The Approximate Optimality of Simple Schedules for Half-Duplex Multi-Relay Networks

Martina Cardone[†], Daniela Tuninetti[*] and Raymond Knopp[†]
[†]Eurecom, Biot, 06410, France, Email: {cardone, knopp}@eurecom.fr
[*] University of Illinois at Chicago, Chicago, IL 60607, USA, Email: danielat@uic.edu

***Abstract*—In ISIT 2012 Brahma, Özgür and Fragouli conjectured that in a half-duplex diamond relay network (a Gaussian noise network without a direct source-destination link and with N non-interfering relays) an approximately optimal relay scheduling (achieving the cut-set upper bound to within a constant gap uniformly over all channel gains) exists with at most $N+1$ *active* states (only $N+1$ out of the 2^N possible relay listen-transmit configurations have a strictly positive probability). Such relay scheduling policies are said to be *simple*. In ITW 2013 we conjectured that simple relay policies are optimal for any half-duplex Gaussian multi-relay network, that is, simple schedules are not a consequence of the diamond network's sparse topology. In this paper we formally prove the conjecture beyond Gaussian networks. In particular, for any memoryless half-duplex N-relay network for which the cut-set bound is approximately optimal to within a constant gap under some conditions (satisfied for example by Gaussian networks), an optimal schedule exists with at most $N+1$ active states. The key step of our proof is to write the minimum of a submodular function by means of its Lovász extension and use the greedy algorithm for submodular polyhedra to highlight structural properties of the optimal solution. This, together with the saddle-point property of min-max problems and the existence of optimal basic feasible solutions in linear programs, proves the claim.**

I. Introduction

Adding relaying stations to today's cellular infrastructure promises to boost both coverage and network throughput. Although higher performances could be attained with Full-Duplex (FD) relays, due to practical restrictions, such as the inability to perfectly cancel the self-interference, currently employed relays operate in Half-Duplex (HD).

This paper studies a general memoryless multi-relay network, where the communication between a source and a destination is assisted by N relays operating in HD mode. The capacity of this network is not known in general. In [1] we showed that Noisy Network Coding (NNC) [2] achieves the cut-set upper bound [3] to within $1.96\,(N+2)$ bits per channel use for a general Gaussian noise multi-relay network, universally over all channel gains, thus improving on previously known constant gap results. In general, finding the capacity of a HD multi-relay network is a combinatorial problem since the cut-set upper bound is the minimum between 2^N bounds (one for each possible cut in the network), each of which is a linear combination of 2^N relay states (since each relay can either transmit or receive). Thus, as the number of relays increases, optimizing the cut-set bound becomes prohibitively complex. Identifying structural properties of the cut-set upper bound, or of a constant gap approximation of the cut-set upper bound, is therefore critical for efficient numerical evaluations and can have important practical consequences for the design of reduced complexity / simple relaying policies.

In [4], the authors analyzed the Gaussian HD diamond relay network, a multi-relay network without a direct source-destination link, with $N=2$ non-interfering relays and proved that at most $N+1=3$ states, out of the $2^N=4$ possible ones, suffice to characterize the capacity to within a constant gap. We say that these $N+1$ states are *active* and form an (approximately optimal) *simple* schedule. In [5], Brahma *et al* verified through extensive numerical evaluations that in Gaussian HD diamond networks with $N \leq 7$ relays an optimal (to within a constant gap) schedule has at most $N+1$ active states and conjectured this to be true for any N. In [6], Brahma *et al*'s conjecture was proved for Gaussian HD diamond networks with $N \leq 6$ relays; the proof is based on certain properties of submodularity and on linear programming duality; the proof technique does not appear to easily generalize to an arbitrary N. Our numerical experiments in [1] showed that Brahma *et al*'s conjecture on the existence of optimal simple schedules for diamond HD relay networks extends to any Gaussian HD multi-relay network (i.e., not necessarily with a diamond topology) with $N \leq 8$; we conjectured that the same holds for any N. Should our more general version of Brahma *et al*'s conjecture be true, then Gaussian HD multi-relay networks have optimal simple schedules irrespectively of their topology. In [1] we also discussed polynomial time algorithms to determine the optimal simple schedule and extensions beyond relay networks.

Related works on determining the optimal relay scheduling, but not focused on characterizing the minimum number of active states, are available in the literature. For example [7] studied an iterative algorithm to determine the optimal schedule when the relays use decode-and-forward. In [8] the authors proposed a 'grouping' technique to compute the relay schedule that maximizes the approximate capacity of certain Gaussian HD relay networks; because finding a good node grouping is computationally complex, the authors proposed a heuristic approach based on tree decomposition which results in polynomial time algorithms; as for diamond networks in [5], the low-complexity algorithm of [8] relies on the 'simplified' topology of certain networks. *As opposed to these works, we prove that a linear number of states is sufficient to determine an optimal schedule regardless of the network topology.* We also note that in [9], FD relay networks were studied and that "under the assumption of independent inputs and noises, the cut-set bound is submodular" [9, Theorem 1], a result that we shall use in the derivation of our main result.

The main result of this paper is a formal proof of Brahma *et al*'s conjecture beyond the Gaussian noise case. In particular, we prove that for any HD network with N relays for which the cut-set bound is approximately optimal to within a constant gap under some conditions (precisely stated in Theorem 1) the optimal relay policy is simple. The key idea is to use the Lovász extension and the greedy algorithm for submodular polyhedra to highlight structural properties of the minimum of a submodular function. Then, by using the saddle-point property of min-max problems and the existence of optimal basic feasible solutions for Linear Programs (LPs), an (approximately) optimal relay policy with the claimed number of active states can be shown. A polynomial time algorithm to find the optimal simple relay schedule is also discussed.

The rest of the paper is organized as follows. Section II describes the general memoryless HD multi-relay network. Section III summarizes some known results for submodular functions and LPs and then proves the main result. Finally, Section IV concludes the paper.

II. System Model

A memoryless relay network has one source (node 0), one destination (node $N+1$), and N relays (indexed from 1 to N). It consists of $N+1$ input alphabets $(\mathcal{X}_1, \cdots, \mathcal{X}_N, \mathcal{X}_{N+1})$ (here $\mathcal{X}_i$ is the input alphabet of node i except for the source / node 0 where, for notation convenience, we use $\mathcal{X}_{N+1}$ rather than $\mathcal{X}_0$), $N+1$ output alphabets $(\mathcal{Y}_1, \cdots, \mathcal{Y}_N, \mathcal{Y}_{N+1})$ (here $\mathcal{Y}_i$ is the output alphabet of node i), and a memoryless channel transition probability $\mathbb{P}_{Y_{[1:N+1]}|X_{[1:N+1]}}$. Codes, achievable rates and capacity are defined in the usual way (see for example [1]).

In this general memoryless framework, each relay can listen and transmit at the same time, i.e., it is a FD node. HD channels are a special case of the memoryless FD framework in the following sense [10]. With a slight abuse of notation compared to the previous paragraph, we let the channel input of the k-th relay, $k \in [1:N]$, be the pair (X_k, S_k), where $X_k \in \mathcal{X}_k$ as before and $S_k \in [0:1]$ is the *state* random variable that indicates whether the k-th relay is in receive-mode ($S_k = 0$) or in transmit-mode ($S_k = 1$). In the HD case the channel transition probability is specified as $\mathbb{P}_{Y_{[1:N+1]}|X_{[1:N+1]},S_{[1:N]}}$. In particular, when the k-th relay, $k \in [1:N]$, is listening ($S_k = 0$) the outputs are independent of X_k, while when the k-th relay is transmitting ($S_k = 1$) its output Y_k is independent of all other random variables.

The capacity C of the HD multi-relay network is not known in general, but can be upper bounded by the cut-set bound [1]

$$\mathsf{C} \le \max_{\mathbb{P}_{X_{[1:N+1]},S_{[1:N]}}} \min_{\mathcal{A}\subseteq[1:N]} I_{\mathcal{A}}^{(\text{rand})}, \quad \text{where} \tag{1}$$

$$I_{\mathcal{A}}^{(\text{rand})} := I\left(X_{N+1}, X_{\mathcal{A}^c}, S_{\mathcal{A}^c}; Y_{N+1}, Y_{\mathcal{A}}|X_{\mathcal{A}}, S_{\mathcal{A}}\right) \tag{2}$$

$$\le H(S_{\mathcal{A}^c}) + I_{\mathcal{A}}^{(\text{fix})}, \tag{3}$$

$$I_{\mathcal{A}}^{(\text{fix})} := I\left(X_{N+1}, X_{\mathcal{A}^c}; Y_{N+1}, Y_{\mathcal{A}}|X_{\mathcal{A}}, S_{[1:N]}\right) \tag{4}$$

$$= \sum_{s\in[0:1]^N} \lambda_s \ f_s(\mathcal{A}), \ \text{for } \lambda_s := \mathbb{P}[S_{[1:N]} = s] \text{ and} \tag{5}$$

$$f_s(\mathcal{A}) := I\left(X_{N+1}, X_{\mathcal{A}^c}; Y_{N+1}, Y_{\mathcal{A}}|X_{\mathcal{A}}, S_{[1:N]} = s\right). \tag{6}$$

In the following, we use interchangeably the notation $s \in [0:1]^N$ to index all possible binary vectors of length N, as well as, $s \in [0:2^N-1]$ to indicate the decimal representation of a binary vector of length N. $I_{\mathcal{A}}^{(\text{rand})}$ in (2) is the mutual information across the network cut $\mathcal{A} \subseteq [1:N]$ when a *random schedule* is employed, i.e., information is conveyed from the relays to the destination by switching between listen and transmit modes of operation at random times [10] (see the term $H(S_{\mathcal{A}^c}) \le |\mathcal{A}^c| \le N$ in (3), which implies that a fixed schedule is optimal to within N bits). $I_{\mathcal{A}}^{(\text{fix})}$ in (4) is the mutual information with a *fixed schedule*, i.e., the time instants at which a relay transitions between listen and transmit modes of operation are fixed and known to all nodes in the network [10] (see the term $S_{[1:N]}$ in the conditioning in (4)).

III. Main Result

We next consider networks for which the following holds: there exists a product input distribution

$$\mathbb{P}_{X_{[1:N+1]}|S_{[1:N]}} = \prod_{i\in[1:N+1]} \mathbb{P}_{X_i|S_{[1:N]}} \tag{7a}$$

for which we can evaluate the set function $I_{\mathcal{A}}^{(\text{fix})}$ in (4) for all $\mathcal{A} \subseteq [1:N]$ and bound the capacity as

$$\mathsf{C}' - \mathsf{G}_1 \le \mathsf{C} \le \mathsf{C}' + \mathsf{G}_2, : \ \mathsf{C}' := \max_{\mathbb{P}_{S_{[1:N]}}} \min_{\mathcal{A}\subseteq[1:N]} I_{\mathcal{A}}^{(\text{fix})}, \tag{7b}$$

with G_1 and G_2 being non-negative constants that may depend on N but not on the channel transition probability. In other words, we concentrate on networks for which using independent inputs and a fixed relay schedule in the cut-set bound provides both an upper bound, to within G_2 bits, and a lower bound, to within G_1 bits, on the (unknown) capacity C.

For example, for a general Gaussian multi-relay network with independent noises, fixed schedules and independent Gaussian inputs are optimal to within $\mathsf{G}_1 + \mathsf{G}_2 \le 1.96(N+2)$ bits universally over all channel gains [1]. In [1] we conjectured that the optimal schedule in this case would be a simple one, i.e., the optimal probability mass function $\mathbb{P}_{S_{[1:N]}}$ in (7b) is such that at most $N+1$ entries have a strictly positive probability. This paper proves that not only is the conjecture true for the Gaussian noise case, but it also holds in more generality. The main result of the paper is:

Theorem 1. *Under the assumptions in* (7) *and of*

$$\mathbb{P}_{Y_{[1:N+1]}|X_{[1:N+1]},S_{[1:N]}} = \prod_{i\in[1:N+1]} \mathbb{P}_{Y_i|X_{[1:N+1]},S_{[1:N]}}, \tag{7c}$$

i.e., "independent noises", and if the functions in (6) *are not a function of* $\left\{\lambda_s, s \in [0:1]^N\right\}$*, i.e., they can depend on the state* s *but not on the* λ_s*, then simple relay policies are optimal*

[1] For notation convenience, we refer to $\mathcal{A}$ as the set that contains the relays which are in the cut of the destination / node $N+1$.

in (7b), *i.e., the optimal probability mass function $\mathbb{P}_{S_{[1:N]}}$ has at most $N+1$ non-zero entries / active states.*

We first summarize some properties of submodular functions and LPs in Section III-A, we then prove Theorem 1 in Section III-B, we discuss the computational complexity of finding optimal simple schedules in Section III-C and conclude with an example of a network with $N=2$ relays in order to illustrate some of the steps in the proof in Section III-D.

A. Submodular Functions, LPs and Saddle-point Property

The following are standard results in submodular function optimization [11] and LPs [12].

Definition 1 (Submodular function, Lovász extension and greedy solution for submodular polyhedra). *A set-function $f : 2^N \to \mathbb{R}$ is submodular if and only if, for all subsets $\mathcal{A}_1, \mathcal{A}_2 \subseteq [1:N]$, we have $f(\mathcal{A}_1) + f(\mathcal{A}_2) \geq f(\mathcal{A}_1 \cup \mathcal{A}_2) + f(\mathcal{A}_1 \cap \mathcal{A}_2)$ [2]. Note that submodular functions are closed under non-negative linear combinations.*

For a submodular function f such that $f(\emptyset)=0$, the Lovász extension is a function defined as

$$\widehat{f}(\mathbf{w}) := \max_{\mathbf{x}\in P(f)} \mathbf{w}^T\mathbf{x}, \quad \forall \mathbf{w}\in\mathbb{R}_+^N, \tag{8}$$

where $P(f)$ is the submodular polyhedron defined as

$$P(f) := \left\{ \mathbf{x}\in\mathbb{R}^N : \sum_{i\in\mathcal{A}} x_i \leq f(\mathcal{A}),\ \forall \mathcal{A}\subseteq[1:N] \right\}.$$

The optimal $\mathbf{x}$ in (8) can be found by the greedy algorithm for submodular polyhedra and has components

$$x_{\pi_i} = f(\{\pi_1,\dots,\pi_i\}) - f(\{\pi_1,\dots,\pi_{i-1}\}), \forall i\in[1:N],$$

where π is a permutation of $[1:N]$ such that the weights $\mathbf{w}\in\mathbb{R}_+^N$ are ordered as $w_{\pi_1}\geq w_{\pi_2}\geq\ldots\geq w_{\pi_N}$. Note that the Lovász extension is a piecewise linear convex function.

Proposition 2 (Minimum of submodular functions). *Let f be a submodular function such that $f(\emptyset)=0$ and $\widehat{f}$ its Lovász extension. The minimum of the submodular function satisfies*

$$\min_{\mathcal{A}\subseteq[1:N]} f(\mathcal{A}) = \min_{\mathbf{w}\in[0:1]^N} \widehat{f}(\mathbf{w}) = \min_{\mathbf{w}\in[0,1]^N} \widehat{f}(\mathbf{w}),$$

i.e., $\widehat{f}(\mathbf{w})$ attains its minimum at a vertex of $[0,1]^N$.

Definition 2 (Basic feasible solution). *Consider the LP*

$$\begin{aligned} &\text{maximize} && \mathbf{c}^T\mathbf{x} \\ &\text{subject to} && \mathbf{A}\mathbf{x}\leq\mathbf{b} \quad \mathbf{x}\geq 0, \end{aligned}$$

where $\mathbf{x}\in\mathbb{R}^n$ is the vector of unknowns, $\mathbf{b}\in\mathbb{R}^m$ and $\mathbf{c}\in\mathbb{R}^n$ are vectors of known coefficients, and $\mathbf{A}\in\mathbb{R}^{m\times n}$ is a known matrix of coefficients. If $m<n$, a solution for the LP with at most m non-zero values is called a basic feasible solution.

Proposition 3 (Optimality of basic feasible solutions). *If a LP is feasible, then an optimal solution is at a vertex of the (non-empty and convex) feasible set $S = \{\mathbf{x}\in\mathbb{R}^n : \mathbf{A}\mathbf{x}\leq\mathbf{b}, \mathbf{x}\geq 0\}$. Moreover, if there is an optimal solution, then an optimal basic feasible solution exists as well.*

Proposition 4 (Saddle-point property). *Let $\phi(x,y)$ be a function of two vector variables $x\in\mathcal{X}$ and $y\in\mathcal{Y}$. By the minimax inequality we have*

$$d^\star := \max_{y\in\mathcal{Y}}\min_{x\in\mathcal{X}} \phi(x,y) \leq \min_{x\in\mathcal{X}}\max_{y\in\mathcal{Y}} \phi(x,y) := p^\star,$$

and equality holds, i.e., $p^\star = d^\star$, if the following holds: (i) $\mathcal{X}$ and $\mathcal{Y}$ are both convex and one of them is compact; (ii) $\phi(x,y)$ is convex in x and concave in y; (iii) $\phi(x,y)$ is continuous.

B. Proof of Theorem 1

The objective is to show that simple relay policies are optimal in (7b). The proof consists of the following steps:

1) We first show that the function $I_{\mathcal{A}}^{(\text{fix})}$ defined in (4) is submodular under the assumptions in (7).
2) By using Proposition 2, we show that the problem in (7b) can be recast into an equivalent max-min problem.
3) With Proposition 4 we show that the max-min problem is equivalent to solve a min-max problem. The min-max problem is then shown to be equivalent to solve $N!$ max-min problems, for each of which we obtain an optimal basic feasible solution by Proposition 3 with the claimed maximum number of non-zero entries.

STEP 1: We show that $I_{\mathcal{A}}^{(\text{fix})}$ in (4) is submodular. The result in [9, Theorem 1] showed that $f_s(\mathcal{A})$ in (6) is submodular for each relay state $s\in[0:1]^N$ under the assumption of independent inputs and independent noises (the same work provided an example of a diamond network with correlated inputs, and showed that in this case the cut-set bound is neither submodular nor supermodular). Since submodular functions are closed under non-negative linear combinations (see Definition 1), this implies that $I_{\mathcal{A}}^{(\text{fix})} = \sum_{s\in[0:1]^N}\lambda_s\ f_s(\mathcal{A})$ is submodular under the assumptions of Theorem 1.

STEP 2: Given that $I_{\mathcal{A}}^{(\text{fix})}$ in (4) is submodular, we would like to use Proposition 2 to 'replace' the minimization over the subsets of $[1:N]$ in (7b) with a minimization over the cube $[0:1]^N$. Since $I_{\emptyset}^{(\text{fix})} = I\left(X_{[1:N+1]}; Y_{N+1}|S_{[1:N]}\right)\geq 0$ in general, we define a new submodular function $g(\mathcal{A}) := I_{\mathcal{A}}^{(\text{fix})} - I_{\emptyset}^{(\text{fix})}$ and proceed as in (9) at the top of the next page to show that the problem in (7b) is equivalent to

$$\mathsf{C}' = \max_{\lambda_{\text{vect}}}\min_{\mathbf{w}\in[0,1]^N}\Big\{[1,\mathbf{w}^T]\ \mathbf{H}_{\pi,f}\lambda_{\text{vect}}\Big\}, \tag{14}$$

where λ_{vect} is the probability mass function of $S_{[1:N]}$ (in particular, $\lambda_{\text{vect}} := [\lambda_s]\in\mathbb{R}_+^{2^N\times 1}$ where $\lambda_s := \mathbb{P}[S_{[1:N]}=s]\in[0,1]$, for $s\in[0:1]^N$ such that $\sum_{s\in[0:1]^N}\lambda_s = 1$), $\mathbf{H}_{\pi,f}\in\mathbb{R}^{(N+1)\times 2^N}$ and $\mathbf{F}_\pi\in\mathbb{R}^{(N+1)\times 2^N}$ are defined in (10) at the top of the next page, $\mathbf{P}_\pi\in\mathbb{R}^{(N+1)\times(N+1)}$ is the permutation matrix that maps $[1,w_1,\dots,w_N]$ into $[1,w_{\pi_1},\dots,w_{\pi_N}]$, and $f_s(\mathcal{A})$ was defined in (6). We thus express our original optimization problem as the max-min problem in (14).

[2] A set-function f is supermodular if and only if $-f$ is submodular, and it is modular if it is both submodular and supermodular.

$$\min_{\mathcal{A}\subseteq[1:N]} I_{\mathcal{A}}^{(\text{fix})} = I_{\emptyset}^{(\text{fix})} + \min_{\mathcal{A}\subseteq[1:N]} g(\mathcal{A})$$
$$= I_{\emptyset}^{(\text{fix})} + \min_{\mathbf{w}\in[0,1]^N} \begin{bmatrix} w_{\pi_1} & w_{\pi_2} & \dots & w_{\pi_N}\end{bmatrix} \begin{bmatrix} g(\{\pi_1\}) - g(\emptyset) \\ \vdots \\ g(\{\pi_1,\dots,\pi_N\}) - g(\{\pi_1,\dots,\pi_{N-1}\}) \end{bmatrix}$$
$$= I_{\emptyset}^{(\text{fix})} + \min_{\mathbf{w}\in[0,1]^N} \begin{bmatrix} w_{\pi_1} & w_{\pi_2} & \dots & w_{\pi_N}\end{bmatrix} \begin{bmatrix} I_{\{\pi_1\}}^{(\text{fix})} - I_{\emptyset}^{(\text{fix})} \\ \vdots \\ I_{\{\pi_1,\dots,\pi_N\}}^{(\text{fix})} - I_{\{\pi_1,\dots,\pi_{N-1}\}}^{(\text{fix})} \end{bmatrix}$$
$$= \min_{\mathbf{w}\in[0,1]^N} \begin{bmatrix} 1 & w_{\pi_1} & w_{\pi_2} & \dots & w_{\pi_N}\end{bmatrix} \begin{bmatrix} I_{\emptyset}^{(\text{fix})} \\ I_{\{\pi_1\}}^{(\text{fix})} - I_{\emptyset}^{(\text{fix})} \\ \vdots \\ I_{\{\pi_1,\dots,\pi_N\}}^{(\text{fix})} - I_{\{\pi_1,\dots,\pi_{N-1}\}}^{(\text{fix})} \end{bmatrix} := \min_{\mathbf{w}\in[0,1]^N} \left\{[1,\mathbf{w}^T]\ \mathbf{H}_{\pi,f}\right\}; \quad (9)$$

$$\mathbf{H}_{\pi,f} := \mathbf{P}_\pi \underbrace{\begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ -1 & 1 & 0 & \dots & 0 \\ 0 & -1 & 1 & \dots & 0 \\ \vdots & & & & \\ 0 & 0 & \dots & -1 & 1 \end{bmatrix}}_{(N+1)\times(N+1)} \mathbf{F}_\pi \text{ where } \mathbf{F}_\pi := \begin{bmatrix} f_0(\emptyset) & \dots & f_{2^N-1}(\emptyset) \\ f_0(\{\pi_1\}) & \dots & f_{2^N-1}(\{\pi_1\}) \\ f_0(\{\pi_1,\pi_2\}) & \dots & f_{2^N-1}(\{\pi_1,\pi_2\}) \\ & \dots & \\ f_0(\{\pi_1,\dots,\pi_N\}) & \dots & f_{2^N-1}(\{\pi_1,\dots,\pi_N\}) \end{bmatrix}; \quad (10)$$

$$g(\mathcal{A}) = I_{\mathcal{A}}^{(\text{fix})} - I_{\emptyset}^{(\text{fix})}, \mathcal{A}\subseteq[1:2]: \quad \widehat{g}(w_1,w_2) = \begin{cases} w_1 g(\{1\}) + w_2\left[g(\{1,2\}) - g(\{1\})\right] & \text{if } w_1 \ge w_2 \\ w_2 g(\{2\}) + w_1\left[g(\{1,2\}) - g(\{2\})\right] & \text{if } w_2 \ge w_1 \end{cases}; \quad (11)$$

$$P_1 : \max_{\lambda_{\text{vect}}} \min_{0\le w_2\le w_1\le 1} \underbrace{\begin{bmatrix} 1 & w_1 & w_2 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1\end{bmatrix}}_{=\begin{bmatrix}1-w_1 & w_1-w_2 & w_2\end{bmatrix}} \underbrace{\begin{bmatrix} f_0(\emptyset) & f_1(\emptyset) & f_2(\emptyset) & f_3(\emptyset) \\ f_0(\{1\}) & f_1(\{1\}) & f_2(\{1\}) & f_3(\{1\}) \\ f_0(\{1,2\}) & f_1(\{1,2\}) & f_2(\{1,2\}) & f_3(\{1,2\}) \end{bmatrix}}_{\mathbf{F}_{\pi^I}} \begin{bmatrix}\lambda_0 \\ \lambda_1 \\ \lambda_2 \\ \lambda_3\end{bmatrix}; \quad (12)$$

$$\begin{array}{lll} P_2: & \text{maximize} & \tau \\ & \text{subject to} & \tau \le f_0(\emptyset)\lambda_0 + f_1(\emptyset)\lambda_1 + f_2(\emptyset)\lambda_2 + f_3(\emptyset)\lambda_3, \\ & & \tau \le f_0(\{1\})\lambda_0 + f_1(\{1\})\lambda_1 + f_2(\{1\})\lambda_2 + f_3(\{1\})\lambda_3, \\ & & \tau \le f_0(\{1,2\})\lambda_0 + f_1(\{1,2\})\lambda_1 + f_2(\{1,2\})\lambda_2 + f_3(\{1,2\})\lambda_3, \\ & & \lambda_0+\lambda_1+\lambda_2+\lambda_3 = 1, \quad \lambda_i \ge 0\ i\in[0:3] \end{array}; \quad (13)$$

STEP 3: In order to solve (14) we would like to reverse the order of min and max. We note that the function $\phi(\lambda_{\text{vect}}, \mathbf{w}) := [1, \mathbf{w}^T]\ \mathbf{H}_{\pi,f}\lambda_{\text{vect}}$ satisfies the properties in Proposition 4 (it is continuous, convex in $\mathbf{w}$ by the convexity of the Lovász extension and linear, thus concave, in λ_{vect}; moreover the optimization domain in both variables is compact). Thus, we now focus on the problem

$$\mathsf{C}' = \min_{\mathbf{w}\in[0,1]^N} \max_{\lambda_{\text{vect}}} \left\{[1,\mathbf{w}^T]\ \mathbf{H}_{\pi,f}\lambda_{\text{vect}}\right\}, \quad (15)$$

which can be equivalently rewritten as

$$\mathsf{C}' = \min_{\pi\in\mathcal{P}_N} \min_{\mathbf{w}_\pi\in[0:1]^N} \max_{\lambda_{\text{vect}}} \left\{[1,\mathbf{w}_\pi^T]\ \mathbf{H}_{\pi,f}\lambda_{\text{vect}}\right\} \quad (16)$$
$$= \min_{\pi\in\mathcal{P}_N} \max_{\lambda_{\text{vect}}} \min_{\mathbf{w}_\pi\in[0:1]^N} \left\{[1,\mathbf{w}_\pi^T]\ \mathbf{H}_{\pi,f}\lambda_{\text{vect}}\right\}, \quad (17)$$

where $\mathcal{P}_N$ is the set of all the $N!$ permutations of $[1:N]$. In (16), for each permutation $\pi\in\mathcal{P}_N$, we first find the optimal λ_{vect}, and then find the optimal $\mathbf{w}_\pi : w_{\pi_1} \ge w_{\pi_2} \ge \dots w_{\pi_N}$. This is equivalent to (17), where again by Proposition 4, for each permutation $\pi\in\mathcal{P}_N$, we first find the optimal $\mathbf{w}_\pi : w_{\pi_1} \ge w_{\pi_2} \ge \dots w_{\pi_N}$, and then find the optimal λ_{vect}.

Let's now consider the inner optimization in (17), that is,

$$P_1 : \max_{\lambda_{\text{vect}}} \min_{\mathbf{w}_\pi\in[0:1]^N} \left\{[1,\mathbf{w}_\pi^T]\ \mathbf{H}_{\pi,f}\lambda_{\text{vect}}\right\}. \quad (18)$$

From Proposition 2 we know that, for a given $\pi\in\mathcal{P}_N$, the optimal $\mathbf{w}_\pi$ is a vertex of the cube $[0:1]^N$. For a given $\pi\in\mathcal{P}_N$, there are $N+1$ vertices whose coordinates are ordered according to π. In (18), for each of the $N+1$ feasible vertices of $\mathbf{w}_\pi$, it is easy to see that the product $[1,\mathbf{w}_\pi^T]\ \mathbf{H}_{\pi,f}$ is equal to a row of the matrix $\mathbf{F}_\pi$. By considering all possible $N+1$

feasible vertices compatible with π we obtain all the $N+1$ rows of the matrix $\mathbf{F}_\pi$. Hence, P_1 is equivalent to

$$P_2: \begin{array}{ll} \text{maximize} & \tau \\ \text{subject to} & \mathbf{1}_{(N+1)}\tau \leq \mathbf{F}_\pi \lambda_{\text{vect}} \\ \text{and} & \mathbf{1}_{2^N}^T \lambda_{\text{vect}} = 1, \ \lambda_{\text{vect}} \geq 0. \end{array} \quad (19)$$

The LP P_2 [3] has $n = 2^N + 1$ optimization variables (2^N values for λ_{vect} and 1 value for τ), $m = N + 2$ constraints, and is feasible (consider for example the uniform distribution of λ_{vect} and $\tau = 0$). Therefore, by Proposition 3, P_2 has an optimal basic feasible solution with at most $m = N + 2$ non-zero values. Since $\tau > 0$ (otherwise the channel capacity would be zero), it means that λ_{vect} has at most $N + 1$ non-zero entries.

Since for each $\pi \in \mathcal{P}_N$ the optimal λ_{vect} in (17) has at most $N+1$ non-zero values, so also for the optimal permutation the corresponding optimal λ_{vect} has at most $N+1$ non-zero values. This shows that the optimal schedule in the original problem in (7b) is simple and concludes the proof of Theorem 1.

C. On the complexity of finding the optimal simple schedule

Our proof method seems to suggest that finding the optimal schedule requires the solution of $N!$ LPs. Since $\log(N!) = O(N \log(N/e))$, the computational complexity of this approach would be prohibitive for large N. One can envisage that by using an iterative method that alternates between the submodular function minimization over $\mathbf{w}$ (solvable in strongly polynomial time in N by the Schrijver's algorithm) and the LP maximization over λ_{vect} (by the ellipsoid method, the worst-case dual LP is solvable in polynomial time in N) a polynomial time algorithm that converges to the optimal solution by the saddle-point property could be designed.

D. Example

For $N = 2$, P_1 in (18) requires an optimization over $\mathbf{w} = [w_1, w_2] \in [0, 1]^2$. From Proposition 2, the optimal $\mathbf{w}$ is one of the vertices $(0, 0), (0, 1), (1, 0), (1, 1)$. We must consider $|\mathcal{P}_2| = 2! = 2$ possible permutations: π^I for which $w_1 \geq w_2$, and π^{II} for which $w_2 \geq w_1$. For $N = 2$ the Lovász extension of a submodular function g is given in (11) at the top of the previous page (see also eq.(9)), which results in the problem P_1 in (12) at the top of the previous page when considering $\mathbf{w}_{\pi^I}$ (a similar reasoning holds for $\mathbf{w}_{\pi^{II}}$ but it is not reported here for sake of space). The vertices compatible with π^I are $[w_1, w_2] \in \{(0, 0), (1, 0), (1, 1)\}$, which result in $[1 - w_1, w_1 - w_2, w_2] \in \{(0, 0, 0), (0, 1, 0), (0, 0, 1)\}$. This implies that P_2 in (13) at the top of the previous page is the minimum of three functions, each given by one of the rows of $\mathbf{F}_{\pi_I}$ multiplied by $\lambda_{\text{vect}} = [\lambda_0, \lambda_1, \lambda_2, \lambda_3]$. P_2 has hence 4 constraints (3 from the rows of $\mathbf{F}_{\pi_I}$ and 1 from λ_{vect}) and 5 unknowns (1 value for τ and 4 entries of λ_{vect}). Thus, by Proposition 3, P_2 has an optimal basic feasible solution with at most 4 non-zero values, of which one is τ and thus the other 3 belong to λ_{vect}. By [4] and our generalization in [1], we know that either λ_0 or λ_3 is zero, thus giving the desired optimal simple schedule.

[3]Note that P_2 is a LP if and only if each $f_s, s \in [0 : 1^N]$ in $\mathbf{F}_\pi$ does not depend on λ_s, i.e., under the assumption in Theorem 1.

As mentioned earlier, the result of this paper proves the conjecture for Gaussian SISO networks with N relays and arbitrary topology. Our framework immediately extends to Gaussian networks with MIMO relays and independent noises since also in this setting independent inputs are optimal in the cut-set upper bound to within a constant gap for all choices of the channel matrices (see [13, Section IV]).

IV. Conclusions

In this work we studied networks with N half-duplex relays. For such networks, the capacity achieving scheme must be optimized over the 2^N possible listen-transmit relay configurations. This paper formally proved that, if noises are independent and independent inputs are approximately optimal in the cut-set bound, then the approximately optimal schedule only uses at most $N + 1$ relay configurations.

Acknowledgment. Eurecom's research is partially supported by its industrial partners, by the EU Celtic+ Framework Program Project SHARING, and by a 2014 Qualcomm Innovation Fellowship. The work of Dr. D. Tuninetti was partially funded by NSF under award number 1218635. D. Tuninetti would like to acknowledge insightful discussions with Dr. Salim El Rouayheb on submodular functions and with Dr. Farzad Parvaresh on polynomial time algorithms.

References

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

[2] S. Lim, Y.-H. Kim, A. El Gamal, and S.-Y. Chung, "Noisy network coding," *IEEE Transactions on Information Theory*, vol. 57, no. 5, pp. 3132 –3152, May 2011.

[3] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge U.K.: Cambridge University Press, 2011.

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

[5] S. Brahma, A. Özgür, and C. Fragouli, "Simple schedules for half-duplex networks," in *IEEE International Symposium on Information Theory (ISIT), 2012*, July 2012, pp. 1112 –1116.

[6] S. Brahma and C. Fragouli, "Structure of optimal schedules in diamond networks," in *IEEE International Symposium on Information Theory (ISIT), 2014*, June 2014, pp. 641–645.

[7] 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, September 2012.

[8] R. H. Etkin, F. Parvaresh, I. Shomorony, and A. S. Avestimehr, "Computing half-duplex schedules in Gaussian relay networks via min-cut approximations," *IEEE Transactions on Information Theory*, vol. 60, no. 11, November 2014.

[9] F. Parvaresh and R. Etkin, "Efficient capacity computation and power optimization for relay networks," *IEEE Transactions on Information Theory*, vol. 60, no. 3, pp. 1782–1792, March 2014.

[10] G. Kramer, "Models and theory for relay channels with receive constraints," in *42nd Annual Allerton Conference on Communication, Control, and Computing*, September 2004, pp. 1312–1321.

[11] F. Bach, "Learning with sub modular functions: a convex optimization perspective," *Foundations and Trends®in Machine Learning*, vol. 6, no. 2-3, pp. 145–373, December 2013.

[12] V. Chvátal, *Linear Programming*. W.H.Freeman, 1983.

[13] M. Cardone, D. Tuninetti, and R. Knopp, "On the optimality of simple schedules for networks with multiple half-duplex relays," *submitted to IEEE Transactions on Information Theory, available at arXiv:1412.1993*, December 2014.