When an Energy-Efficient Scheduling is Optimal for Half-Duplex Relay Networks?

Sarthak Jain, Martina Cardone, Soheil Mohajer
University of Minnesota, Minneapolis, MN 55455, USA, Email: {jain0122, mcardone, soheil}@umn.edu

***Abstract*—This paper considers a diamond network with n interconnected relays, namely a network where a source communicates with a destination by hopping information through n communicating/interconnected relays. Specifically, the main focus of the paper is on characterizing sufficient conditions under which the $n+1$ states (out of the 2^n possible ones) in which at most one relay is transmitting suffice to characterize the approximate capacity, that is the Shannon capacity up to an additive gap that only depends on n. Furthermore, under these sufficient conditions, closed form expressions for the approximate capacity and scheduling (that is, the fraction of time each relay should receive and transmit) are provided. A similar result is presented for the dual case, where in each state at most one relay is in receive mode.**

I. Introduction

Computing the Shannon capacity of a wireless relay network is an open problem. In a half-duplex n-relay network, each relay can either transmit or receive at a given time instant and therefore a *scheduling* question arises: What fraction of time each relay in the network should be scheduled to receive/transmit information so that rates close to the Shannon capacity of the network can be achieved?

For an n-relay half-duplex network, there are 2^n possible receive/transmit configuration states, because each relay can either be scheduled for reception or transmission. However, in [1], it has been surprisingly shown that only $n+1$ out of these 2^n possible states are sufficient to achieve the network *approximate capacity*, i.e., an additive gap approximation of the Shannon capacity, where the gap is only a function of n. This result opens novel research directions, such as characterizing a set of $n+1$ *critical* states for each network efficiently (in polynomial time in n).

In this work, we investigate the question above in the context of diamond networks with n interconnected relays, where the source communicates with the destination by hopping information through n half-duplex relays that can communicate with each other. In particular, we analyze the linear deterministic approximation of the Gaussian noise channel, and characterize sufficient conditions under which at most one relay is required to transmit at any given time to achieve the approximate capacity. This 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. The other advantage of a schedule with at most one relay in transmit mode is that it simplifies the synchronization problem at the destination. Our result can 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. The proposed scheduling and approximate capacity can be translated to obtain similar results for the practically relevant Gaussian noise channel.

This research was supported in part by NSF under Award #1907785.

To the best of our knowledge, this is the first work that provides network conditions that suffice to characterize a set of $n+1$ critical states for arbitrary values of n in relay networks where, in addition to broadcasting and signal superposition, we also have signal interference at the relays.

Related Work. The cut-set bound has been shown to offer a constant (i.e., which only depends on n) additive gap approximation of the Shannon capacity for Gaussian relay networks [2]–[6]. Such approximation, for an n-relay Gaussian half-duplex network can be computed by solving a linear program involving 2^n cut constraints and 2^n variables corresponding to the receive/transmit configurations of the n half-duplex relays. However, it has been shown that it suffices to operate the network in only $n+1$ states out of the 2^n possible ones to achieve the approximate capacity [1]. Finding this set of $n+1$ critical states in polynomial time in n for half-duplex Gaussian relay networks is an open problem. These critical states and the approximate capacity can be computed in polynomial time for the following networks: (i) $n=2$ relay half-duplex diamond networks with *non-interconnected* relays [7] and *interconnected* relays [8]; (ii) line networks [9]; (iii) a special class of layered networks [10]; and (iv) diamond networks with n non-interconnected relays under certain network conditions expressed in [11]. We highlight that the result presented in this paper subsumes the result for diamond networks studied in [11] (where there is no signal interference among the relays) and our recent result in [8] for $n=2$.

Paper Organization. Section II introduces the notation, describes the Gaussian and the linear deterministic half-duplex diamond network with n interconnected relays and summarizes known capacity results. Section III presents the 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) suffice to characterize the approximate capacity of the binary-valued linear deterministic

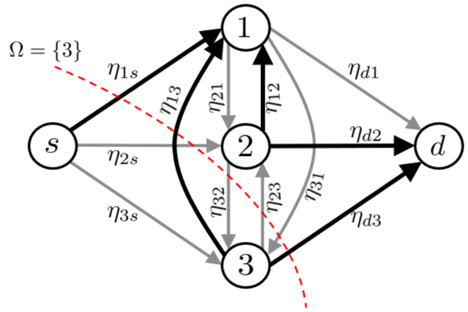

Fig. 1: Diamond network with $n = 3$ interconnected relays (with cut $\Omega = \{3\}$ and state $\mathcal{S} = \{2, 3\}$).

approximation of the Gaussian noise channel. Some of the proofs can be found in the appendix.

II. Notation and System Model

Notation: We denote the set of integers $\{i, \ldots, m\}$ by $[i : m]$, and $\{1, \ldots, m\}$ 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 $\mathcal{B}$. Matrix columns and rows are indexed beginning from 0 (instead of 1). $\lfloor\cdot\rfloor$ 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 diamond network with n interconnected relays consists of a source (node s) that wishes to communicate with a destination (node d) through n *interconnected* relays. At each time instant t, the input/output relationship of this network is described as

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

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) $X_i(t)$ is the channel input at node i that satisfies the unit average power constraint $\mathbb{E}[|X_i(t)|^2] \le 1$ for $i \in \{s\} \cup [n]$; (iii) 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 $h_{ds} = 0$ and, since the relays operate in half-duplex mode, without loss of generality we let $h_{ii} = 0$; (iv) $Z_i(t) \sim \mathcal{CN}(0, 1)$ is the complex additive white Gaussian noise at node $i \in \{d\} \cup [n]$; and finally, (v) $Y_i(t)$ is the received signal at node $i \in \{d\} \cup [n]$.

The Shannon capacity C^{G} of the network in (1) is not known for general n. However, the capacity can be approximated within a constant $O(n)$ bit gap. More precisely, we can focus on the binary *linear deterministic* approximation of the Gaussian noise network model [2], for which the approximate capacity is known and provides an approximation for C^{G}. The linear deterministic model (a.k.a. ADT model [2]) corresponding to the Gaussian noise network in (1) has an input-output relationship given by

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

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

$$\eta_{ij} = \left\lceil \log|h_{ij}|^2 \right\rceil^+, \ i \in [n]\cup\{d\}, j \in \{s\}\cup[n], i \ne j.$$

Here, the vectors $X_s(t)$, $X_i(t)$, $Y_d(t)$, and $Y_i(t)$ with $i \in [n]$ are binary of length $\eta = \max \eta_{ij}$, where the maximization is taken over all channels η_{ij} in the network; $\mathbf{D}$ is the so-called $\eta \times \eta$ shift matrix, and $S_i(t), i \in [n]$ is the i-th relay binary-value state random variable.

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

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

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

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

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 [2].

It turns out that $|\mathsf{C}^{\mathrm{G}} - \mathsf{C}^{\mathrm{LD}}| \le \kappa$, where $\kappa = O(n)$ is independent of the channel gains and operating SNR and hence, C^{LD} in (3) provides an approximation for the Shannon capacity of the network in (1)[1].

Example 1. Consider the relay diamond network with $n = 3$ interconnected relays 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 $\{d\} \cup (\Omega^c \cap$

[1]We highligth that schemes such as quantize-map-and-forward [2] and noisy network coding [4], together with the cut set bound, allow to characterize the capacity of Gaussian relay networks up to a constant additive gap.

$\mathcal{S}^c) = \{d, 1\}$. The input-output relationship for this cut and state is given by

$$\begin{bmatrix} Y_d \\ Y_1 \end{bmatrix} = \begin{bmatrix} \mathbf{D}^{\eta} & \mathbf{D}^{\eta-\eta_{d3}} \\ \mathbf{D}^{\eta-\eta_{1s}} & \mathbf{D}^{\eta-\eta_{13}} \end{bmatrix} \begin{bmatrix} X_s \\ X_3 \end{bmatrix}.$$

Therefore, we have

$$f_{\mathcal{S}}^{\Omega} = \det\left(\mathbf{F}_{\{2,3\}}^{\{3\}}\right) = \det \begin{bmatrix} \mathbf{D}^{\eta} & \mathbf{D}^{\eta-\eta_{d3}} \\ \mathbf{D}^{\eta-\eta_{1s}} & \mathbf{D}^{\eta-\eta_{13}} \end{bmatrix}. \qquad \diamond$$

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) of the linear deterministic network and can be readily translated into a similar result for the original noisy Gaussian channel model in (1).

III. Main Result: 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, $\eta_{1s} \leq \eta_{2s} \leq \cdots \leq \eta_{ns}$. 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

$$\mathbf{P}_{i,j} = \begin{cases} -f_{\{j\}}^{[i:n]}, & (i,j) \in [n+1]^2, \\ 0, & (i,j) = (0,0), \\ 1, & \text{otherwise}, \end{cases} \tag{4}$$

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

$$\mathbb{S} \triangleq \{\{1\}, \{2\}, \dots, \{n\}, \varnothing\},$$

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. *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 2. Consider the relay network in Fig. 1 with link capacities given by $\eta_{1s} = 1, \eta_{2s} = 3, \eta_{3s} = 5, \eta_{d1} = 6, \eta_{d2} = 5, \eta_{d3} = 3, \eta_{12} = 3, \eta_{21} = 4, \eta_{32} = 5, \eta_{23} = 3, \eta_{31} = 2$ and $\eta_{13} = 4$. For this network, the matrix $\mathbf{P}$ is given by

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

It is not difficult to check that $\det(\mathbf{P}) \neq 0$ and $\mathsf{P}_{(4)} = \det\left(\mathbf{P}_{[4],[0:3]}\right) \geq 0$, i.e., the conditions in Theorem 1 are satisfied. Thus, operating this network in $\mathbb{S} = \{\{1\}, \{2\}, \{3\}, \varnothing\}$ achieves C^{LD} in (3). $\diamond$

Remark 1. *Note that* $\mathbb{S}$ *consists of all the states where at most one relay is transmitting, while the rest of the relays are receiving. A similar condition can be obtained for the optimality of the states* $\mathbb{S}' = \{[n], [n]\setminus\{1\}, [n]\setminus\{2\}, \dots, [n]\setminus\{n\}\}$ *(where at most one relay is in receive mode).*

Remark 2. *For the case when the relays are non-interconnected, i.e.,* $\eta_{ij} = 0$ *for all* $(i,j) \in [n]^2$*, the result in Theorem 1 subsumes the result in [11]. Moreover, for* $n = 2$*, the result in Theorem 1 is equivalent to the one in [8], where we characterized the set of at most* 3 *states that suffice to achieve* C^{LD} *in* (3) *for* $n = 2$.

Remark 3. *The conditions in Theorem 1 are a consequence of the relaying scheme 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 that each relay decodes is equal to the amount of unique linearly independent bits that each relay transmits. The conditions in Theorem 1 ensure the feasibility of this scheme.*

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

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

$$f_{\mathcal{S}}^{\Omega} \geq \max_{i \in \Omega^c \cap \mathcal{S}^c} \eta_{is} + \max_{j \in \Omega \cap \mathcal{S}} \eta_{dj}, \tag{6}$$

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

Proof. Let

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

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

$$\left[\begin{array}{c|c} \mathbf{D}^{\eta-\eta_{ds}} & \mathbf{D}^{\eta-\eta_{dj^\star}} \\ \hline \mathbf{D}^{\eta-\eta_{i^\star s}} & \mathbf{D}^{\eta-\eta_{i^\star j^\star}} \end{array}\right] = \left[\begin{array}{c|c} \mathbf{0}_{\eta\times\eta} & \mathbf{D}^{\eta-\eta_{dj^\star}} \\ \hline \mathbf{D}^{\eta-\eta_{i^\star s}} & \mathbf{D}^{\eta-\eta_{i^\star j^\star}} \end{array}\right],$$

the rank of which is $\eta_{i^\star s} + \eta_{dj^\star}$. 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 $\{\eta_{dj} : j \in \Omega \cap \mathcal{S}\}$ and hence, $\mathrm{rank}(\mathbf{F}_{\mathcal{S}}^{\Omega}) = \max_{j \in \Omega\cap\mathcal{S}} \eta_{dj}$. Similarly, if $\Omega \cap \mathcal{S} = \varnothing$, then $\mathbf{F}_{\mathcal{S}}^{\Omega}$ has only one column and hence, $\mathrm{rank}(\mathbf{F}_{\mathcal{S}}^{\Omega}) = \max_{i \in \Omega^c\cap\mathcal{S}^c} \eta_{is}$. This concludes the proof of Proposition 1. $\square$

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

$$f_{\mathcal{S}}^{\Omega_1} + f_{\mathcal{S}}^{\Omega_2} \geq f_{\mathcal{S}}^{\Omega_1\cap\Omega_2} + f_{\mathcal{S}}^{\Omega_1\cup\Omega_2}, \tag{7}$$

for any subsets $\Omega_1, \Omega_2 \subseteq [n]$. Similarly, for a given cut $\Omega \subseteq [n]$, $f^{\Omega}_{\mathcal{S}}$ is submodular in $\mathcal{S}$, that is, $f^{\Omega}_{\mathcal{S}_1} + f^{\Omega}_{\mathcal{S}_2} \geq f^{\Omega}_{\mathcal{S}_1\cup\mathcal{S}_2} + f^{\Omega}_{\mathcal{S}_1\cap\mathcal{S}_2}$ for any subsets $\mathcal{S}_1, \mathcal{S}_2 \subseteq [n]$.

Proof. The proof is given in [1]. □

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 C^{U} provides an upper bound on the optimum value of (3), i.e., $\mathsf{C}^{\mathrm{U}} \geq \mathsf{C}^{\mathrm{LD}}$. 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 $\mathsf{C}^{\mathrm{U}} = t^\star$. Finally, in Section IV-C we show that the proposed solution is feasible for the original problem in (3), implying that $\mathsf{C}^{\mathrm{LD}} \geq t^\star$. Therefore, putting the three results together, we get $t^\star = \mathsf{C}^{\mathrm{U}} \geq \mathsf{C}^{\mathrm{LD}} \geq t^\star$. This shows that $\mathsf{C}^{\mathrm{LD}} = t^\star$ that can be attained by $\boldsymbol{\lambda}^\star$ that satisfies the claim of Theorem 1, which 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

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

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., $\mathsf{C}^{\mathrm{U}} \geq \mathsf{C}^{\mathrm{LD}}$.

B. An Optimal Solution for the Optimization Problem in (8)

We show that under the conditions of Theorem 1, the states in $\mathbb{S}$ are optimal for the upper bound C^{U} on the approximate capacity obtained by solving (8), that is, there exists an optimal solution with $(t, \lambda_{\mathbb{S}}) \geq 0$ and $\lambda_{\mathcal{S}} = 0$, for all $\mathcal{S} \notin \mathbb{S}$. In particular, we leverage the Karush–Kuhn–Tucker (KKT) conditions to prove the following proposition.

Proposition 3. *Let $\det(\mathbf{P}) \neq 0$ and $\frac{(-1)^{n+1}\mathsf{P}_{(n+1)}}{\det(\mathbf{P})} \geq 0$. Then, $(\boldsymbol{\lambda}^\star, t^\star)$ given by*

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

and $\lambda^\star_{\varnothing} = \lambda^\star_{\{n+1\}}$ are such that:

(i) $\lambda^\star_{\{i\}} \geq 0$, for every $i \in [n+1]$;
(ii) and $(\boldsymbol{\lambda}^\star, t^\star)$ is an optimal solution for the optimization problem in (8).

Consequently, we have $\mathsf{C}^U = t^\star$.

We present the proof of claim *(i)* in Appendix A. The proof of claim *(ii)* is presented below.

Proof of claim (ii) of Proposition 3. The proof leverages the KKT conditions. More precisely, for KKT multipliers $\boldsymbol{\mu} = (\mu_p, \mu_1, \ldots, \mu_{n+1})$ and $\boldsymbol{\sigma} = (\sigma_{\mathcal{S}} : \mathcal{S} \subseteq [n])$, the Lagrangian for the optimization problem in (8) is given by

$$\begin{aligned} \mathcal{L}(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\lambda}, t) = &- t + \sum_{i\in[n+1]} \mu_i (t - g_{[i:n]}) \\ &+ \mu_p (g_p - 1) - \sum_{\mathcal{S}\subseteq[n]} \sigma_{\mathcal{S}} \lambda_{\mathcal{S}}. \end{aligned} \tag{10}$$

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

$$\begin{bmatrix} \mu_p & \mu_1 & \ldots & \mu_n & \mu_{n+1} \end{bmatrix} \mathbf{P} = \begin{bmatrix} 1 & 0 & \ldots & 0 & 0 \end{bmatrix}, \tag{11}$$

and

$$\sigma_{\mathcal{S}} = \mu_p - \sum_{i=1}^{n+1} \mu_i f^{[i:n]}_{\mathcal{S}}, \tag{12}$$

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 (11) and (12) together with $(\boldsymbol{\lambda}^\star, t^\star)$ satisfy the following four groups of conditions.

• *Primal Feasibility.* We first show that, if $\det(\mathbf{P}) \neq 0$ and $\frac{(-1)^{n+1}\mathsf{P}_{(n+1)}}{\det(\mathbf{P})} \geq 0$, then the solution in (9) is feasible for the optimization problem in (8). Note that by setting $\lambda_{\mathcal{S}} = 0$, for all $\mathcal{S} \subseteq [n]$ with $\mathcal{S} \neq \mathbb{S}$ in (8), we obtain an optimization problem over $(n+2)$ variables (including t and $\lambda_{\mathcal{S}}$ for $\mathcal{S} \in \mathbb{S}$). We also force constraints $t \leq g_{[i:n]}$ for $i \in [n+1]$ and $g_p \leq 1$ to hold with equality. This leads to a system of $(n+2)$ linear equations in $(n+2)$ variables, given by

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

where $\mathbf{P}$ is the matrix defined in (4). The solution of this system of linear equations is given in (9). Therefore, the solution $(\boldsymbol{\lambda}^\star, t^\star)$ automatically satisfies the constraints of (8) corresponding to $g_{[i:n]}$ for all $i \in [n+1]$ and to g_p. Furthermore, the feasibility of the remaining constraints, namely $\lambda^\star_{\mathcal{S}} \geq 0$ for $\mathcal{S} \in \mathbb{S}$, is established by claim *(i) of Proposition 3.*

• Complementary Slackness. *We need to show that $(\boldsymbol{\mu}, \boldsymbol{\sigma})$ and the optimum solution $(\boldsymbol{\lambda}^\star, t^\star)$ given in (9) satisfy*

- *$\mu_i (t^\star - g^\star_{[i:n]}) = 0$ for all $i \in [n+1]$;*
- *$\mu_p (g^\star_p - 1) = 0$;*
- *and $\sigma_{\mathcal{S}} \lambda^\star_{\mathcal{S}} = 0$ for every $\mathcal{S} \subseteq [n]$.*

The first and the second conditions are readily implied by (9) for any choice of $\boldsymbol{\mu}$. Moreover, the third condition holds for $\mathcal{S} \notin \mathbb{S}$, since we have $\lambda^\star_{\mathcal{S}} = 0$ 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* (12) *and the j-th column of the matrix identity in* (11) *imply that*

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

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

• Stationarity. *We aim to prove that, when evaluated in $\boldsymbol{\mu}$ as in* (11) *and $\sigma_{\mathcal{S}}$ in* (12)*, the derivatives of the Lagrangian in* (10) *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

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

in which we used the definition of $\sigma_{\mathcal{S}}$ in (12).

• Dual Feasibility. *In this last part, we need to prove that the KKT multipliers in* (11) *and* (12) *are non-negative.*

Towards proving this claim for the μ variables, we first use induction to show that sign(μ_i) = sign(μ_j) for all $i, j \in [n+1]$. Then, we argue that since $\sum_{i=1}^{n+1} \mu_i = 1$ (from the first column of the matrix identity in (11)*), we have $\mu_1, \mu_2, \ldots, \mu_{n+1} \geq 0$. Lastly, for each $j \in [n+1]$, the j-th column of the same identity implies that $\mu_p = \sum_{i=1}^{n+1} \mu_i f_{\{j\}}^{[i:n]}$, from which it directly follows that $\mu_p \geq 0$. We refer to Appendix B for a more detailed proof of this argument.*

Next, we focus on the KKT multipliers $\sigma_{\mathcal{S}}, \mathcal{S} \subseteq [n]$ in (12)*. For an arbitrary state $\mathcal{S} = \{a_1, a_2, \ldots, a_k\}$, we can write*

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

where the inequality is due to Proposition 2, i.e., $f_{\mathcal{S}}^{\Omega}$ is submodular in $\mathcal{S}$. Thus, we get $\sigma_{\mathcal{S}} = \mu_p - \sum_{i=1}^{n+1} \mu_i f_{\mathcal{S}}^{[i:n]} \geq 0$. This concludes the proof of claim (ii) of Proposition 3. □

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

In Section IV-B, we have shown that the solution $(\boldsymbol{\lambda}^\star, t^\star)$ given in (9) is optimal for the optimization problem in (8). This implies that $t^\star = \mathsf{C}^{\mathrm{U}} \geq \mathsf{C}^{\mathrm{LD}}$ 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) is a feasible solution for the optimization problem in (3), which in turn implies $\mathsf{C}^{\mathrm{LD}} \geq t^\star$, 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, $\mathsf{C}^{LD} \geq t^\star$.*

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 $t^\star \leq g_{\Omega}^\star = \sum_{\mathcal{S} \subseteq [n]} \lambda_{\mathcal{S}}^\star f_{\mathcal{S}}^{\Omega}$, which exclusively appear in (3).

Towards this end, we consider an arbitrary cut $\Omega = [a_1 : b_1] \cup [a_2 : b_2] \cup \ldots \cup [a_k : b_k] \subseteq [n]$, where $a_1 \leq b_1 \leq a_2 \leq b_2 \leq \ldots \leq a_k \leq b_k$. We also define $a_{k+1} = n+1$. Recall from Proposition 2 that, for a given state $\mathcal{S}$, the function $f_{\mathcal{S}}^{\Omega}$ 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,

$$f_{\mathcal{S}}^{\Omega \cup [a_{i+1}:n]} + f_{\mathcal{S}}^{[b_i+1:n]} \geq f_{\mathcal{S}}^{\Omega \cup [a_i:n]} + f_{\mathcal{S}}^{[a_{i+1}:n]}, \tag{13}$$

for every $i \in [k]$. Moreover, note that for $a_{k+1} = n+1$, we have $[a_{k+1} : n] = [n+1 : n] = \varnothing$, and $\Omega \subseteq [a_1 : n]$. Therefore, we can write

$$\begin{aligned}&\sum_{\mathcal{S}\subseteq[n]} \lambda_{\mathcal{S}}^\star f_{\mathcal{S}}^{\Omega} \\ &= \sum_{\mathcal{S}\subseteq[n]} \lambda_{\mathcal{S}}^\star \left[\sum_{i=1}^{k}\left(f_{\mathcal{S}}^{\Omega\cup[a_{i+1}:n]} - f_{\mathcal{S}}^{\Omega\cup[a_i:n]}\right) + f_{\mathcal{S}}^{\Omega\cup[a_1:n]}\right] \\ &\geq \sum_{\mathcal{S}\subseteq[n]} \lambda_{\mathcal{S}}^\star \left[\sum_{i=1}^{k}\left(f_{\mathcal{S}}^{[a_{i+1}:n]} - f_{\mathcal{S}}^{[b_i+1:n]}\right) + f_{\mathcal{S}}^{[a_1:n]}\right] \\ &= \sum_{i=1}^{k+1}\sum_{\mathcal{S}\subseteq[n]} \lambda_{\mathcal{S}}^\star f_{\mathcal{S}}^{[a_i:n]} - \sum_{i=1}^{k}\sum_{\mathcal{S}\subseteq[n]} \lambda_{\mathcal{S}}^\star f_{\mathcal{S}}^{[b_i+1:n]} \\ &= \sum_{i=1}^{k+1} g_{[a_i:n]}^\star - \sum_{i=1}^{k} g_{[b_i+1:n]}^\star \\ &= (k+1)t^\star - kt^\star = t^\star,\end{aligned}$$

where the inequality follows from (13). This implies that $t^\star \leq \sum_{\mathcal{S}\subseteq[n]} \lambda_{\mathcal{S}}^\star f_{\mathcal{S}}^{\Omega}$ for any $\Omega \subseteq [n]$ and hence, concludes the claim of Proposition 4. This also completes the proof of Theorem 1. □

Appendix A
Proof of claim (i) in Proposition 3

We start by noting that, under the conditions in Theorem 1, namely $\det(\mathbf{P}) \neq 0$ and $\frac{(-1)^{n+1}\mathsf{P}_{(n+1)}}{\det(\mathbf{P})} \geq 0$, we readily have that $\lambda^\star_\varnothing$ in Proposition 3 is larger than or equal to zero, i.e.,

$$\lambda^\star_\varnothing = \lambda^\star_{\{n+1\}} = (-1)^{n+1}\frac{\mathsf{P}_{(n+1)}}{\det(\mathbf{P})} \geq 0.$$

Thus, in what follows we focus on showing that $\lambda^\star_{\{i\}} \geq 0$ for all $i \in [n]$. Towards this end, we first highlight that $(\boldsymbol{\lambda}^\star, t^\star)$ in Proposition 3 are the solutions of a system of linear equations constructed as follows: (i) setting $\lambda_{\mathcal{S}} = 0$, for all $\mathcal{S} \subseteq [n]$ with $\mathcal{S} \neq \mathbb{S}$ 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

$$\mathbf{P}\begin{bmatrix} t & \lambda_{\{1\}} & \dots & \lambda_{\{n\}} & \lambda_\varnothing\end{bmatrix}^T = \begin{bmatrix}1 & 0 & \dots & 0 & 0\end{bmatrix}^T, \quad (14)$$

and hence, the equation corresponding to the row $i+1$ for $i \in [0:n]$ of (14) is given by

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

where the equality in (a) follows from: (i) letting $\eta_{js} = 0$ for $j \notin [n]$; (ii) the fact that $\eta_{1s} \leq \eta_{2s} \leq \dots \leq \eta_{ns}$; and (iii) using the property in Proposition 1.

Using the above expression, we can now get the following recursive relation for $\lambda^\star_{\{i\}}$ for $i \in [n]$,

$$\begin{aligned}
&\lambda^\star_{\{i\}}\left(\eta_{is} - \eta_{(i-1)s}\right) \\
&= (\eta_{is} - t^\star) + \sum_{j=i+1}^{n} \lambda^\star_{\{j\}}\left(f^{[i+1:n]}_{\{j\}} - \eta_{is}\right),
\end{aligned} \quad (15)$$

where $\eta_{js} = 0$ for $j \notin [n]$.

We now proceed to the proof that the $\lambda^\star_{\{i\}}$'s for all $i \in [n]$ obtained from (15) are such that $\lambda^\star_{\{i\}} \geq 0$. In particular, the proof is based on contradiction and leverages the property of $f^\Omega_{\mathcal{S}}$ given in Lemma 1 (proof in Appendix C) and the property of $\mathbf{P}$ given in Lemma 2 (proof in Appendix D).

Lemma 1. *If $\eta_{1s} \leq \eta_{2s} \leq \dots \leq \eta_{(n-1)s} \leq \eta_{ns}$, then $f^{[a:n]}_{\{j\}} - \eta_{(a-1)s} \geq f^{[b:n]}_{\{j\}} - \eta_{(b-1)s}$ for all $n \geq j \geq a \geq b \geq 1$.*

Lemma 2. *If $\eta_{1s} \leq \eta_{2s} \leq \dots \leq \eta_{ns}$ and there exists an $i \in [n]$ such that $\eta_{is} = \eta_{(i-1)s} = f^{[i:n]}_{\{i\}}$, then $\det(\mathbf{P}) = 0$.*

Let $\mathcal{K} = \{i : \lambda^\star_{\{i\}} < 0, i \in [n]\}$, and k be the maximum element of $\mathcal{K}$. We also let $\mathcal{L} = \{j \in [1:k-1] : \eta_{js} = \eta_{(j-1)s}\}$, and ℓ be the maximum element of $\mathcal{L}$, with $\ell = 0$ if $\mathcal{L}$ is empty. Recall that $\eta_{js} = 0$ for $j \notin [n]$.

We now show that $\mathcal{K}$ cannot be non-empty (hence proving that $\lambda^\star_{\{i\}} \geq 0$ for all $i \in [n]$), using two main steps: (1) using the recursive relation in (15), we show that $\lambda^\star_{\{j\}} \leq 0$ for all $j \in [\ell+1:k-1]$; and (2) using the facts that $\lambda^\star_{\{j\}} \leq 0$ for all $j \in [\ell+1:k-1]$ and $\lambda_k < 0$ by assumption, we show that $(\boldsymbol{\lambda}^\star, t^\star)$ does not satisfy the equation corresponding to row $\ell+1$ of (14), thereby giving us the required contradiction.

• **Step 1:** $\lambda^\star_{\{j\}} \leq 0$ *for all* $j \in [\ell+1:k-1]$.
Since ℓ is the maximum element of $\mathcal{L}$, $\eta_{is} > \eta_{(i-1)s}$ for all $i \in [\ell+1:k-1]$. Therefore, using (15) we have that

$$\text{sign}(\lambda^\star_{\{i\}}) = \text{sign}\left((\eta_{is} - t^\star) + \sum_{j=i+1}^{n} \lambda^\star_{\{j\}}\left(f^{[i+1:n]}_{\{j\}} - \eta_{is}\right)\right).$$

Using the above relation, we show by induction that $\lambda^\star_{\{i\}} \leq 0$, for every $i \in [\ell+1:k-1]$.

⇒ *Recursive Case.* Assume that $\lambda^\star_{\{i\}} \leq 0$ for $i \in [k-m+1:k-1]$, for some $m \in [1:k-\ell-1]$. Using the above equation of sign$(\lambda^\star_{\{i\}})$ for $i = k-m$, we obtain

$$\begin{aligned}
&\left(\eta_{(k-m)s} - t^\star\right) + \sum_{j=k-m+1}^{n} \lambda^\star_{\{j\}}\left(f^{[k-m+1:n]}_{\{j\}} - \eta_{(k-m)s}\right) \\
&\overset{(a)}{\leq} \left(\eta_{(k-m)s} - t^\star\right) + \sum_{j=k+1}^{n} \lambda^\star_{\{j\}}\left(f^{[k-m+1:n]}_{\{j\}} - \eta_{(k-m)s}\right) \\
&\overset{(b)}{\leq} \left(\eta_{(k-m)s} - t^\star\right) + \sum_{j=k+1}^{n} \lambda^\star_{\{j\}}\left(f^{[k+1:n]}_{\{j\}} - \eta_{ks}\right) \\
&= \left(\eta_{(k-m)s} - \eta_{ks}\right) \\
&\quad + \left(\left(\eta_{ks} - t^\star\right) + \sum_{j=k+1}^{n} \lambda^\star_{\{j\}}\left(f^{[k+1:n]}_{\{j\}} - \eta_{ks}\right)\right) \\
&\overset{(c)}{=} \left(\eta_{(k-m)s} - \eta_{ks}\right) + \lambda^\star_{\{k\}}\left(\eta_{ks} - \eta_{(k-1)s}\right) \overset{(d)}{\leq} 0,
\end{aligned} \quad (16)$$

where the labeled inequalities/equalities follow from: (a) the assumptions $\lambda^\star_{\{j\}} \leq 0$ for $j \in [k-m+1:k-1]$ and $\lambda_{\{k\}} < 0$ and the fact that $\left(f^{[k-m+1:n]}_{\{j\}} - \eta_{(k-m)s}\right) \geq 0$ for $j \in [k-m+1:n]$ by Proposition 1; (b) the fact that $\lambda^\star_{\{j\}} \geq 0$ for $j \in [k+1:n]$ (since k is the maximum element of $\mathcal{K}$) and using Lemma 1 from which $f^{[k+1:n]}_{\{j\}} - \eta_{ks} \geq f^{[k-m+1:n]}_{\{j\}} - \eta_{(k-m)s}$; (c) using (15) for $k = i$; and (d) the fact that $\eta_{1s} \leq \eta_{2s} \leq \dots \leq \eta_{ns}$ and the assumption $\lambda_{\{k\}} < 0$.

⇒ *Base Case.* The base case is $m = 1$ and it readily follows by noting that (16) holds for $m = 1$.

• **Step 2:** $(\boldsymbol{\lambda}^\star, t^\star)$ *does not satisfy the equation corresponding to row* $\ell+1$ *of* (14).
Since $\ell \in \mathcal{L}$, we have that $\eta_{\ell s} = \eta_{(\ell-1)s}$. Therefore,

equation (15) corresponding to $i = \ell$ is given by

$$
\begin{aligned}
0 &= (\eta_{\ell s} - t^\star) + \sum_{j=\ell+1}^{n} \lambda^\star_{\{j\}} \left(f^{[\ell+1:n]}_{\{j\}} - \eta_{\ell s} \right) \\
&\overset{\text{(a)}}{\leq} (\eta_{\ell s} - t^\star) + \sum_{j=k+1}^{n} \lambda^\star_{\{j\}} \left(f^{[\ell+1:n]}_{\{j\}} - \eta_{\ell s} \right) \\
&\overset{\text{(b)}}{\leq} (\eta_{\ell s} - t^\star) + \sum_{j=k+1}^{n} \lambda^\star_{\{j\}} \left(f^{[k+1:n]}_{\{j\}} - \eta_{k s} \right) \\
&= (\eta_{\ell s} - \eta_{ks}) + (\eta_{ks} - t^\star) + \sum_{j=k+1}^{n} \lambda^\star_{\{j\}} \left(f^{[k+1:n]}_{\{j\}} - \eta_{ks} \right) \\
&\overset{\text{(c)}}{=} (\eta_{\ell s} - \eta_{ks}) + \lambda^\star_{\{k\}} \left(\eta_{ks} - \eta_{(k-1)s} \right) \overset{\text{(d)}}{<} 0,
\end{aligned} \tag{17}
$$

where the labeled inequalities/equalities follow from: (a) the facts that $\lambda^\star_{\{j\}} \leq 0$ for $j \in [\ell+1 : k-1]$ and $\lambda_{\{k\}} < 0$, and the fact that $f^{[\ell+1:n]}_{\{j\}} - \eta_{\ell s} \geq 0$ for $j \in [\ell+1 : n]$ by Proposition 1; (b) the fact that $\lambda^\star_{\{j\}} \geq 0$ for $j \in [k+1 : n]$ (since k is the maximum element of $\mathcal{K}$) and using Lemma 1 from which $f^{[k+1:n]}_{\{j\}} - \eta_{ks} \geq f^{[\ell+1:n]}_{\{j\}} - \eta_{\ell s}$ for $j \in [k+1 : n]$; (c) using (15) for $k = i$; and (d) (strict inequality) the fact that η_{ks} cannot be equal to $\eta_{\ell s}$ without making the matrix $\mathbf{P}$ singular, hence violating the condition $\det(\mathbf{P}) \neq 0$ in Theorem 1. This is proved in the remaining part of this appendix.

By letting $\lambda^\star_{\{0\}} = 0$, the difference between the equations corresponding to $i = k$ and $i = k-1$ in (15) is given by equation (18), at the top of the next page. Since $\ell \in \mathcal{L}$, we know that $\eta_{\ell s} = \eta_{(\ell-1)s}$. Now, if $\eta_{ks} = \eta_{\ell s} = \eta_{(\ell-1)s}$, then it implies that $\eta_{ks} = \eta_{(k-1)s} = \eta_{(k-2)s}$; this follows since $\ell \in [1 : k-1]$ and $\eta_{1s} \leq \eta_{2s} \leq \cdots \leq \eta_{ns}$. Moreover, if $\eta_{ks} = \eta_{(k-1)s} = \eta_{(k-2)s}$, then equation (18) becomes

$$
\lambda^\star_{\{k\}} \left(f^{[k:n]}_{\{k\}} - \eta_{(k-1)s} \right) = \sum_{j=k+1}^{n} \lambda^\star_{\{j\}} \left(f^{[k+1:n]}_{\{j\}} - f^{[k:n]}_{\{j\}} \right). \tag{19}
$$

Since by assumption $\lambda^\star_{\{k\}} < 0$ and $f^{[k:n]}_{\{k\}} - \eta_{(k-1)s} \geq 0$ from Proposition 1, we have that the left-hand side of (19) is smaller than or equal to zero. Moreover, since $\lambda^\star_{\{j\}} \geq 0$ for all $j \in [k+1 : n]$ (because k is the maximum element of $\mathcal{K}$) and $f^{[k+1:n]}_{\{j\}} \geq f^{[k:n]}_{\{j\}}$ (from Lemma 1), then the right-hand side of (19) is greater than or equal to zero. Thus, we must have that both sides of (19) are equal to zero, which implies that $f^{[k:n]}_{\{k\}} = \eta_{(k-1)s} = \eta_{ks}$. However, from Lemma 2 we have that this makes $\mathbf{P}$ singular, which contradicts the assumption of Theorem 1 that $\det(\mathbf{P}) \neq 0$. Therefore, η_{ks} cannot be equal to $\eta_{\ell s}$ and (17) is strictly less than 0, completing the proof of claim (i) in Proposition 3.

Appendix B
Proof that the KKT multipliers in (11) are non-negative

We start by showing that $\text{sign}(\mu_i) = \text{sign}(\mu_j)$ for all $i, j \in [n+1]$. In particular, we show this by induction.

$\Rightarrow$ *Recursive Case.* Assume that $\text{sign}(\mu_i) = \text{sign}(\mu_j)$ for all $i, j \in [k-1]$, where $k \in [2 : n+1]$. Then, we show by induction that $\text{sign}(\mu_{k-1}) = \text{sign}(\mu_k)$. Towards this end, we consider column $k-1$ and column $n+1$ of the identity in (11), which provides us with the relation $\mu_p = \sum_{j=1}^{n+1} f^{[j:n]}_{\{k-1\}} \mu_j = \sum_{j=1}^{n+1} f^{[j:n]}_{\varnothing} \mu_j$. Then, we have

$$
\sum_{j=1}^{n+1} f^{[j:n]}_{\{k-1\}} \mu_j - \sum_{j=1}^{n+1} f^{[j:n]}_{\varnothing} \mu_j = 0,
$$

or similarly

$$
\begin{aligned}
&\sum_{j=1}^{k-1} \left(f^{[j:n]}_{\{k-1\}} - f^{[j:n]}_{\varnothing} \right) \mu_j + \left(f^{[k:n]}_{\{k-1\}} - f^{[k:n]}_{\varnothing} \right) \mu_k \\
&+ \sum_{j=k+1}^{n+1} \left(f^{[j:n]}_{\{k-1\}} - f^{[j:n]}_{\varnothing} \right) \mu_j = 0,
\end{aligned}
$$

and hence,

$$
\begin{aligned}
&\sum_{j=1}^{k-1} \left(f^{[j:n]}_{\{k-1\}} - \eta_{(j-1)s} \right) \mu_j \\
&- \left(\eta_{(k-1)s} - \eta_{(k-2)s} \right) \mu_k = 0,
\end{aligned} \tag{20}
$$

where $\eta_{0s} = 0$ and where the last implication follows from Proposition 1. Moreover, from (6) we have that $\left(f^{[j:n]}_{\{k-1\}} - \eta_{(j-1)s} \right) \geq 0$ for $j \in [k-1]$, and also $\left(\eta_{(k-1)s} - \eta_{(k-2)s} \right) \geq 0$. It therefore follows that μ_k has the same sign as $\mu_1, \mu_2, \ldots, \mu_{k-1}$.

$\Rightarrow$ *Base Case.* The base case is $k = 2$ and it readily follows by noting that (20) holds for $k = 2$.

Thus, by induction, we obtain that $\mu_1, \mu_2, \ldots, \mu_{n+1}$ all have the same sign. Since from the first column of the identity in (11) we have $\sum_{i=1}^{n+1} \mu_i = 1$, it follows that $\mu_1, \mu_2, \ldots, \mu_{n+1} \geq 0$. Moreover, since $\mu_p = \sum_{j=1}^{n+1} f^{[j:n]}_{\{i\}} \mu_j$ for all $i \in [n+1]$ it directly follows that $\mu_p \geq 0$. This concludes the proof.

Appendix C
Proof of Lemma 1

The matrix $\mathbf{F}^{[a:n]}_{\{j\}}, j \in [a : n]$ is given as follows

$$
\mathbf{F}^{[a:n]}_{\{j\}} = \left[\begin{array}{c|c}
\mathbf{0}_{\eta \times \eta} & \mathbf{D}^{\eta - \eta_{dj}} \\
\hline
\mathbf{D}^{\eta - \eta_{1s}} & \mathbf{D}^{\eta - \eta_{1j}} \\
\hline
\mathbf{D}^{\eta - \eta_{2s}} & \mathbf{D}^{\eta - \eta_{2j}} \\
\hline
\vdots & \vdots \\
\hline
\mathbf{D}^{\eta - \eta_{(a-1)s}} & \mathbf{D}^{\eta - \eta_{(a-1)j}}
\end{array} \right].
$$

Since $a \geq b$, the matrix $\mathbf{F}^{[b:n]}_{\{j\}}$ is given by the first b block rows of $\mathbf{F}^{[a:n]}_{\{j\}}$. Since $\eta_{1s} \leq \eta_{2s} \leq \cdots \leq \eta_{ns}$, from the definition of $\mathbf{D}^{\eta - m}$ in (2) and recalling that we index rows and columns of a matrix starting from zero, it follows that columns $[\eta_{(b-1)s} : \eta - 1]$ are zero in $\mathbf{F}^{[b:n]}_{\{j\}}$.

Now, consider the lower left block of $\mathbf{F}^{[a:n]}_{\{j\}}$, namely $\mathbf{D}^{\eta - \eta_{(a-1)s}}$. From (2), for every $j \in [\eta_{(b-1)s} : \eta_{(a-1)s} - 1]$, the row $\eta - \eta_{(a-1)s} + j$ of the matrix $\mathbf{D}^{\eta - \eta_{(a-1)s}}$ has a 1 in column j and zero elsewhere. Therefore, these rows form

$$
\begin{aligned}
&\sum_{j\in[k-2]\cup\{n+1\}} \lambda^\star_{\{j\}}\left(\eta_{ks}-\eta_{(k-1)s}\right)+\lambda^\star_{\{k\}}\left(\eta_{(k-1)s}-f^{[k:n]}_{\{k\}}\right)\\
&+\lambda^\star_{\{k-1\}}\left(\eta_{ks}-\eta_{(k-2)s}\right)+\sum_{j=k+1}^{n}\lambda^\star_{\{j\}}\left(f^{[k+1:n]}_{\{j\}}-f^{[k:n]}_{\{j\}}\right)=0,
\end{aligned}
\tag{18}
$$

additional linearly independent rows in the matrix $\mathbf{F}^{[a:n]}_{\{j\}}$ compared to $\mathbf{F}^{[b:n]}_{\{j\}}$. Since there are $\eta_{(a-1)s}-\eta_{(b-1)s}$ such rows, it follows that $f^{[a:n]}_{\{j\}} \geq f^{[b:n]}_{\{j\}}+\left(\eta_{(a-1)s}-\eta_{(b-1)s}\right)$. This concludes the proof of Lemma 1.

Appendix D
Proof of Lemma 2

We show that, if $\eta_{1s}\leq\eta_{2s}\leq\cdots\leq\eta_{ns}$ and there exists an $i\in[n]$ such that $\eta_{is}=\eta_{(i-1)s}=f^{[i:n]}_{\{i\}}$, then columns i 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, we start by noting that $\mathbf{P}_{j,i}=-f^{[j:n]}_{\{i\}}$ and $\mathbf{P}_{j,(n+1)}=-f^{[j:n]}_{\varnothing}$ where $j\in[n+1]$.

Now, if $j\in[i+1:n+1]$, then using Proposition 1 and the fact that $\eta_{is}=\eta_{(i-1)s}$, we have that $f^{[j:n]}_{\{i\}}$=$f^{[j:n]}_{\varnothing}$. Thus, for all $j\in[i+1:n+1]$ we have that $\mathbf{P}_{j,i}=\mathbf{P}_{j,(n+1)}$.

We now focus on the case $j\in[1:i]$ and show that $f^{[j:n]}_{\{i\}}$=$f^{[j:n]}_{\varnothing}$ for all $j\in[i]$. Recall that $f^{[i:n]}_{\{i\}}=\text{rank}\left(\mathbf{F}^{[i:n]}_{\{i\}}\right)$, where $\mathbf{F}^{[i:n]}_{\{i\}}$ is defined as

$$
\mathbf{F}^{[i:n]}_{\{i\}}=\left[\begin{array}{c|c}
\mathbf{0}_{\eta\times\eta} & \mathbf{D}^{\eta-\eta_{di}}\\ \hline
\mathbf{D}^{\eta-\eta_{1s}} & \mathbf{D}^{\eta-\eta_{1i}}\\ \hline
\mathbf{D}^{\eta-\eta_{2s}} & \mathbf{D}^{\eta-\eta_{2i}}\\ \hline
\vdots & \vdots\\ \hline
\mathbf{D}^{\eta-\eta_{(i-1)s}} & \mathbf{D}^{\eta-\eta_{(i-1)i}}
\end{array}\right],
$$

where $\mathbf{D}^{\eta-m}$ is defined in (2). From the above equation, it follows that

$$
\begin{aligned}
f^{[i:n]}_{\{i\}} &\geq \max_{x\in[i-2]\cup d}\left\{\text{rank}\left(\mathbf{D}^{\eta-\eta_{(i-1)s}}\right)+\text{rank}\left(\mathbf{D}^{\eta-\eta_{xi}}\right)\right\}\\
&\overset{(a)}{=}\max_{x\in[i-2]\cup d}\left\{\eta_{(i-1)s}+\eta_{xi}\right\}\\
&\overset{(b)}{=}\eta_{(i-1)s},
\end{aligned}
$$

where the labeled equalities follow from: (a) the fact that the rank of $\mathbf{D}^{\eta-m}$ is equal to m; and (b) the assumption that $\eta_{is}=\eta_{(i-1)s}=f^{[i:n]}_{\{i\}}$ which implies that $\eta_{xi}=0$ for all $x\in[i-2]\cup d$. Moreover, for $j\in[i]$ the matrix $\mathbf{F}^{[j:n]}_{\{i\}}$ consists of the first j block rows of $\mathbf{F}^{[i:n]}_{\{i\}}$, and its entire column block is zero since $\eta_{xi}=0$ for all $x\in[i-2]\cup d$. Thus,

$$
f^{[j:n]}_{\{i\}}=\max_{\ell\in[j-1]}\{\eta_{\ell s}\}=\eta_{(j-1)s}=f^{[j:n]}_{\varnothing},
$$

where the second equality follows from the assumption $\eta_{1s}\leq\eta_{2s}\leq\cdots\leq\eta_{ns}$. Thus, also for $j\in[i]$ we have that $\mathbf{P}_{j,i}=\mathbf{P}_{j,(n+1)}$. This concludes the proof of Lemma 2.

References

[1] 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.

[2] 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.

[3] 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.

[4] 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.

[5] 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.

[6] 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.

[7] 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.

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

[9] 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.

[10] 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.

[11] 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.