\bigcup_{j\in[n+1]}\mathcal{B}(j,B+1)\right)=0. \quad (26)
\end{aligned}
$$

E. Evaluation of the Achievable Rate

The analysis in Section V-D shows that all the information bits in (22) sent by the source during block B can be decoded by the destination by the end of block $B+1$. Therefore, the proposed scheme can achieve a stationary rate of $|\mathcal{A}(B)|$ per block. From (22) we have

$$
\begin{aligned}
&\frac{1}{T}|\mathcal{A}(B)| = \frac{1}{T}\sum_{j=1}^{n+1}|\mathsf{TW}(j,B)|f_{\{j\}}^{\varnothing} = \frac{1}{T}\sum_{j=1}^{n+1}T\lambda_{\{j\}}f_{\{j\}}^{\varnothing}\\
&= \sum_{j=1}^{n+1}(-1)^j\frac{\mathsf{P}_{(j)}}{\det(\mathbf{P})}f_{\{j\}}^{\varnothing} = \frac{1}{\det(\mathbf{P})}\sum_{j=1}^{n+1}(-1)^j\mathsf{P}_{(j)}f_{\{j\}}^{[n+1:n]}\\
&\stackrel{(a)}{=}\frac{1}{\det(\mathbf{P})}\left[\mathsf{P}_{(0)}-\det(\mathbf{P}_{n+1\to 0})\right] = \frac{\mathsf{P}_{(0)}}{\det(\mathbf{P})}, \quad (27)
\end{aligned}
$$

where (a) follows from the Laplace expansion with respect to the top row of matrix $\mathbf{P}_{n+1\to 0}$, given by $\det(\mathbf{P}_{n+1\to 0}) = \mathsf{P}_{(0)}+\sum_{j=1}^{n+1}(-1)^j\left(-f_{\{j\}}^{[n+1:n]}\right)\mathsf{P}_{(j)} = 0$. It is worth recalling that $\mathbf{P}_{n+1\to 0}$ is identical to $\mathbf{P}$, except its top row which is replaced by row $n+1$ of $\mathbf{P}$. Note that (27) proves an achievable rate which is equal to C^{U} in (10), as shown in Proposition 2.

VI. Conclusion

In this work, we have analyzed the linear deterministic approximation of the Gaussian noise model of a half-duplex relay network with *arbitrary* topology. Our main contribution is two-fold.

First, we have presented sufficient conditions to achieve the approximate capacity by operating the network with an *energy-efficient* schedule. Under such sufficient conditions, we have also provided closed form expressions for the optimal schedule and the approximate capacity.

Second, we have designed a time-block relaying scheme that operates in the energy-efficient states. The proposed scheme has several appealing practical features, such as: it ensures that the destination decodes all the information bits sent by the source in block B by the end of block $B+1$, it explicitly provides the information bits that each relay is exclusively responsible to store and forward to the destination, and it achieves a rate that is equal to the approximate capacity whenever the derived sufficient conditions are satisfied.

Appendix A
Proof of Proposition 1

Note that the conditions of Theorem 1 immediately implies that $\lambda_{\varnothing}^\star = \lambda_{\{n+1\}}^\star \geq 0$. Next, we prove that $\lambda_{\{i\}}^\star \geq 0$ for all $i\in[n]$. Towards this end, let $(\boldsymbol{\lambda}^\star, t^\star)$ be the solution of a system of linear equations constructed as follows: (i) setting $\lambda_{\mathcal{S}} = 0$, for all $\mathcal{S}\subseteq[n]$ with $\mathcal{S}\notin\mathbb{S}$ in (8); and (ii) forcing constraints $t\leq g_{[i:n]}$ for $i\in[n+1]$ and $g_p\leq 1$ in (8) to hold with equality. This system of $(n+2)$ linear equations in $(n+2)$ variables, is given by

$$
\mathbf{P}\begin{bmatrix}t & \lambda_{\{1\}} & \cdots & \lambda_{\{n\}} & \lambda_{\varnothing}\end{bmatrix}^T = \begin{bmatrix}1 & 0 & \cdots & 0 & 0\end{bmatrix}^T, \quad (28)
$$

and hence, $\boldsymbol{\lambda}^\star$ in Proposition 1 is indeed the solution of the above system of linear equations. The equation corresponding to the row $i+1$ for $i\in[0:n]$ of (28) is given by (30) at the top of the next page, where we define $\eta_{-1,s}=\eta_{0,s}=0$ and $\lambda_{\{0\}}^\star = 0$. In (a), we used Lemma 1 to evaluate $f_{\{j\}}^{[i+1:n]}$ as well as the fact that $\eta_{1,s}\leq\eta_{2,s}\leq\cdots\leq\eta_{n,s}$, and (b) follows from the identity $\lambda_{\varnothing}^\star+\sum_{j=1}^n\lambda_{\{j\}}^\star = 1$.

Using the above equation, we can recursively express $\lambda_{\{i\}}^\star$ in terms of $\{\lambda_{\{j\}}^\star : j>i\}$ for $i\in[0:n]$, which is given by

$$
\begin{aligned}
&\lambda_{\{i\}}^\star\left(\max\{\eta_{i,s},\eta_{d,s}\}-\max\{\eta_{(i-1),s},\eta_{d,s}\}\right)\\
&= (\max\{\eta_{i,s},\eta_{d,s}\}-t^\star)+\\
&\qquad\sum_{j=i+1}^{n}\lambda_{\{j\}}^\star\left(f_{\{j\}}^{[i+1:n]}-\max\{\eta_{i,s},\eta_{d,s}\}\right).
\end{aligned} \quad (29)
$$

Before we prove the claim, we present the following lemmas which will be used in the proof.

Lemma 3. *If* $\eta_{1,s}\leq\eta_{2,s}\leq\cdots\leq\eta_{n,s}$*, then*

$$
f_{\{j\}}^{[a:n]}-\max\{\eta_{a-1,s},\eta_{d,s}\}\geq f_{\{j\}}^{[b:n]}-\max\{\eta_{b-1,s},\eta_{d,s}\},
$$

for all $1\leq b<a\leq j\leq n$.

Lemma 4. *If* $\eta_{1,s}\leq\eta_{2,s}\leq\cdots\leq\eta_{n,s}$ *and there exists* $j\in[n]$ *such that*

$$
\max\{\eta_{j,s},\eta_{d,s}\} = \max\{\eta_{j-1,s},\eta_{d,s}\} = f_{\{j\}}^{[j:n]},
$$

then $\det(\mathbf{P}) = 0$.

$$
\begin{aligned}
t^\star &= \sum_{j=1}^{n+1} \lambda^\star_{\{j\}} f^{[i+1:n]}_{\{j\}} \\
&\overset{(a)}{=} \lambda^\star_{\varnothing} \max\{\eta_{i,s},\eta_{d,s}\} + \sum_{j=1}^{i-1} \lambda^\star_{\{j\}} \max\{\eta_{i,s},\eta_{d,s}\} + \lambda^\star_{\{i\}} \max\{\eta_{(i-1),s},\eta_{d,s}\} + \sum_{j=i+1}^{n} \lambda^\star_{\{j\}} f^{[i+1:n]}_{\{j\}} \\
&\overset{(b)}{=} \left(1 - \sum_{j=i}^{n} \lambda^\star_{\{j\}}\right) \max\{\eta_{i,s},\eta_{d,s}\} + \lambda^\star_{\{i\}} \max\{\eta_{i-1,s},\eta_{d,s}\} + \sum_{j=i+1}^{n} \lambda^\star_{\{j\}} f^{[i+1:n]}_{\{j\}} \\
&= \max\{\eta_{i,s},\eta_{d,s}\} + \lambda^\star_{\{i\}}(\max\{\eta_{i-1,s},\eta_{d,s}\} - \max\{\eta_{i,s},\eta_{d,s}\}) + \sum_{j=i+1}^{n} \lambda^\star_{\{j\}} \Big(f^{[i+1:n]}_{\{j\}} - \max\{\eta_{i,s},\eta_{d,s}\}\Big), \qquad (30)
\end{aligned}
$$

The proofs of Lemma 3 and Lemma 4 are in Appendix H and in Appendix I, respectively.

We now proceed to show that the $\lambda^\star_{\{i\}}$'s obtained from (29) (and hence, the $\lambda^\star_{\{i\}}$'s in Proposition 1) are non-negative for all $i \in [n]$. The proof is based on contradiction.

We start by assuming that the claim is wrong. Let $\mathcal{K} = \left\{ i : \lambda^\star_{\{i\}} < 0, i \in [n] \right\}$, and k be the maximum element of $\mathcal{K}$. We also define $\mathcal{L} = \{j \in [k-1] : \max\{\eta_{j,s},\eta_{d,s}\} = \max\{\eta_{(j-1),s},\eta_{d,s}\}\}$, and $\ell = \max \mathcal{L} \cup \{0\}$, where $\eta_{0,s} = 0$. Our goal is to show that $\mathcal{K}$ is an empty set and hence, $\lambda^\star_{\{i\}} \geq 0$ for all $i \in [n]$. We first use a *backward induction* to prove that for every $i \in [\ell+1 : k]$ we have $\lambda^\star_{\{i\}} \leq 0$. Note that for the base case of $i = k$, the assumption of $k \in \mathcal{K}$ implies that $\lambda^\star_{\{k\}} < 0$. Now, consider some $i \in [\ell+1 : k-1]$ and assume $\lambda^\star_{\{i+1\}}, \dots, \lambda^\star_{\{k\}} \leq 0$. Our goal is to show that $\lambda^\star_{\{i\}} \leq 0$. Note that for $\mathcal{S} = \{j\}$ and $\Omega = [i+1:n]$ with $j \in [i+1:n]$ we have $\mathcal{S}^c \cap \Omega^c = [i]$ and hence, Lemma 1 implies that $f^{[i+1:n]}_{\{j\}} \geq \max_{x\in[i]} \max\{\eta_{x,s},\eta_{d,s}\} = \max\{\eta_{i,s},\eta_{d,s}\}$, where the last equality follows since $\eta_{1,s} \leq \eta_{2,s} \leq \cdots \leq \eta_{n,s}$. Therefore, the coefficients of the form $\left(f^{[i+1:n]}_{\{j\}} - \max\{\eta_{i,s},\eta_{d,s}\}\right)$ in (29) are non-negative. Then, starting from (29), we can write the chain of inequalities in (31) at the top of the next page, where (a) holds by the induction assumption that $\left\{\lambda^\star_{\{i+1\}}, \dots, \lambda^\star_{\{k\}}\right\}$ are all non-positive, (b) is due to the fact that $\eta_{i,s} \leq \eta_{k,s}$ for $i \leq k$, (c) follows from Lemma 3 for $i < k < j$, and (d) follows from (29) with $i = k$. Finally, since $i > \ell$ we have $\max\{\eta_{i,s},\eta_{d,s}\} - \max\{\eta_{(i-1),s},\eta_{d,s}\} > 0$, which together with (31) implies $\lambda^\star_{\{i\}} \leq 0$.

Now, consider (31) for $i = \ell = \max \mathcal{L} \cup \{0\}$. Recall that if $\ell = 0$ we have $\lambda^\star_{\{\ell\}} = 0$, an if $\ell \in \mathcal{L}$ we have $\max\{\eta_{\ell,s},\eta_{d,s}\} - \max\{\eta_{\ell-1,s},\eta_{d,s}\} = 0$. Thus, the left-hand-side of (31) equals zero for i. Then, the chain of inequalities in (31) is feasible if and only if the four inequalities labeled by (a), (b), (c) and (e) hold with equality. From (b) we can conclude that $\max\{\eta_{\ell,s},\eta_{d,s}\} = \max\{\eta_{k,s},\eta_{d,s}\}$, and since $\eta_{\cdot,s}$'s are sorted in an increasing order, we have

$$
\begin{aligned}
\max\{\eta_{\ell,s},\eta_{d,s}\} &= \max\{\eta_{(\ell+1),s},\eta_{d,s}\} = \cdots \\
&= \max\{\eta_{(k-1),s},\eta_{d,s}\} = \max\{\eta_{k,s},\eta_{d,s}\}. \qquad (32)
\end{aligned}
$$

Thus, since $\max\{\eta_{\ell,s},\eta_{d,s}\} = \max\{\eta_{(\ell+1),s},\eta_{d,s}\}$ and $\ell = \max \mathcal{L} \cup \{0\}$, we can conclude that $\ell = k-1$ (otherwise $\ell+1$ also belongs to $\mathcal{L}$ and hence, ℓ is not the maximum entry of $\mathcal{L} \cup \{0\}$).

Now, for (a) to hold with equality, we can conclude that the first summation is zero. However, since $\lambda^\star_{\{i\}} \leq 0$ for $i \in [\ell+1 : k]$ and $f^{[i+1:n]}_{\{j\}} - \max\{\eta_{i,s},\eta_{d,s}\} \geq 0$, each term in the summation should be zero. In particular, for the term corresponding to $j = k$, since $\lambda^\star_{\{k\}} < 0$, we get

$$
\max\{\eta_{\ell,s},\eta_{d,s}\} = f^{[\ell+1:n]}_{\{k\}} = f^{[\ell+1:n]}_{\{\ell+1\}}, \qquad (33)
$$

where the second equality holds since, as we have shown above, $\ell+1 = k$. Therefore, from (32) and (33) we can conclude that the conditions of Lemma 4 are satisfied for $j = \ell+1$, and thus, Lemma 4 implies that $\det(\mathbf{P}) = 0$. This last conclusion is in contradiction with the assumption of Proposition 1. In other words, in order to have $\det(\mathbf{P}) \neq 0$ we need $\mathcal{K}$ to be an empty set and hence, $\lambda^\star_{\{i\}} \geq 0$ for all $i \in [n]$. This completes the proof of Proposition 1. □

Appendix B
Proof of Proposition 3

We let $\mathcal{L} = \left\{ i \in [n] : \max\{\eta_{i,s},\eta_{d,s}\} = \max\{\eta_{(i-1),s},\eta_{d,s}\} \right\}$ and $\ell = \max \mathcal{L} \cup \{0\}$. The proof consists of three parts. First, we show that $\mu_j = 0$ for all $j \in [\ell]$. Then, we prove that all non-zero μ_i's have the same sign for $i \in [\ell+1 : n+1]$. This together with the fact that $\sum_{i=1}^{n+1} \mu_i = 1$ guarantees $\mu_i \geq 0$ for all $i \in [n+1]$. Finally, $\mu_p \geq 0$ is implied from $\mu_p = \sum_{j=1}^{n+1} f^{[j:n]}_{\{i\}} \mu_j$.

Recall from (12) that μ_i's for $i \in [n+1] \cup \{p\}$ can be obtained by solving the equation

$$
\boldsymbol{\mu}\mathbf{P} = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 \end{bmatrix}. \qquad (34)
$$

Note that for any $k \in [n]$, the kth column of the identity in (34), is given by

$$
\boldsymbol{\mu}\mathbf{P}[[0:n+1],k] = \mu_p - \sum_{j=1}^{n+1} f^{[j:n]}_{\{k\}} \mu_j = 0. \qquad (35)
$$

Similarly, rewriting the equation for column $n+1$, we get

$$
\boldsymbol{\mu}\mathbf{P}[[0:n+1],n+1] = \mu_p - \sum_{j=1}^{n+1} f^{[j:n]}_{\varnothing} \mu_j = 0. \qquad (36)
$$

$$\lambda^\star_{\{i\}}\left(\max\{\eta_{i,s},\eta_{d,s}\}-\max\{\eta_{(i-1),s},\eta_{d,s}\}\right)=\left(\max\{\eta_{i,s},\eta_{d,s}\}-t^\star\right)+\sum_{j=i+1}^{n}\lambda^\star_{\{j\}}\left(f_{\{j\}}^{[i+1:n]}-\max\{\eta_{i,s},\eta_{d,s}\}\right)$$

$$=\left(\max\{\eta_{i,s},\eta_{d,s}\}-t^\star\right)+\sum_{j=i+1}^{k}\lambda^\star_{\{j\}}\left(f_{\{j\}}^{[i+1:n]}-\max\{\eta_{i,s},\eta_{d,s}\}\right)+\sum_{j=k+1}^{n}\lambda^\star_{\{j\}}\left(f_{\{j\}}^{[i+1:n]}-\max\{\eta_{i,s},\eta_{d,s}\}\right)$$

$$\overset{\text{(a)}}{\leq}\left(\max\{\eta_{i,s},\eta_{d,s}\}-t^\star\right)+\sum_{j=k+1}^{n}\lambda^\star_{\{j\}}\left(f_{\{j\}}^{[i+1:n]}-\max\{\eta_{i,s},\eta_{d,s}\}\right)$$

$$\overset{\text{(b)}}{\leq}\left(\max\{\eta_{k,s},\eta_{d,s}\}-t^\star\right)+\sum_{j=k+1}^{n}\lambda^\star_{\{j\}}\left(f_{\{j\}}^{[i+1:n]}-\max\{\eta_{i,s},\eta_{d,s}\}\right)$$

$$\overset{\text{(c)}}{\leq}\left(\max\{\eta_{k,s},\eta_{d,s}\}-t^\star\right)+\sum_{j=k+1}^{n}\lambda^\star_{\{j\}}\left(f_{\{j\}}^{[k+1:n]}-\max\{\eta_{k,s},\eta_{d,s}\}\right)\overset{\text{(d)}}{=}\lambda^\star_{\{k\}}\left(\max\{\eta_{k,s},\eta_{d,s}\}-\max\{\eta_{(k-1),s},\eta_{d,s}\}\right)\overset{\text{(e)}}{\leq}0, \quad (31)$$

Subtracting (35) from (36), we get

$$\begin{aligned}0&=\boldsymbol{\mu}\left(\mathbf{P}[[0:n+1],n+1]-\mathbf{P}[[0:n+1],k]\right)\\&=\sum_{j=1}^{n+1}\left(f_{\{k\}}^{[j:n]}-f_{\varnothing}^{[j:n]}\right)\mu_j\\&=\sum_{j=1}^{k}\left(f_{\{k\}}^{[j:n]}-f_{\varnothing}^{[j:n]}\right)\mu_j+\left(f_{\{k\}}^{[k+1:n]}-f_{\varnothing}^{[k+1:n]}\right)\mu_{k+1}\\&\quad+\sum_{j=k+2}^{n+1}\left(f_{\{k\}}^{[j:n]}-f_{\varnothing}^{[j:n]}\right)\mu_j\\&\overset{\text{(a)}}{=}\sum_{j=1}^{k}\left(f_{\{k\}}^{[j:n]}-\max\{\eta_{(j-1),s},\eta_{d,s}\}\right)\mu_j\\&\quad-\left(\max\{\eta_{k,s},\eta_{d,s}\}-\max\{\eta_{k-1,s},\eta_{d,s}\}\right)\mu_{k+1},\end{aligned}\quad (37)$$

where in (a) we used Lemma 1 to get $f_{\varnothing}^{[j:n]}=\max\{\eta_{j-1,s},\eta_{d,s}\}$, $f_{\{k\}}^{[k+1:n]}=\max\{\eta_{k-1,s},\eta_{d,s}\}$, and $f_{\{k\}}^{[j:n]}=\max\{\eta_{(j-1),s},\eta_{d,s}\}$ for $j\geq k+2$. Recall that $\eta_{0,s}=0$.

Note that the equation in (37) holds for every $k\in[n]$. Moreover, if $k<\ell$, such an equation only involves variables $\{\mu_1,\dots,\mu_\ell\}$. Lastly, for $k=\ell$, we have $\max\{\eta_{\ell,s},\eta_{d,s}\}=\max\{\eta_{(\ell-1),s},\eta_{d,s}\}$ and hence, the coefficient of $\mu_{\ell+1}$ will be zero. Thus, equation (37) for $k=\ell$ reduces to

$$\sum_{j=1}^{\ell}\left(f_{\{\ell\}}^{[j:n]}-\max\{\eta_{(j-1),s},\eta_{d,s}\}\right)\mu_j=0. \quad (38)$$

Hence, (37) for $k\in[\ell-1]$ and (38) provide a total of ℓ equations in ℓ variables $\{\mu_1,\dots,\mu_\ell\}$.

Let $\mathbf{Q}$ be an $(n+2)\times\ell$ matrix where its kth column[3] is $\mathbf{P}[[0:n+1],n+1]-\mathbf{P}[[0:n+1],k+1]$. Note that $\mathbf{Q}$ is obtained by elementary column operations on $\mathbf{P}$, and since $\mathbf{P}$ is full-rank, so is $\mathbf{Q}$, i.e., $\operatorname{rank}(\mathbf{Q})=\ell$. Moreover, (37) and (38) imply that the jth row of $\mathbf{Q}$ is zero, for $j\in\{0,\ell+1,\ell+2,\dots,n+1\}$. Hence, the remaining ℓ rows should be linearly independent, which implies that $\mathbf{Q}[[\ell],[0:\ell-1]]$ is full rank. Therefore, the unique solution for the system of equations obtained from (37) and (38), i.e., $[\mu_1,\cdots,\mu_\ell]\,\mathbf{Q}[[\ell],[0:\ell-1]]=\mathbf{0}_{1\times\ell}$, is an all-zero vector, i.e., $[\mu_1,\cdots,\mu_\ell]=\mathbf{0}_{1\times\ell}$. Next, we use induction to show that all non-zero μ_i's have the same sign, for all $i\in[\ell+1:n+1]$. First note that (37) for $k=\ell+1$ together with the fact that $\mu_i=0$ for $i\in[\ell]$ implies that

$$\begin{aligned}0&=\sum_{j=1}^{\ell+1}\left(f_{\{\ell+1\}}^{[j:n]}-\max\{\eta_{(j-1),s},\eta_{d,s}\}\right)\mu_j\\&\quad-\left(\max\{\eta_{(\ell+1),s},\eta_{d,s}\}-\max\{\eta_{\ell,s},\eta_{d,s}\}\right)\mu_{\ell+2}\\&=\left(f_{\{\ell+1\}}^{[\ell+1:n]}-\max\{\eta_{\ell,s},\eta_{d,s}\}\right)\mu_{\ell+1}\\&\quad-\left(\max\{\eta_{(\ell+1),s},\eta_{d,s}\}-\max\{\eta_{\ell,s},\eta_{d,s}\}\right)\mu_{\ell+2}.\end{aligned}$$

Note that $\max\{\eta_{(\ell+1),s},\eta_{d,s}\}-\max\{\eta_{\ell,s},\eta_{d,s}\}>0$ since ℓ is the maximum element of $\mathcal{L}\cup\{0\}$ and the left side link capacities are arranged in increasing order. Moreover, Lemma 1 implies that $\left(f_{\{\ell+1\}}^{[\ell+1:n]}-\max\{\eta_{\ell,s},\eta_{d,s}\}\right)\geq 0$. Therefore, we either have $\mu_{\ell+2}=0$, or $\operatorname{sign}(\mu_{\ell+2})=\operatorname{sign}(\mu_{\ell+1})$. This establishes the base case of the induction. Now, assume that our claim holds for every $j\leq k>\ell$, i.e., all non-zero μ_j's have the same sign for $j\leq k$. Then, from (37) we have

$$\begin{aligned}&\left(\max\{\eta_{k,s},\eta_{d,s}\}-\max\{\eta_{(k-1),s},\eta_{d,s}\}\right)\mu_{k+1}\\&\quad=\sum_{j=1}^{k}\left(f_{\{k\}}^{[j:n]}-\max\{\eta_{(j-1),s},\eta_{d,s}\}\right)\mu_j.\end{aligned}$$

Similar to the base case, we also note that $\max\{\eta_{k,s},\eta_{d,s}\}-\max\{\eta_{(k-1),s},\eta_{d,s}\}>0$, and that $f_{\{k\}}^{[j:n]}-\max\{\eta_{(j-1),s},\eta_{d,s}\}\geq 0$. Therefore, μ_{k+1} is either zero or its sign is identical to the one of the non-zero μ_j's with $j\in[\ell+1:k]$. This completes the induction, from which we can conclude that all non-zero μ_i's have the same sign for $i\in[\ell+1:n+1]$. Finally, since $\sum_{k=1}^{n+1}\mu_k=1$, this common sign has to be positive. Lastly, from $\mu_p=\sum_{j=1}^{n+1}f_{\{i\}}^{[j:n]}\mu_j$ for all $i\in[n+1]$, it directly follows that $\mu_p\geq 0$. This concludes the proof of Proposition 3. □

APPENDIX C
PROOF OF PROPOSITION 5

Note that each block of the transfer matrix $\mathbf{F}_{\{i\}}^{[i:n]}$ is an $\eta\times\eta$ matrix of the form $\mathbf{D}^{\eta-x}$, for some $x\in[0:\eta]$, i.e., each column in a block of $\mathbf{F}_{\{i\}}^{[i:n]}$ is an all-zero vector with possibly

[3] Recall that columns and rows of matrix $\mathbf{P}$ are indexed by numbers in $[0:n+1]$.

a single 1, where by construction the position of the 1 shifts one unit to the bottom from column j to column $j+1$. Since the same argument holds for all the blocks, it is not difficult to verify that for any $\ell = u\eta + v$ (for some $u \in \{0,1\}$ and $v \in [\eta]$) we have

$$\mathbf{D}^{\otimes i} c_\ell = \begin{cases} c_{\ell+1} & \text{if } v < \eta, \\ \mathbf{0}_{i\eta\times 1} & \text{if } v = \eta, \end{cases} \tag{39}$$

where $\mathbf{D}^{\otimes i} := \mathbf{I}_i \otimes \mathbf{D}$, with $\otimes$ denoting the Kronecker product. Now, if c_j is a linear combination of the columns in $\{c_\ell : \ell \in [j-1]\}$, we have $c_j = \sum_{\ell=1}^{j-1} \alpha_\ell c_\ell$. Then, for $j = k\eta + p$ with $p \in [\eta-1]$, by leveraging (39), we obtain

$$\begin{aligned} c_{j+1} = \mathbf{D}^{\otimes i} c_j &= \mathbf{D}^{\otimes i} \sum_{\ell\in[j-1]} \alpha_\ell c_\ell = \sum_{\ell\in[j-1]} \alpha_\ell \mathbf{D}^{\otimes i} c_\ell \\ &= \sum_{\ell\in[j-1]\setminus\{\eta\}} \alpha_\ell c_{\ell+1} = \sum_{\ell\in[2:j]\setminus\{\eta+1\}} \alpha_{\ell-1} c_\ell \\ &= \sum_{\ell\in[2:j-1]\setminus\{\eta+1\}} \alpha_{\ell-1} c_\ell + \alpha_{j-1} c_j \\ &= \sum_{\ell\in[2:j-1]\setminus\{\eta+1\}} \alpha_{\ell-1} c_\ell + \sum_{\ell\in[j-1]} \alpha_{j-1}\alpha_\ell c_\ell, \end{aligned}$$

which shows that c_{j+1} is a linear combination of the columns $\{c_\ell : \ell \in [j-1]\}$.

The proof of the second claim of Proposition 5 is similar to that of the first one. Here, row j of each block is an all-zero vector of length η, with possibly a single 1, the position of which shifts to the right, from one row to the next row. Applying this argument to the rows in one row-block, we can verify that for $\ell = u\eta + v$ (with some $u \in [0 : i-1]$ and $v \in [\eta]$) we have

$$r_\ell \mathbf{D}^{\otimes 2} = \begin{cases} r_{\ell-1} & \text{if } v > 1, \\ \mathbf{0}_{1\times 2\eta} & \text{if } v = 1. \end{cases}$$

Now, if r_j is a linear combination of $\{r_\ell : \ell \in [j-1]\}$, we have $r_j = \sum_{\ell\in[j-1]} \alpha_\ell r_\ell$. Therefore, for $j = k\eta + p$ with $p \in [2:\eta]$ we can write

$$\begin{aligned} r_{j-1} = r_j \mathbf{D}^{\otimes 2} &= \sum_{\ell\in[j-1]} \alpha_\ell r_\ell \mathbf{D}^{\otimes 2} \\ &= \sum_{\ell\in[1:j-1]\setminus\{1,\eta+1,\dots,(i-1)\eta+1\}} \alpha_\ell r_{\ell-1} = \sum_{\ell\in[j-2]\setminus\{\eta,\dots,(i-1)\eta\}} \alpha_{\ell+1} r_\ell, \end{aligned}$$

which expresses r_{j-1} as a linear combination of $\{r_\ell : \ell \in [j-2]\}$. This concludes the proof. $\square$

Appendix D
Proof of Corollary 1

We know that $\text{rank}\left(\mathbf{F}^{[i:n]}_{\{i\}}\right) = f^{[i:n]}_{\{i\}}$ and hence, there should be $f^{[i:n]}_{\{i\}}$ linearly independent columns in $\mathbf{F}^{[i:n]}_{\{i\}}$. Moreover, the first column-block of $\mathbf{F}^{[i:n]}_{\{i\}}$ is $\mathbf{F}^{[i+1:n]}_{\{i\}}$, which has $f^{[i+1:n]}_{\{i\}}$ linearly independent columns. Therefore, there should exist a collection of $f^{[i:n]}_{\{i\}} - f^{[i+1:n]}_{\{i\}}$ columns in the second block-column of $\mathbf{F}^{[i:n]}_{\{i\}}$ that, together with the $f^{[i+1:n]}_{\{i\}}$ in the first block-column, form a set of $f^{[i:n]}_{\{i\}}$ linearly independent vectors. We start selecting the columns in the first block from left-to-right. It follows from Proposition 5-*(i)* that, once we reach a column c_j that is a linear combination of the previously selected columns $\{c_\ell : \ell \in [j-1]\}$, then the remaining columns in the same block are also linear combinations of $\{c_\ell : \ell \in [j-1]\}$. Thus, we end up selecting the first (from the left) $f^{[i+1:n]}_{\{i\}}$ columns from the first column-block of $\mathbf{F}^{[i:n]}_{\{i\}}$. We continue by examining columns from the second column-block of $\mathbf{F}^{[i:n]}_{\{i\}}$. Again, Proposition 5-*(i)* implies that once we reach a column c_j in the second column-block that cannot be selected (because it is a linear combination of previously selected columns), all the columns c_k with $k > j$ are also linear combinations of the previously selected columns and hence, they cannot be selected. Hence, the collection of $f^{[i:n]}_{\{i\}} - f^{[i+1:n]}_{\{i\}}$ columns should be chosen from the most left columns of the second column-block of $\mathbf{F}^{[i:n]}_{\{i\}}$. This implies that the columns in $\mathfrak{C}_i$ are linearly independent.

Similarly, there exists a set of $f^{[i:n]}_{\{i\}}$ linearly independent rows in $\mathbf{F}^{[i:n]}_{\{i\}}$. We start scanning the rows from the top to the bottom, and add a row to the collection if it is linearly independent from the set of already selected rows. Note that Proposition 5-*(ii)* implies that, in any row-block $k \in [i]$, once we observe a row r_j that can be added to the collection, then all the remaining rows in the same row-block and below r_j can also be added to the collection. Hence, the set of selected rows in each row-block should be located at the bottom of the block. Now, recall that the first row-block is $\mathbf{F}^{[1:n]}_{\{i\}}$ which has $f^{[1:n]}_{\{i\}} = \text{rank}\left(\mathbf{F}^{[1:n]}_{\{i\}}\right)$ linearly independent rows which are located in its lowest part, i.e., rows $\left[\eta - f^{[1:n]}_{\{i\}} + 1 : \eta\right]$. The concatenation of the first and second row-blocks forms the transfer matrix $\mathbf{F}^{[2:n]}_{\{i\}}$ with a total of $f^{[2:n]}_{\{i\}} = \text{rank}\left(\mathbf{F}^{[2:n]}_{\{i\}}\right)$ linearly independent rows, out of which $f^{[1:n]}_{\{i\}}$ were already selected. Hence, the remaining $f^{[2:n]}_{\{i\}} - f^{[1:n]}_{\{i\}}$ rows are located at the bottom of the second row-block, i.e., $\left[2\eta - \left(f^{[2:n]}_{\{i\}} - f^{[1:n]}_{\{i\}}\right) + 1 : 2\eta\right]$. Repeating the same argument for row-block $k \in [i]$, we observe that the rows in $\left[k\eta - \left(f^{[k:n]}_{\{i\}} - f^{[k-1:n]}_{\{i\}}\right) + 1 : k\eta\right]$ will be added to the collection, until the set of linearly independent rows in $\mathfrak{R}_i$ is formed. This concludes the proof of Corollary 1. $\square$

Appendix E
Proof of Proposition 6

Let $\mathbf{P}_{i\to j}$ be a copy of the matrix $\mathbf{P}$ defined in (5), except that its row j is replaced by the row i of $\mathbf{P}$ (recall that for the matrix $\mathbf{P}$, we index rows and columns starting from 0). Since $\mathbf{P}_{i\to j}$ has two identical rows, it is rank-deficient and hence, its determinant is zero. Now, using the definition of $\mathcal{B}(i,B)$ in (23), we have the set of equalities in (40) at the top of the next page, where (a) follows from the definition of $\mathsf{RW}(i,B,j)$, in the two equalities labeled as (b) we replaced the values of $\lambda_{\{i\}}$'s from Proposition 1, in (c) we used the Laplace expansion of the determinant with respect to the first

$$
\begin{aligned}
|\mathcal{B}(i,B)| &= \sum_{j\in[n+1]\setminus\{i\}} |\mathsf{RW}(i,B,j)|\cdot\left(f_{\{j\}}^{[i+1:n]}-f_{\{j\}}^{[i:n]}\right) \overset{(a)}{=} \sum_{j\in[n+1]\setminus\{i\}} T\lambda_{\{j\}}\left(f_{\{j\}}^{[i+1:n]}-f_{\{j\}}^{[i:n]}\right)\\
&\overset{(b)}{=} \frac{T}{\det(\mathbf{P})}\left[\sum_{j\in[n+1]\setminus\{i\}}(-1)^j\mathsf{P}_{(j)}f_{\{j\}}^{[i+1:n]} - \sum_{j\in[n+1]\setminus\{i\}}(-1)^j\mathsf{P}_{(j)}f_{\{j\}}^{[i:n]}\right]\\
&\overset{(c)}{=} \frac{T}{\det(\mathbf{P})}\Big[\Big(-\det(\mathbf{P}_{i+1\to 0})-(-1)^i\mathsf{P}_{(i)}f_{\{i\}}^{[i+1:n]}+\mathsf{P}_{(0)}\Big)-\Big(-\det(\mathbf{P}_{i\to 0})-(-1)^i\mathsf{P}_{(i)}f_{\{i\}}^{[i:n]}+\mathsf{P}_{(0)}\Big)\Big]\\
&\overset{(d)}{=} \frac{T}{\det(\mathbf{P})}\left[(-1)^i\mathsf{P}_{(i)}f_{\{i\}}^{[i:n]}-(-1)^i\mathsf{P}_{(i)}f_{\{i\}}^{[i+1:n]}\right]\\
&= T\cdot(-1)^i\frac{\mathsf{P}_{(i)}}{\det(\mathbf{P})}\left(f_{\{i\}}^{[i:n]}-f_{\{i\}}^{[i+1:n]}\right) \overset{(b)}{=} T\lambda_{\{i\}}\left(f_{\{i\}}^{[i:n]}-f_{\{i\}}^{[i+1:n]}\right) \overset{(e)}{=} |\mathsf{TW}(i,B)|\left(f_{\{i\}}^{[i:n]}-f_{\{i\}}^{[i+1:n]}\right),
\end{aligned}
\tag{40}
$$

row for $\mathbf{P}_{i+1\to 0}$ and $\mathbf{P}_{i\to 0}$, that leads to

$$\det(\mathbf{P}_{i\to 0}) = \mathsf{P}_{(0)} + \sum_{j\in[n+1]}(-1)^j\left(-f_{\{j\}}^{[i:n]}\right)\mathsf{P}_{(j)},$$

and a similar expression for $\det(\mathbf{P}_{i+1\to 0})$, (d) holds since for matrices with identical rows we have $\det(\mathbf{P}_{i+1\to 0}) = \det(\mathbf{P}_{i\to 0}) = 0$, and finally in (e) we used the definition of $\mathsf{TW}(i,B)$ in (16). This completes the proof of Proposition 6. □

Appendix F
Proof of Proposition 7

We prove the proposition by induction over $i\in[n]$. For $i=1$, recall from (25) that relay 1 transmits a bit on its level ℓ only if it satisfies $\ell\in\left[f_{\{1\}}^{[1:n]}-f_{\{1\}}^{[2:n]}\right]$. On the one hand, if $\eta_{d,s}\geq\eta_{d,1}$, then from Lemma 1 $f_{\{1\}}^{[1:n]}-f_{\{1\}}^{[2:n]}=\max\{\eta_{d,1},\eta_{d,s}\}-\eta_{d,s}=0$ and hence, relay 1 remains silent. In this case, the destination receives all the information bits sent by the source from its top $f_{\{1\}}^{[2:n]}=\eta_{d,s}$ levels, without any interference. On the other hand, if $\eta_{d,s}<\eta_{d,1}$, then relay 1 transmits on its top $\eta_{d,1}-\eta_{d,s}$ levels, while the bits in $\overline{\mathcal{A}}(i,B)$ are sent on the top $\eta_{d,s}$ levels of the source. Then, the received bit at level $\eta-k$ of the destination is given by

$$
\begin{aligned}
Y_d^t(\eta-k) &= X_s^t(\eta_{d,s}-k)+X_1^t(\eta_{d,1}-k)\\
&= \begin{cases} X_s^t(\eta_{d,s}-k) & \text{if } k\in[0:\eta_{d,s}-1],\\ X_1^t(\eta_{d,1}-k) & \text{if } k\in[\eta_{d,s}:\eta_{d,1}-1],\end{cases}\\
&= \begin{cases} a_{t,B}^1(\eta_{d,s}-k) & \text{if } k\in[0:\eta_{d,s}-1],\\ b_{t,B}^1(\eta_{d,1}-k) & \text{if } k\in[\eta_{d,s}:\eta_{d,1}-1],\end{cases}
\end{aligned}
$$

where we let $b_{t,B}^i(\ell)$ be the (selected) bit sent by relay i on its level ℓ at a time instance $t\in\mathsf{TW}(i,B)$. This shows that the bits sent from the source and relay 1 are never combined and thus, the destination can decode all the bits in $\mathcal{B}(1,B)=\{b_{t,B}^1(\ell):t\in\mathsf{TW}(1,B),\ell\in[\eta_{d,1}-\eta_{d,s}]\}$ and those in $\overline{\mathcal{A}}(1,B)$. This establishes the basis of the induction.

Now, let's assume that the claim holds for every relay $j<i$, i.e., the destination can decode $\overline{\mathcal{A}}(j,B)\cup\mathcal{B}(j,B)$ for every $j<i$. We will show that for relay i the claim also holds. Let $t\in\mathsf{TW}(i,B)$ be a time slot at which relay i transmits, and consider the channel model from $\{s,i\}$ to $\{d,1,2,\dots,i-1\}$ at time t, where $\begin{bmatrix} Y_d^t & Y_1^t & \dots & Y_{i-1}^t\end{bmatrix}^T$ is equal to

$$
\mathbf{F}_{\{i\}}^{[i:n]}\begin{bmatrix}X_s^t\\X_i^t\end{bmatrix} = \left[\begin{array}{c|c} \mathbf{D}^{\eta-\eta_{d,s}} & \mathbf{D}^{\eta-\eta_{d,i}}\\ \hline \mathbf{D}^{\eta-\eta_{1,s}} & \mathbf{D}^{\eta-\eta_{1,i}}\\ \vdots & \vdots\\ \hline \mathbf{D}^{\eta-\eta_{(i-1),s}} & \mathbf{D}^{\eta-\eta_{(i-1),i}}\end{array}\right]\begin{bmatrix}X_s^t\\X_i^t\end{bmatrix}. \tag{41}
$$

Note that column j of $\mathbf{F}_{\{i\}}^{[i:n]}$ is zero for $j\in\left[f_{\{i\}}^{[i+1:n]}+1:\eta\right]$. Moreover, from (25) we have $X_i^t(\ell)=0$ for $\ell>f_{\{i\}}^{[i:n]}-f_{\{i\}}^{[i+1:n]}$. Therefore, from (41), $\begin{bmatrix} Y_d^t & Y_1^t & \dots & Y_{i-1}^t\end{bmatrix}^T$ reduces to

$$\mathbf{F}_{\{i\}}^{[i:n]}[[1:i\eta],\mathfrak{C}_i]\cdot\begin{bmatrix} X_s^t\left(\left[1:f_{\{i\}}^{[i+1:n]}\right]\right)\\ X_i^t\left(\left[1:f_{\{i\}}^{[i:n]}-f_{\{i\}}^{[i+1:n]}\right]\right)\end{bmatrix},$$

where $\mathfrak{C}_i$ is defined in (19). Then, focusing only on the rows in $\mathfrak{R}_i$ defined in (20), we get

$$
\begin{aligned}
&\begin{bmatrix} Y_d^t\left(\left[\eta-(f_{\{i\}}^{[1:n]}-f_{\{i\}}^{[0:n]})+1:\eta\right]\right)\\ Y_1^t\left(\left[\eta-(f_{\{i\}}^{[2:n]}-f_{\{i\}}^{[1:n]})+1:\eta\right]\right)\\ \vdots\\ Y_{i-1}^t\left(\left[\eta-(f_{\{i\}}^{[i:n]}-f_{\{i\}}^{[i-1:n]})+1:\eta\right]\right)\end{bmatrix}\\
&= \mathbf{F}_{\{i\}}^{[i:n]}[\mathfrak{R}_i,\mathfrak{C}_i]\cdot\begin{bmatrix} X_s^t\left(\left[1:f_{\{i\}}^{[i+1:n]}\right]\right)\\ X_i^t\left(\left[1:f_{\{i\}}^{[i:n]}-f_{\{i\}}^{[i+1:n]}\right]\right)\end{bmatrix}.
\end{aligned}
\tag{42}
$$

Recall from (21) that $X_s^t(\ell)=a_{t,B}^i(\ell)$ for $\ell\in\left[f_{\{i\}}^{[i+1:n]}\right]$. Similarly, (25) implies $X_i^t(\ell)=b_{t,B}^i(\ell)$ for $\ell\in\left[f_{\{i\}}^{[i:n]}-f_{\{i\}}^{[i+1:n]}\right]$. Proposition 5 guarantees that $\mathbf{F}_{\{i\}}^{[i:n]}[\mathfrak{R}_i,\mathfrak{C}_i]$ in (42) is full-rank. Therefore, we can recover the X matrix on the right-hand side of (42) from the Y matrix in the left-hand side of (42). Thus, we arrive at

$$
\begin{aligned}
H\Bigg(&\left\{a_{t,B}^i(\ell):\ell\in\left[f_{\{i\}}^{[i+1:n]}\right]\right\},\left\{b_{t,B}^i(\ell):\ell\in\left[f_{\{i\}}^{[i:n]}-f_{\{i\}}^{[i+1:n]}\right]\right\}\\
&\Bigg|\bigcup_{j=1}^{i-1}\left\{Y_j^t(\ell):\ell\in\left[\eta-\left(f_{\{i\}}^{[j+1:n]}-f_{\{i\}}^{[j:n]}\right)+1:\eta\right]\right\},\\
&\left\{Y_d^t(\ell):\ell\in\left[\eta-\left(f_{\{i\}}^{[1:n]}-f_{\{i\}}^{[0:n]}\right)+1:\eta\right]\right\}\Bigg)=0,
\end{aligned}
\tag{43}
$$

for every $t\in\mathsf{TW}(i,B)$. Summing up (43) for all $t\in\mathsf{TW}(i,B)$, and using $H(V_1,V_2|U_1,U_2)\leq H(V_1|U_1)+H(V_2|U_2)$, we arrive at (44) at the top of the next page.

$$H\left(\overline{\mathcal{A}}(i,B),\mathcal{B}(i,B)\middle|\bigcup_{j=1}^{i-1}\left\{Y_j^t(\ell)\colon t\in\mathsf{TW}(i,B),\ell\in\left[\eta-\left(f_{\{i\}}^{[j+1:n]}-f_{\{i\}}^{[j:n]}\right)+1:\eta\right]\right\},\right.$$
$$\left.\left\{Y_d^t(\ell)\colon t\in\mathsf{TW}(i,B),\ell\in\left[\eta-\left(f_{\{i\}}^{[1:n]}-f_{\{i\}}^{[0:n]}\right)+1:\eta\right]\right\}\right)=0, \tag{44}$$

Moreover, the induction assumption for $j<i$ implies that

$$\begin{aligned}&H\Bigg(\left\{Y_j^t(\ell)\!:\!t\!\in\!\mathsf{TW}(i,B),\ell\!\in[\eta\!-\!(f_{\{i\}}^{[j+1:n]}\!-\!f_{\{i\}}^{[j:n]})\!+\!1\!:\!\eta]\right\}\Bigg|\\&\qquad Y_d^t(\ell)\!:\!t\!\in[(B\!-\!1)T\!+\!1\!:\!BT],\ell\in[\eta]\Bigg)\\&\overset{(a)}{=}H\Bigg(\left\{Y_j^t(\eta\!-\!\ell)\!:\!t\!\in\!\mathsf{RW}(j,B,i),\ell\!\in\![0\!:\!f_{\{i\}}^{[j+1:n]}\!-\!f_{\{i\}}^{[j:n]}\!-\!1]\right\}\Bigg|\\&\qquad Y_d^t(\ell)\!:\!t\!\in[(B\!-\!1)T\!+\!1\!:\!BT],\ell\in[\eta]\Bigg)\\&\overset{(b)}{=}H\left(\mathcal{B}(j,B,i)\middle|Y_d^t(\ell)\!:\!t\in[(B\!-\!1)T\!+\!1\!:\!BT],\ell\in[\eta]\right)\\&\overset{(c)}{\le}H\left(\mathcal{B}(j,B)\middle|Y_d^t(\ell)\!:\!t\!\in\![(B\!-\!1)T\!+\!1\!:\!BT],\ell\!\in\![\eta]\right)\overset{(d)}{=}0,\end{aligned}\tag{45}$$

where (a) follows from (18) for $j<i$, (b) and (c) follow from (24) and (23), respectively, and (d) is guaranteed by the induction assumption for $j<i$.

Next, combining (44) and (45) we arrive at (46) at the top of the next page, which shows the claim of the induction for i. This completes the proof of Proposition 7. □

Appendix G
Proof of Proposition 8

Consider an information bit in $\underline{\mathcal{A}}(B)$, which is of the form $a_{t,B}^i(\ell)$ for some $i\in[n+1]$ and $\ell\in\left[f_{\{i\}}^{[i+1:n]}+1:f_{\{i\}}^{\varnothing}\right]$ and it is sent at time instance $t\in\mathsf{TW}(i,B)$. Consequently, we have $a_{t,B}^i(\ell)\in\underline{\mathcal{A}}(i,B)\subseteq\underline{\mathcal{A}}(B)$. Since $f_{\{i\}}^{[i+1:n]}\le f_{\{i\}}^{[i+2:n]}\le\cdots\le f_{\{i\}}^{[n:n]}\le f_{\{i\}}^{[n+1:n]}=f_{\{i\}}^{\varnothing}$, for every $\ell\in\left[f_{\{i\}}^{[i+1:n]}+1:f_{\{i\}}^{\varnothing}\right]$, there exists some $j\in[i+1:n]$ such that $\ell\in\left[f_{\{i\}}^{[j:n]}+1:f_{\{i\}}^{[j+1:n]}\right]$. In the following, we will show that $a_{t,B}^i(\ell)$ will be received by relay j and stored in its $\mathcal{B}(j,B+1)$. Note that the set $\left[f_{\{i\}}^{[j:n]}+1:f_{\{i\}}^{[j+1:n]}\right]$ is empty if $\eta_{d,s}>\eta_{j,s}$. Hence, in the rest of the proof, we may assume $\eta_{d,s}\le\eta_{j,s}$, which implies $f_{\{i\}}^{[j+1:n]}=\max\{\eta_{j,s},\eta_{d,s}\}=\eta_{j,s}$.

Recall from (21) that this information bit will be precoded as $a_{t,B}^i(\ell)-X_i^t(\ell+\eta_{j,i}-\eta_{j,s})$, where $X_i^t(k)$ is the bit sent by relay i at its level k. Then, the received bit at relay j on level $\eta-k$ for $k\in\left[0:f_{\{i\}}^{[j+1:n]}-f_{\{i\}}^{[j:n]}-1\right]$ is given by

$$\begin{aligned}Y_j^t(\eta-k)&=X_s^t(\eta_{j,s}-k)+X_i^t(\eta_{j,i}-k)\\&=\left(a_{t,B}^i(\eta_{j,s}-k)-X_i^t(\eta_{j,s}-k+\eta_{j,i}-\eta_{j,s})\right)\\&\quad+X_i^t(\eta_{j,i}-k)=a_{t,B}^i(\eta_{j,s}-k),\end{aligned}$$

which is an interference-free information bit. To show that $a_{t,B}^i(\ell)\in\mathcal{B}(j,B+1)$, we need to show that $t\in\mathsf{RW}(j,B+1,i)$ and $\ell=\eta_{j,s}-k$ for some $k\in\left[0\!:\!f_{\{i\}}^{[j+1:n]}-f_{\{i\}}^{[j:n]}-1\right]$. The first claim is an immediate consequence of (18) and the fact that $j\in[i+1:n]$. To show the second claim, we note that since $\ell\in\left[f_{\{i\}}^{[j:n]}+1:f_{\{i\}}^{[j+1:n]}\right]$ we have

$$\begin{aligned}k&=\eta_{j,s}-\ell\ge\eta_{j,s}-f_{\{i\}}^{[j+1:n]}=0,\\k&=\eta_{j,s}-\ell\le\eta_{j,s}-f_{\{i\}}^{[j:n]}-1=f_{\{i\}}^{[j+1:n]}-f_{\{i\}}^{[j:n]}-1,\end{aligned}$$

where we have used Lemma 1. The above inequalities imply that $k=\eta_{j,s}-\ell\in\left[0\!:\!f_{\{i\}}^{[j+1:n]}\!-\!f_{\{i\}}^{[j:n]}\!-\!1\right]$. This together with the fact that $t\in\mathsf{RW}(j,B\!+\!1,i)$ leads to $a_{t,B}^i(\ell)\in\mathcal{B}(j,B\!+\!1,i)\subseteq\mathcal{B}(j,B\!+\!1)$. This completes the proof of Proposition 8. □

Appendix H
Proof of Lemma 3

Recall that $\mathbf{F}_{\{j\}}^{[a:n]}$ for $j\in[a:n]$ is the transfer matrix from $X_{\{s,j\}}$ to $Y_{\{d,1,\dots,a-1\}}$, given by

$$\mathbf{F}_{\{j\}}^{[a:n]}=\left[\begin{array}{c|c}\mathbf{D}^{\eta-\eta_{d,s}}&\mathbf{D}^{\eta-\eta_{d,j}}\\\hline\mathbf{D}^{\eta-\eta_{1,s}}&\mathbf{D}^{\eta-\eta_{1,j}}\\\hline\vdots&\vdots\\\hline\mathbf{D}^{\eta-\eta_{a-1,s}}&\mathbf{D}^{\eta-\eta_{a-1,j}}\end{array}\right].$$

Similarly, for $b<a$, matrix $\mathbf{F}_{\{j\}}^{[b:n]}$ is given by the top b block rows of $\mathbf{F}_{\{j\}}^{[a:n]}$. Since $\eta_{1,s}\le\cdots\le\eta_{n,s}$, from the definition of $\mathbf{D}^{\eta-m}$ in (2), it follows that columns $[\max\{\eta_{(b-1),s},\eta_{d,s}\}+1:\eta]$ in $\mathbf{F}_{\{j\}}^{[b:n]}$ are zero, where $\eta_{0,s}=0$. Now, consider the lowest left block of $\mathbf{F}_{\{j\}}^{[a:n]}$, namely $\mathbf{D}^{\eta-\eta_{(a-1),s}}$. From (2), for every $\ell\in[\max\{\eta_{(b-1),s},\eta_{d,s}\}+1:\max\{\eta_{(a-1),s},\eta_{d,s}\}]$, the row $\eta-\eta_{(a-1),s}+\ell$ of the matrix $\mathbf{D}^{\eta-\eta_{(a-1),s}}$ has a one in column ℓ and zero elsewhere. Since $\mathbf{F}_{\{j\}}^{[b:n]}$ is fully zero in these columns, the row $\eta-\eta_{(a-1),s}+\ell$ of the matrix $\mathbf{F}_{\{j\}}^{[a:n]}$ is linearly independent from all rows in $\mathbf{F}_{\{j\}}^{[b:n]}$. Thus, $\mathbf{F}_{\{j\}}^{[a:n]}$ has at least $\max\{\eta_{(a-1),s},\eta_{d,s}\}-\max\{\eta_{(b-1),s},\eta_{d,s}\}$ additional linearly independent rows compared to $\mathbf{F}_{\{j\}}^{[b:n]}$. Thus, we get $f_{\{j\}}^{[a:n]}\ge f_{\{j\}}^{[b:n]}+(\max\{\eta_{a-1,s},\eta_{d,s}\}-\max\{\eta_{b-1,s},\eta_{d,s}\})$. This concludes the proof of Lemma 3. □

Appendix I
Proof of Lemma 4

We show that, under the assumption of $\eta_{1,s}\le\eta_{2,s}\le\cdots\le\eta_{n,s}$, if there exists some $j\in[n]$ such that $\max\{\eta_{j,s},\eta_{d,s}\}=\max\{\eta_{(j-1),s},\eta_{d,s}\}=f_{\{j\}}^{[j:n]}$, then columns j and $n+1$ of the matrix $\mathbf{P}$ are identical, i.e., $\mathbf{P}$ is singular and hence, $\det(\mathbf{P})=0$. Towards this end recall that from (5), $\mathbf{P}_{i,j}=-f_{\{j\}}^{[i:n]}$ for $(i,j)\in[n+1]\times[n]$ and the column j of the matrix $\mathbf{P}$ is given by

$$\begin{bmatrix}1&-f_{\{j\}}^{[1:n]}&-f_{\{j\}}^{[2:n]}&\cdots&-f_{\{j\}}^{[n:n]}&-f_{\{j\}}^{[n+1:n]}\end{bmatrix}^T.$$

$$H\left(\overline{\mathcal{A}}(i,B),\mathcal{B}(i,B)\middle|\left\{Y_d^t(\ell):t\in[(B-1)T+1:BT],\ell\in[\eta]\right\}\right)$$
$$\leq H\left(\overline{\mathcal{A}}(i,B),\mathcal{B}(i,B),\bigcup_{j=1}^{i-1}\left\{Y_j^t(\ell):t\in\mathsf{TW}(i,B),\ell\in\left[\eta-\left(f_{\{i\}}^{[j+1:n]}-f_{\{i\}}^{[j:n]}\right)+1:\eta\right]\right\}\middle|\left\{Y_d^t(\ell):t\in[(B-1)T+1:BT],\ell\in[\eta]\right\}\right)$$
$$\leq H\left(\overline{\mathcal{A}}(i,B),\mathcal{B}(i,B)\middle|\bigcup_{j=1}^{i-1}\left\{Y_j^t(\ell):t\in\mathsf{TW}(i,B),\ell\in\left[\eta-\left(f_{\{i\}}^{[j+1:n]}-f_{\{i\}}^{[j:n]}\right)+1:\eta\right]\right\},\right.$$
$$\left.\left\{Y_d^t(\ell):t\in\mathsf{TW}(i,B),\ell\in\left[\eta-\left(f_{\{i\}}^{[1:n]}-f_{\{i\}}^{[0:n]}\right)+1:\eta\right]\right\}\right)$$
$$+H\left(\bigcup_{j=1}^{i-1}\left\{Y_j^t(\ell):t\in\mathsf{TW}(i,B),\ell\in\left[\eta-\left(f_{\{i\}}^{[j+1:n]}-f_{\{i\}}^{[j:n]}\right)+1:\eta\right]\right\}\middle|\left\{Y_d^t(\ell):t\in[(B-1)T+1:BT],\ell\in[\eta]\right\}\right)=0, \quad (46)$$

Now, we evaluate each entry of the vector $\mathbf{P}[[0:n+1],j]$, for $j\in[n]$. Consider some $i\in[j+2:n+1]$. Using Lemma 1 for $\mathcal{S}=\{j\}$ and $\Omega=[i:n]$ with $\Omega\cap\mathcal{S}=\varnothing$ we have

$$f_{\{j\}}^{[i:n]}=\max_{t\in[i-1]\setminus\{j\}}\max\{\eta_{t,s},\eta_{d,s}\}=\max\{\eta_{i-1,s},\eta_{d,s}\}. \quad (47)$$

Similarly, for $i=j+1$ using Lemma 1 we get

$$f_{\{j\}}^{[i:n]}=\max_{t\in[i-1]\setminus\{j\}}\max\{\eta_{t,s},\eta_{d,s}\}=\max\{\eta_{j-1,s},\eta_{d,s}\}$$
$$\stackrel{(a)}{=}\max\{\eta_{j,s},\eta_{d,s}\}=\max\{\eta_{i-1,s},\eta_{d,s}\}, \quad (48)$$

where the equality in (a) follows from the assumption of the lemma. It remains to evaluate $f_{\{j\}}^{[i:n]}$ for $i\in[j]$. Towards this end, consider $\mathbf{F}_{\{j\}}^{[j:n]}$ which is defined as

$$\mathbf{F}_{\{j\}}^{[j:n]}=\left[\begin{array}{c|c}\mathbf{D}^{\eta-\eta_{d,s}} & \mathbf{D}^{\eta-\eta_{d,j}}\\ \hline \mathbf{D}^{\eta-\eta_{1,s}} & \mathbf{D}^{\eta-\eta_{1,j}}\\ \hline \vdots & \vdots\\ \hline \mathbf{D}^{\eta-\eta_{j-1,s}} & \mathbf{D}^{\eta-\eta_{j-1,j}}\end{array}\right],$$

where $\mathbf{D}^{\eta-m}$ is given in (2). Focusing on the first column-block of $\mathbf{F}_{\{j\}}^{[j:n]}$ we observe that each row in this $\eta\times j\eta$ matrix is either zero or if not zero and appearing in a row-block $t\in[j-2]$, it also appears in its lowest block, $\mathbf{D}^{\eta-\eta_{j-1,s}}$; this follows since $\eta_{1,s}\leq\eta_{2,s}\leq\cdots\leq\eta_{n,s}$. Hence, we have

$$\text{rank}\left(\left[\mathbf{D}^{\eta-\eta_{d,s}}\,\middle|\,\mathbf{D}^{\eta-\eta_{1,s}}\,\middle|\dots\middle|\,\mathbf{D}^{\eta-\eta_{(j-1),s}}\right]^T\right)$$
$$=\text{rank}\left(\left[\mathbf{D}^{\eta-\eta_{d,s}}\quad\mathbf{D}^{\eta-\eta_{(j-1),s}}\right]^T\right)$$
$$=\max\{\eta_{(j-1),s},\eta_{d,s}\}\stackrel{(a)}{=}f_{\{j\}}^{[j:n]}\stackrel{(b)}{=}\text{rank}\left(\mathbf{F}_{\{j\}}^{[j:n]}\right),$$

where in (a) we used the assumption of the lemma, and (b) follows from (4). In other words, the (column)-rank of $\mathbf{F}_{\{j\}}^{[j:n]}$ equals the (column)-rank of its first block column, or, every column in the second column block of $\mathbf{F}_{\{j\}}^{[j:n]}$ can be written as a linear combination of the columns in the first column block of $\mathbf{F}_{\{j\}}^{[j:n]}$. Next, note that for every $i\in[j]$ the matrix

$$\mathbf{F}_{\{j\}}^{[i:n]}=\left[\begin{array}{c|c}\mathbf{D}^{\eta-\eta_{d,s}} & \mathbf{D}^{\eta-\eta_{d,j}}\\ \hline \mathbf{D}^{\eta-\eta_{1,s}} & \mathbf{D}^{\eta-\eta_{1,j}}\\ \hline \vdots & \vdots\\ \hline \mathbf{D}^{\eta-\eta_{i-1,s}} & \mathbf{D}^{\eta-\eta_{i-1,j}}\end{array}\right]$$

is a submatrix of $\mathbf{F}_{\{j\}}^{[j:n]}$ and hence, the same conclusion holds for $\mathbf{F}_{\{j\}}^{[i:n]}$, i.e., each column in the second column-block of $\mathbf{F}_{\{j\}}^{[i:n]}$ is also a linear combination of the columns in the first column block of $\mathbf{F}_{\{j\}}^{[i:n]}$. Thus, the rank of $\mathbf{F}_{\{j\}}^{[i:n]}$ equals the rank of its first column block, which leads to

$$f_{\{j\}}^{[i:n]}=\text{rank}\left(\mathbf{F}_{\{j\}}^{[i:n]}\right)$$
$$=\text{rank}\left(\left[\mathbf{D}^{\eta-\eta_{d,s}}\,\middle|\,\mathbf{D}^{\eta-\eta_{1,s}}\,\middle|\dots\middle|\,\mathbf{D}^{\eta-\eta_{i-1,s}}\right]^T\right)$$
$$=\max_{y\in[i-1]\cup\{d\}}\text{rank}\left(\mathbf{D}^{\eta-\eta_{y,s}}\right)$$
$$\stackrel{(a)}{=}\max_{y\in[i-1]}\max\{\eta_{y,s},\eta_{d,s}\}\stackrel{(b)}{=}\max\{\eta_{i-1,s},\eta_{d,s}\}, \quad (49)$$

where (a) is due to the fact that the rank of $\mathbf{D}^{\eta-m}$ equals m, and (b) follows from the fact that $\eta_{1,s}\leq\eta_{2,s}\leq\cdots\leq\eta_{n,s}$. Hence, using (47)–(49), the entries of $\mathbf{P}[[0:n+1],j]$ for $j\in[n]$ can be evaluated, and are given by

$$\left[1\ \ -\max\{\eta_{0,s},\eta_{d,s}\}\ \cdots\ -\max\{\eta_{n,s},\eta_{d,s}\}\right]^T,$$

where $\eta_{0,s}=0$. Using (5) for $j=n+1$ and Lemma 1 for $\mathcal{S}=[n+1:n]=\varnothing$ and $\Omega=[i:n]$ with $\Omega\cap\mathcal{S}=\varnothing$, $\mathbf{P}_{i,n+1}$ equals

$$-f_{\{n+1\}}^{[i:n]}=-\max_{t\in[i-1]}\max\{\eta_{t,s},\eta_{d,s}\}=-\max\{\eta_{(i-1),s},\eta_{d,s}\},$$

which shows columns j and $(n+1)$ of matrix $\mathbf{P}$, are identical. This concludes the proof. □

References

[1] S. Jain, M. Cardone, and S. Mohajer, "Operating half-duplex diamond networks with two interfering relays," in *IEEE Information Theory Workshop (ITW)*, July 2021.

[2] ——, "When an Energy-Efficient Scheduling is Optimal for Half-Duplex Relay Networks?" in *IEEE International Symposium on Information Theory (ISIT)*, July 2021.

[3] A. Asadi, Q. Wang, and V. Mancuso, "A survey on device-to-device communication in cellular networks," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 4, pp. 1801–1819, Fourthquarter 2014.

[4] M. Hasan and E. Hossain, "Distributed resource allocation for relay-aided device-to-device communication: A message passing approach," *IEEE Trans. Wireless Commun.*, vol. 13, no. 11, pp. 6326–6341, November 2014.

[5] S. Biswas, S. Vuppala, J. Xue, and T. Ratnarajah, "On the performance of relay aided millimeter wave networks," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 576–588, April 2016.

[6] Y. Niu, Y. Li, D. Jin, L. Su, and A. V. Vasilakos, "A survey of millimeter wave communications (mmwave) for 5G: Opportunities and challenges," *Wireless Networks*, vol. 21, no. 8, pp. 2657–2676, Nov. 2015. [Online]. Available: https://doi.org/10.1007/s11276-015-0942-z

[7] T. Aktas, G. Quer, T. Javidi, and R. R. Rao, "From connected vehicles to mobile relays: Enhanced wireless infrastructure for smarter cities," in *IEEE Global Communications Conference (GLOBECOM)*, December 2016, pp. 1–6.

[8] J. Scheim and N. Lavi, "Vehicular relay nodes for cellular deployment: Downlink channel modeling and analysis," in *IEEE International Conference on Microwaves, Communications, Antennas and Electronic Systems (COMCAS)*, 2013, pp. 1–5.

[9] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, "Unmanned aerial vehicle with underlaid device-to-device communications: Performance and tradeoffs," *IEEE Trans. Wireless Commun.*, vol. 15, no. 6, pp. 3949–3963, 2016.

[10] Y. Zeng, R. Zhang, and T. J. Lim, "Wireless communications with unmanned aerial vehicles: opportunities and challenges," *IEEE Commun. Mag.*, vol. 54, no. 5, pp. 36–42, May 2016.

[11] M. Duarte, A. Sabharwal, V. Aggarwal, R. Jana, K. K. Ramakrishnan, C. W. Rice, and N. K. Shankaranarayanan, "Design and characterization of a full-duplex multiantenna system for wifi networks," *IEEE Trans. Veh. Technol.*, vol. 63, no. 3, pp. 1160–1177, March 2014.

[12] E. Everett, C. Shepard, L. Zhong, and A. Sabharwal, "Softnull: Many-antenna full-duplex wireless via digital beamforming," *IEEE Trans. Wireless Commun.*, vol. 15, no. 12, pp. 8077–8092, December 2016.

[13] M. Jain, J. I. Choi, T. Kim, D. Bharadia, S. Seth, K. Srinivasan, P. Levis, S. Katti, and P. Sinha, "Practical, real-time, full duplex wireless," in *17th Annual International Conference on Mobile Computing and Networking (MobiCom)*, September 2011, pp. 301–312. [Online]. Available: http://doi.acm.org/10.1145/2030613.2030647

[14] K. E. Kolodziej, B. T. Perry, and J. S. Herd, "In-band full-duplex technology: Techniques and systems survey," *IEEE Transactions on Microwave Theory and Techniques*, vol. 67, no. 7, pp. 3025–3041, 2019.

[15] X. Xia, K. Xu, Y. Wang, and Y. Xu, "A 5g-enabling technology: Benefits, feasibility, and limitations of in-band full-duplex mmimo," *IEEE Vehicular Technology Magazine*, vol. 13, no. 3, pp. 81–90, 2018.

[16] Y. P. E. Wang, X. Lin, A. Adhikary, A. Grovlen, Y. Sui, Y. Blankenship, J. Bergman, and H. S. Razaghi, "A primer on 3gpp narrowband internet of things," *IEEE Commun. Mag.*, vol. 55, no. 3, pp. 117–123, March 2017.

[17] A. S. Avestimehr, S. N. Diggavi, and D. N. C. Tse, "Wireless network information flow: A deterministic approach," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 1872–1905, April 2011.

[18] A. Özgür and S. N. Diggavi, "Approximately achieving Gaussian relay network capacity with lattice-based QMF codes," *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 8275–8294, December 2013.

[19] S. Lim, Y.-H. Kim, A. El Gamal, and S.-Y. Chung, "Noisy network coding," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 3132 –3152, 2011.

[20] S. H. Lim, K. T. Kim, and Y. H. Kim, "Distributed decode-forward for multicast," in *IEEE International Symposium on Information Theory (ISIT)*, June 2014, pp. 636–640.

[21] M. Cardone, D. Tuninetti, R. Knopp, and U. Salim, "Gaussian half-duplex relay networks: improved constant gap and connections with the assignment problem," *IEEE Trans. Inf. Theory*, vol. 60, no. 6, pp. 3559 – 3575, June 2014.

[22] L. Huang and Y. Hong, "On the capacity of a two-hop half-duplex relay channel with a markovian constrained relay," *IEEE Access*, vol. 7, pp. 15 683–15 695, 2019.

[23] L. Ong, M. Motani, and S. J. Johnson, "On capacity and optimal scheduling for the half-duplex multiple-relay channel," *IEEE Transactions on Information Theory*, vol. 58, no. 9, pp. 5770–5784, 2012.

[24] N. Zlatanov, V. Jamali, and R. Schober, "On the capacity of the two-hop half-duplex relay channel," in *2015 IEEE Global Communications Conference (GLOBECOM)*, 2015, pp. 1–7.

[25] M. Cardone, D. Tuninetti, and R. Knopp, "On the optimality of simple schedules for networks with multiple half-duplex relays," *IEEE Trans. Inf. Theory*, vol. 62, no. 7, pp. 4120–4134, July 2016.

[26] Y. H. Ezzeldin, M. Cardone, C. Fragouli, and D. Tuninetti, "Efficiently finding simple schedules in Gaussian half-duplex relay line networks," in *IEEE International Symposium on Information Theory (ISIT)*, June 2017, pp. 471–475.

[27] R. H. Etkin, F. Parvaresh, I. Shomorony, and A. S. Avestimehr, "Computing half-duplex schedules in Gaussian relay networks via min-cut approximations," *IEEE Trans. Inf. Theory*, vol. 60, no. 11, pp. 7204–7220, November 2014.

[28] H. Bagheri, A. S. Motahari, and A. K. Khandani, "On the capacity of the half-duplex diamond channel," in *2010 IEEE International Symposium on Information Theory*, 2010, pp. 649–653.

[29] S. Jain, M. Elyasi, M. Cardone, and S. Mohajer, "On simple scheduling in half-duplex relay diamond networks," in *IEEE International Symposium on Information Theory (ISIT)*, July 2019.

[30] H. Xie, Y. Zhan, G. Zeng, and X. Pan, "Leo mega-constellations for 6g global coverage: Challenges and opportunities," *IEEE Access*, vol. 9, pp. 164 223–164 244, 2021.

[31] M. Chen, R. Chai, and Q. Chen, "Joint route selection and resource allocation algorithm for data relay satellite systems based on energy efficiency optimization," in *2019 11th International Conference on Wireless Communications and Signal Processing (WCSP)*, 2019, pp. 1–6.

[32] A. Nordrum, "Facebook pushes networking tech: The company's terragraph technology will soon be available in commercial gear - [news]," *IEEE Spectrum*, vol. 56, no. 4, pp. 8–9, 2019.

Sarthak Jain is currently pursuing his PhD in the Electrical and Computer Engineering department at the University of Minnesota (UMN) under the supervision of professors Martina Cardone and Soheil Mohajer. He completed his B.E. (Hons.) in Electrical and Electronics Engineering in 2018 from Birla Institute of Technology and Science, Pilani Campus, India. Sarthak's main research interests are in information theory, algorithms, and optimization.

Martina Cardone is currently an Assistant Professor in the Electrical and Computer Engineering department at the University of Minnesota (UMN). Dr. Cardone received her Ph.D. in Electronics and Communications in 2015 from Télécom ParisTech (with work done at Eurecom in Sophia Antipolis, France). From November 2017 to January 2018, she was a postdoctoral associate in the Electrical and Computer Engineering department at UMN. From July 2015 to August 2017, she was a postdoctoral research fellow in the Electrical and Computer Engineering Department at UCLA Henry Samueli School. She is the recipient of the 2022 McKnight Land-Grant Professorship, the NSF CAREER award in 2021, the NSF CRII award in 2019, the second prize in the Outstanding Ph.D. award, Télécom ParisTech, Paris, France and the Qualcomm Innovation Fellowship in 2014. Dr. Cardone's main research interests are in estimation theory, network information theory, network coding, and wireless networks with a special focus on their capacity, security and privacy aspects.

Soheil Mohajer received the B.Sc. degree in electrical engineering from the Sharif University of Technology, Tehran, Iran, in 2004, and the M.Sc. and Ph.D. degrees in communication systems from the Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland, in 2005 and 2010, respectively. He was affiliated as a Post-Doctoral Researcher with Princeton University from 2010 to 2011, and as a Post-Doctoral Fellow with the University of California at Berkeley from 2011 to 2013. He joined the University of Minnesota, Twin Cities, in 2013, where he is currently a McKnight Land-Grant Associate Professor with the Department of Electrical and Computer Engineering. Dr. Mohajer received the NSF CAREER Award in 2018. He currently serves as an editor of IEEE Transactions on Communications. His research interests include information theory and its applications in distributed storage systems, wireless networks, statistical machine learning, and bioinformatics.