Gaussian Half-Duplex Relay Networks: Improved Gap and a Connection with the Assignment Problem

Martina Cardone[†], Daniela Tuninetti[*], Raymond Knopp[†], Umer Salim[‡]
[†]Eurecom, Biot, 06410, France, Email: {cardone, knopp}@eurecom.fr
[*] University of Illinois at Chicago, Chicago, IL 60607, USA, Email: danielat@uic.edu
[‡] Intel Mobile Communications, Sophia Antipolis, 06560, France, Email: umer.salim@intel.com

***Abstract*—This paper studies a Gaussian relay network, where the relays can either transmit or receive at any given time, but not both. Known upper (cut-set) and lower (noisy network coding) bounds on the capacity of a memoryless full-duplex relay network are specialized to the half-duplex case and shown to be to within a constant gap of one another. For fairly broad range of relay network sizes, the derived gap is smaller than what is known in the literature, and it can be further reduced for more structured networks such as diamond networks. It is shown that the asymptotically optimal duration of the listen and transmit phases for the relays can be obtained by solving a linear program; the coefficients of the linear constraints of this linear program are the solution of certain 'assignment problems' for which efficient numerical routines are available; this gives a general interesting connection between the high SNR approximation of the capacity of a MIMO channel and the 'assignment problem' in graph theory. Finally, some results available for diamond networks are extended to general networks. For a general relay network with 2 relays, it is proved that, out of the 4 possible listen/transmit states, at most 3 have a strictly positive probability. Numerical results for a network with $K-2 < 9$ relays show that at most $K-1$ states have a strictly positive probability, which is conjectured to be true for any number of relays.**

I. Introduction

Cooperation between nodes in a network has been proposed as a potential and promising technique to enhance the performance of wireless systems in terms of coverage, throughput, and diversity. The simplest form of collaboration can be modeled as a Relay Channel (RC) [1]. The RC is a multi-terminal network where a source conveys information to a destination with the help of one relay. A relay is said to work in Half-Duplex (HD) mode if at any time / frequency instant it can not simultaneously transmit and receive. The HD modeling assumption is, at present, more practical than the Full-Duplex (FD) one since practical restrictions, such as the inability to perfectly cancel the self-interference, make the implementation of FD relays challenging. Motivated by the undeniable practical importance of the RC, in this paper we analyze a system where the communication between a source and a destination is assisted by multiple HD relays.

Despite the large amount of work on the RC, its capacity is still not known in general. [1] developed two coding schemes, Decode-and-Forward (DF) and Compress-and-Forward (CF), a general upper bound on the capacity (cut-set) and showed capacity for certain classes of RCs. [2] derived upper and lower bounds on the capacity of the Gaussian HD RC; the former is based on the cut-set argument; the latter exploits Partial DF; the relay listen/transmit switch is considered as fixed (i.e., *deterministic switch*). [3] showed that larger achievable rates can be attained with a *random switch* at the relay; moreover it argued that HD constraints can be incorporated into the general memoryless FD framework by considering appropriate *state* random variables. Here we follow the approach of [3] and specialize known FD inner and outer bounds to the HD case.

For Gaussian HD relay networks with K nodes (a source, a destination and $K-2$ relays), [4] showed that quantize-map-forward (QMF) achieves the cut-set upper bound to within $5K$ bits but did not consider random switch in the upper bound. [5] showed that QMF can be implemented with lattice codes and it reduced the gap to $5(K-2)$ while also properly accounting for the random switch in the upper bound. In Theorem 1 we show that, for fairly broad range of relay network sizes (from networks with few relays to up to 239 relays), the gap of [5] can be reduced to $(K/2)\log(4K)$ by using Noisy Network Coding (NNC) [6]. Our derived gap is valid for a general fully connected HD relay network. In the FD literature, it is well known that the gap of $1.26K$ bits for a general network [6] can be reduced to $2\log(K-1)$ bits for a diamond network [7]. In a *diamond network* the network topology is restricted compared to a general network, i.e., the source can not communicate directly with the destination and the relays can not communicate among themselves. In Theorem 2 we show that similar conclusions hold for HD networks, namely, we derive a reduced gap for HD diamond networks.

In a HD network with $K-2$ relays, there are 2^{K-2} possible combinations of listen/transmit states for the relays. For the diamond network with $K-2=2$ relays, [8] showed that at most $K-1=3$ states, out of the $2^{K-2}=4$ possible ones, with strictly positive probability suffice to achieve capacity to within 4 bits; in this case we say that there are $K-1$ *active states*. Inspired by [8], [9] showed that for a special HD diamond network with 3 relays, at most 4 states are active. In [9], it was also numerically verified that for a general HD diamond network with $(K-2)\leq 7$ relays, at most $K-1$ states are active; it was conjectured that this result holds for any K. Here we show that these results do not depend on the particular topology of diamond networks. In particular, in Theorem 5 we analytically show that at most 3 states are active in a general HD network with 2 relays by using a proof technique different from [8]. Based on numerical evidences we conjecture that the conjecture of [9] holds for a general HD relay network, namely that in a network with $K-2$ HD relays, at most $K-1$ states are active, out of the 2^{K-2} possible states.

In [8], [9] it was shown that determining the optimal values of the state probabilities of a diamond network is equivalent to solving a Linear Program (LP). Not surprisingly, the same holds for a general HD network as stated in Theorem 3 where we show, by properly accounting for power allocation and random switch, that a LP gives the asymptotic optimal values of the state probabilities at high SNR, i.e., in a generalized Degrees of Freedom (gDoF) sense [4]. One of the difficulties in determining the gDoF lies in the absence of known numerically efficient ways to evaluate the coefficients of the constraints in the LP, which are log-det expressions similar to the capacity of a MIMO channel. In a HD network with $K-2$ relays, one has to evaluate $2^{2(K-2)}$ such log-det terms. In Theorem 4 we make a connection between the problem of finding the gDoF of a MIMO channel with the graph theoretic problem of determining the *maximum weighted bipartite matching* (MWBM) of a bigraph, also known as assignment problem. As a result, known polynomial time algorithms for the MWBM problem, such as the Hungarian algorithm, can be readily used to efficiently determine the gDoF of a general HD relay network. Citations from the graph theory literature on these problems are omitted for sake of space.

II. System Model

We use the following notation convention: $[n_1 : n_2]$ is the set of integers from n_1 to $n_2 \geq n_1$; $[x]^+ := \max\{0, x\}$ for $x \in \mathbb{R}$; Y^j is a vector of length j with components $(Y_1, \ldots, Y_j)$; for an index set $\mathcal{A}$ we let $Y_{\mathcal{A}} = \{Y_j : j \in \mathcal{A}\}$; $\mathbf{0}_j$ is the all zero row vector of length j; $\mathbf{1}_j$ is the all one column vector of length j; $\mathbf{I}_j$ is the identity matrix of dimension j; $f_1(x) \doteq f_2(x)$ means that $\lim_{x\to+\infty} f_1(x)/f_2(x) = 1$.

A relay network has one source (node 1) that sends a message to one destination (node K) with the help of $K-2$ causal HD relays (numbered from 2 to $K-1$) through a shared memoryless channel. We use here the standard definition of capacity, which we do not repeat for sake of space.

A single-antenna complex-valued power-constrained Gaussian HD relay network has input/output relationship

$$\mathbf{Y} = (\mathbf{I}_K - \mathbf{S})\mathbf{H}\mathbf{S}\mathbf{X} + \mathbf{Z}, \tag{1}$$

where: $\mathbf{Y} := [Y_1, \ldots, Y_K]^T \in \mathbb{C}^K$ is the vector of the received signals, $\mathbf{X} := [X_1, \ldots, X_K]^T \in \mathbb{C}^K$ is the vector of the transmitted signals, $\mathbf{S} := \mathrm{diag}[S_1, \ldots, S_K]$ is the diagonal matrix that indicates the state of the relays (here $S_1 = 1$ since the source always sends, $S_k \in \{0, 1\}$, $k \in [2 : K-1]$, since a relay can either receive or transmit, and $S_K = 0$ since the destination always receives), $\mathbf{H} = [h_{ij}] \in \mathbb{C}^{K\times K}$ is the constant channel matrix known to all terminals, where h_{ij} with $(i, j) \in [1 : K]^2$ represents the channel from node j to node i. Without loss of generality, we assume that the channel inputs are subject to the average power constraint $\mathbb{E}\left[|X_k|^2\right] \leq 1$, $k \in [1 : K-1]$, and that the noises at all nodes are jointly Gaussian with zero mean and unit variance. Furthermore we assume that the noises are independent.

The capacity C of the Gaussian HD relay network in (1) is not known. We say that C is known to within GAP bits if one can show an achievable rate $R^{(\mathrm{in})}$ and an outer bound $R^{(\mathrm{out})}$ such that $R^{(\mathrm{out})} \leq R^{(\mathrm{in})} + \mathsf{GAP}$, and where GAP is a constant that does not depend on the channel gain matrix $\mathbf{H}$ in (1). Knowing the capacity to within a constant gap implies the exact knowledge of the gDoF defined as

$$\mathsf{d} := \lim_{\mathsf{SNR}\to+\infty} \frac{\mathsf{C}}{\log(1+\mathsf{SNR})}, \tag{2}$$

where the channel gains are parameterized as $|h_{ij}|^2 = \mathsf{SNR}^{\beta_{ij}}$ for some fixed set $\{\beta_{ij} \geq 0 : (i, j) \in [1 : K]^2\}$.

III. Capacity to within a constant gap

This section characterizes the capacity of the Gaussian HD relay network in (1) to within a constant gap. To accomplish this, we adapt the cut-set upper bound [10] and the NNC lower bound [6] to the HD case by following the approach proposed in [3]. Our main result is:

Theorem 1 *The cut-set upper bound for a general Gaussian HD relay network with $K-2$ relays is achievable to within*

$$\mathsf{GAP} \leq \max_{\ell\in[0:K-2]} \{\min\{1+\ell, K-1-\ell\} \log(1+\ell) \tag{3}$$

$$+ \min\{1+3\ell, \ell+K-1\}\} \leq (K/2)\log(4K) \text{ bits}. \tag{4}$$

Proof: The proof can be found in Appendix A. ■

The linear part of the gap in (3) is partly due to the random switch at the relays, which can convey at most 1 bit of information per relay per channel use. In the cut-set bound, we upper bounded this amount of information by the number of relays. For large K, the optimal value of ℓ in (3) is $\ell \cong \frac{K-2}{2}$, which provides the looser gap in (4). For fairly broad range of relay network sizes (up to 239 relays), our NNC-based gap in (4) improves on the gap of $5(K-2)$ bits in [5].

The gap in Theorem 1 grows with K and for large K it could be too large to give an approximate capacity characterization. A smaller gap may be obtained by several means:

- Inspired by [6], one could envisage to use more sophisticated bounding techniques. In [11], we proved that a more involved strategy based on a water-filling power allocation reduces the gap to $2.021K$ bits, which is smaller than that in [5] for $K \geq 4$, i.e., for strictly more than one relay.
- By employing other achievable schemes. For example, Partial DF gives a smaller gap than NNC in a single relay case [12]. However, Partial DF seems not to extend easily to networks with an arbitrary number of relays [10], which is the main reason why we considered NNC here.
- By deriving tighter bounds on specific network topologies, such as the diamond network [7], for which we can show

Theorem 2 *The cut-set upper bound for the Gaussian HD diamond network with $K-2$ relays is achievable to within* $\mathsf{GAP} \leq (K-4)\log(2) + 4\log(K) + 2\log(\mathrm{e})$.

Proof: The proof directly follows from Appendix A, by taking into consideration the fact that in the diamond network the channel matrix $\mathbf{H}$ has rank 2 (and thus $\mathrm{Rank}[\mathbf{H}_{\mathcal{A},s}] = 2$) and by using the possibly suboptimal value $\sigma^2 = K/2 - 1$. ■

As expected, the gap in Theorem 2 for the HD diamond network is, in general, smaller than that in (4), which is in line with what happens in FD. However, for the FD diamond network the gap is logarithmic in K [7], while the gap in Theorem 2 is linear in K. This is a consequence of further upper bounding, in the cut-set upper bound, the entropy of $[1,\ S_{2:K-1},\ 0]$ by the number of relays, which is linear in K.

IV. Analysis of optimal schedules

With the result of Theorem 1, and by substituting the expression in (9) from Appendix A into (2), we can evaluate the gDoF of the Gaussian HD relay network. Similarly to [8], [9], the gDoF is the solution of the LP

$$\mathsf{d} = \max\{\mathbf{f}^T\ \mathbf{x}\}\ \ \text{s.t.} \begin{bmatrix} -\mathbf{A} & \mathbf{1}_{2^{K-2}} \\ \mathbf{1}_{2^{K-2}}^T & 0 \end{bmatrix} \mathbf{x} \leq \mathbf{f}, \quad \mathbf{x} \geq 0, \quad (5)$$

where $\mathbf{f}^T := [\mathbf{0}_{2^{K-2}}, 1]$, $\mathbf{x}^T := [\lambda_1, \ldots, \lambda_{2^{K-2}}, \mathsf{d}]$, and the square matrix $\mathbf{A}$, of dimension 2^{K-2}, contains the high-SNR pre-log factors of the different log-determinant terms from the expression in (9) in Appendix A.

We next describe how to evaluate the entries of $\mathbf{A}$ by establishing an interesting connection between this problem and the *assignment problem* in graph theory. The importance of this connection stems for the fact that the computation of the entries of $\mathbf{A}$ can be performed very efficiently by using known polynomial time algorithms for the assignment problem.

We start with few definitions. A weighted bipartite graph, or bigraph, is a graph whose vertices can be separated into two sets such that each edge in the graph has exactly one endpoint in each set. Moreover, a non-negative weight is associated with each edge in the bigraph. A matching, or independent edge set, is a set of edges without common vertices. The MWBM problem, or assignment problem, is defined as a matching where the sum of the edge weights in the matching has the maximal value. The Hungarian algorithm is a polynomial time algorithm that efficiently solves the assignment problem. Equipped with these definitions we can show

Theorem 3 *For a general HD relay network, the $(i,j) \in [1 : 2^{K-2}]^2$ entry of the matrix $\mathbf{A}$ in (5) evaluates to the MWBM of a bigraph with weight matrix*

$$\Big[(\mathbf{I}_K - \mathrm{diag}[1, s_j, 0])\ \mathbf{B}\ \mathrm{diag}[1, s_j, 0]\Big]_{\{K\}\cup\mathcal{A}_i^c, \{1\}\cup\mathcal{A}_i} \quad (6)$$

where $\mathbf{B}$ is the 'SNR-exponent matrix' defined as $[\mathbf{B}]_{\ell k} = \beta_{\ell k} \geq 0 : |h_{\ell k}|^2 = \mathsf{SNR}^{\beta_{\ell k}}, (\ell, k) \in [1 : K]^2$. The index set $\mathcal{A}_i$ and the vector s_j in (6) are determined from the pair of indices (i, j), which gives the position within the matrix $\mathbf{A}$, as explained in the footnote 1.

The proof of Theorem 3 is a straightforward application of the following general result that establishes the equality between the gDoF of a MIMO channel, i.e., its high-SNR pre-log factor, and the MWBM problem.

Theorem 4 *Let $\mathbf{H} \in \mathbb{R}^{k\times n}$ be a full-rank matrix, where without loss of generality $k \leq n$. Let $\mathcal{S}_{n,k}$ be the set of all k-combinations of the integers in $[1 : n]$ and $\mathcal{P}_{n,k}$ be the set of all k-permutations of the integers in $[1 : n]$. Then,*

$$|\mathbf{I}_k + \mathbf{H}\mathbf{H}^H| \doteq \mathsf{SNR}^{\mathit{MWBM}(\mathbf{B})} \quad (7)$$

$$\mathit{MWBM}(\mathbf{B}) := \max_{\varsigma\in\mathcal{S}_{n,k}} \max_{\pi\in\mathcal{P}_{n,k}} \sum_{i=1}^{k} [\mathbf{B}^{\varsigma}]_{i,\pi(i)} \quad (8)$$

where $\mathbf{B}$ is the 'SNR-exponent matrix' defined as $[\mathbf{B}]_{\ell k} = \beta_{\ell k} \geq 0 : |h_{\ell k}|^2 = \mathsf{SNR}^{\beta_{\ell k}}, (\ell, k) \in [1 : K]^2$, and $\mathbf{B}^{\varsigma}$ is a square matrix obtained from $\mathbf{B}$ by retaining all the rows and the columns indexed by ς.

Proof: We only provide a sketch of the proof (the details may be found in [11]). First we prove that $|\mathbf{I}_k + \mathbf{H}\mathbf{H}^H| \doteq |\mathbf{H}\mathbf{H}^H|$ at high SNR, then we apply the Cauchy-Binet formula, the determinant Leibniz formula and the Cauchy-Swartz inequality to $|\mathbf{H}\mathbf{H}^H|$. In Appendix B, we provide two examples on how (8) can be applied. ■

With Theorem 4 we can now compute the matrix $\mathbf{A}$ in (6) for a general HD relay network and so find the gDoF. A closed form solution for the LP in (5) is not available in general. Even for networks with restricted topology, like diamond networks, the analytical solution of (5) is elusive. Nonetheless, it is possible to make statements on the 'sparsity' of an optimal solution for the state probability vector $[\lambda_1, \ldots, \lambda_{2^{K-2}}]$.

For the case of the HD diamond network, it was noted in [9] that, out of the 2^{K-2} states, at most $K-1$ are active. An interesting question is whether this is a consequence of the special topology of a diamond network or whether it holds for any HD relay network. For $K-2=2$ relays we can show:

Theorem 5 *For a general HD relay network with 2 relays, there exists an optimal schedule, i.e., an optimal value of the 2^2 possible listen-receive configurations/states for the relays, that optimizes d in (5) with at most 3 active states.*

Proof: We only provide a sketch of the proof (the details can be found in [11]). The key idea – and main difference with respect to [8] – is to determine the channel parameters under which setting one of the λ's to zero is without loss of optimality. The result is the following generalization of [8]: if

$$[\beta_{21} - \beta_{41}]^+[\beta_{31} - \beta_{41}]^+ \geq [\beta_{42} - \beta_{41}]^+[\beta_{43} - \beta_{41}]^+,$$

[1] Index i refers to a "cut" in the network and index j to a "state of the relays". Both indices range in $[1 : 2^{K-2}]$ and must be seen as the decimal representation of a binary number with $K-2$ bits. $\mathcal{A}_i$, $i \in [1 : 2^{K-2}]$, is the set of those relays who have a one in the corresponding binary representation of i (e.g., for $K-2=3$ and $i=6=1\cdot 2^2+1\cdot 2^1+0\cdot 2^0$ we have $\mathcal{A}_6=\{2,3\}$ and therefore $\mathcal{A}_6^c=\{4\}$). s_j, such that $\mathbb{P}(S_{[2:K-1]}=s_j)=\lambda_j$, $j \in [1 : 2^{K-2}]$, sets the state of a relay to the corresponding bit in the binary representation of j (e.g., for $K-2=3$ and $j=6$ we have $s_6=[1,1,0]$, which means that nodes 2 and 3 are sending and node 4 is receiving).

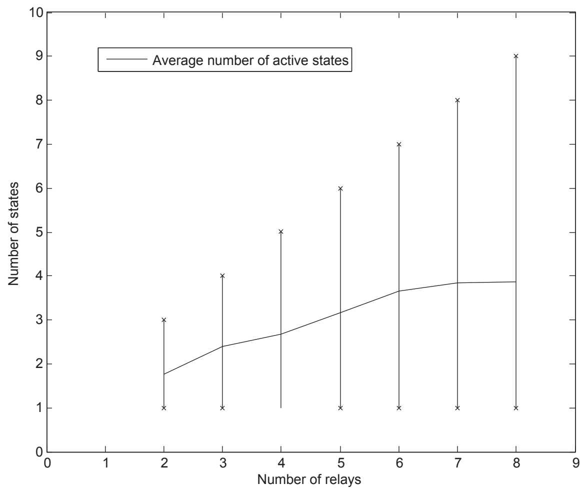

Fig. 1. Average, minimum and maximum number of active states to characterize the capacity of a HD relay network.

then $\lambda_1 = 0$ without loss of optimality, otherwise $\lambda_4 = 0$. ■

We conjecture that, for a general HD relay network with any number of relays, Theorem 5 continues to hold, similarly to the conjecture in [9] for the diamond network. Namely:

Conjecture: *For a general HD relay network with $K-2$ relays, there always exists an optimal schedule that maximizes the gDoF with at most $K-1$ active states.*

The conjecture holds for the case of 2 relays as proved in Theorem 5. We proceeded through the following numerical evaluations: for each value of $K-2 \in [2:8]$, we generated uniformly at random the SNR exponents of the channel gains, we computed the entries of $\mathbf{A}$ in (6) with the Hungarian algorithm, we solved the LP in (5) with the simplex method and we counted the number of constraints that equal the optimal gDoF (which is a known upper bound on the number of non-zero entries of an optimal solution). The minimum and the maximum number of active states were found to be 1 and $K-1$, respectively, as shown in Fig. 1, which also shows the average number of active states computed by giving an equal weight to all the tried channels. Note that the minimum number of active states for a *generic* HD relay network with K–2 relays has to be at least K–1. To see this, consider a 'line network' where the source can only communicate with relay 1, relay 1 can only communicate with relay 2, etc, and relay K–2 can only communicate with the destination; in a line network, $K-1$ non-zero states are necessary to enable the source to communicate with the destination. It is interesting that the minimum number of active states given by $K-1$ appears to be also the required maximum number of active states for optimal gDoF-wise network operation. If the reduction of the number of active states from exponential to linear as conjectured holds, it offers a simpler and more amenable way to design the network [9].

V. Conclusions

In this work we studied the Gaussian HD relay network and characterized its capacity to within a constant gap by considering a random switch at the relays. We showed that this gap may be further reduced by considering more structured systems, such as the diamond network. We conjectured that, in a $(K–2)$-relay network, the optimal schedule has at most K–1 active states, instead of the possible 2^{K-2}. This conjecture has been supported by the analytical proof in the case of 2 relays and in general by numerical results. An interesting feature of our conjecture is the formal proof of the equivalence between the problem of finding the coefficients of the gDoF linear program and the associated assignment problem.

Acknowledgment

The work of D. Tuninetti was partially funded by NSF under award number 0643954. Eurecom's research is partially supported by its industrial partners: BMW Group Research & Technology, IABG, Monaco Telecom, Orange, SAP, SFR, ST Microelectronics, Swisscom and Symantec. The research at Eurecom leading to these results has received funding from the EU FP7 grant agreement CONECT n° 257616. The research at IMC leading to these results has received funding from the EU FP7 grant agreement iJOIN n° 317941.

Appendix A
Proof of Theorem 1

We partition the relays into two sets: those in $\mathcal{A}$ lie on the node 1 side of the cut while those in $\mathcal{A}^c$ lie on the node K side of the cut, for all $\mathcal{A} \subseteq [2:K-1]$, $\mathcal{A}^c := [2:K-1]\backslash\mathcal{A}$.

Upper Bound. The cut-set bound for the general FD network [10] adapted to the HD case [3] gives for each $\mathcal{A}$

$$\begin{aligned}
&\max\{R\} \leq I(X_1, X_{\mathcal{A}}, S_{\mathcal{A}}; Y_K, Y_{\mathcal{A}^c} | X_{\mathcal{A}^c}, S_{\mathcal{A}^c}, S_1{=}1, S_K{=}0)\\
&\overset{\text{(a)}}{\leq} |\mathcal{A}| + \sum_{s\in[1:2^{K-2}]} \lambda_s \log\left|\mathbf{I}_{|\mathcal{A}^c|+1} + \mathbf{H}_{\mathcal{A},s}\mathbf{K}_{\{1\}\cup\mathcal{A},s}\mathbf{H}_{\mathcal{A},s}^H\right|\\
&\overset{\text{(b)}}{\leq} |\mathcal{A}| + \sum_{s\in[1:2^{K-2}]} \lambda_s \ \log\left|\mathbf{I}_{|\mathcal{A}^c|+1} + \mathbf{H}_{\mathcal{A},s}\mathbf{H}_{\mathcal{A},s}^H\right|\\
&\quad + \sum_{s\in[1:2^{K-2}]} \lambda_s \mathrm{Rank}[\mathbf{H}_{\mathcal{A},s}] \log\left(\max\{1, \mathrm{Tr}[\mathbf{K}_{\{1\}\cup\mathcal{A},s}]\}\right)\\
&\overset{\text{(c)}}{\leq} |\mathcal{A}| + \sum_{s\in[1:2^{K-2}]} \lambda_s \log\left|\mathbf{I}_{|\mathcal{A}^c|+1} + \mathbf{H}_{\mathcal{A},s}\mathbf{H}_{\mathcal{A},s}^H\right|\\
&\quad + \min\{1+|\mathcal{A}|, 1+|\mathcal{A}^c|\} \log\max\left\{1, \sum_{s\in[1:2^{K-2}]} \lambda_s \mathrm{Tr}[\mathbf{K}_{\{1\}\cup\mathcal{A},s}]\right\}\\
&\overset{\text{(d)}}{\leq} \sum_{s\in[1:2^{K-2}]} \lambda_s \ \log\left|\mathbf{I}_{|\mathcal{A}^c|+1} + \mathbf{H}_{\mathcal{A},s}\mathbf{H}_{\mathcal{A},s}^H\right| \qquad (9)\\
&\quad + \min\{1+|\mathcal{A}|, 1+|\mathcal{A}^c|\} \log(1+|\mathcal{A}|) + |\mathcal{A}| \qquad (10)
\end{aligned}$$

where the inequalities are due to the following facts:

Inequality (a): chain rule of the mutual information, by considering that the discrete random variable $S_{[2:K-1]}$ has at most 2^{K-2} masses and by letting $\lambda_s := \mathbb{P}[S_{[2:K-1]} = s] \in [0,1]$ for $s \in [1:2^{K-2}]$ such that $\sum_{s\in[1:2^{K-2}]} \lambda_s = 1$. Here we use the convention that "$S_{[2:K-1]} = s$" means that the j-th entry of $S_{[2:K-1]}$ is equal to the j-th digit in the binary expansion of the number s. For example: with $K=5$ and $s=4=1\cdot 2^2+0\cdot 2^1+0\cdot 2^0$, the notation "$S_{[2:K-1]}=s$ means $S_2=1, S_3=0, S_4=0$". $\mathbf{K}_{\mathcal{A},s}$ represents the covariance matrix of $X_{\mathcal{A}}$ conditioned on $[S_{[2:K-1]}=s,\ S_1=1, S_K=0]$ and $\mathbf{H}_{\mathcal{A},s}=[(\mathbf{I}_K-\mathrm{diag}[1,s_j,0])\mathbf{H}\mathrm{diag}[1,s_j,0]]_{\{K\}\cup\mathcal{A}^c,\{1\}\cup\mathcal{A}}$. Because of the power constraint we must have $\sum_{s\in[1:2^{K-2}]}\lambda_s\left[\mathbf{K}_{[1:K],s}\right]_{k,k}\leq 1,\ k\in[1:K]$.

Inequality (b): by exploiting the following relation: $0\preceq\mathbf{K}\preceq\lambda_{\max}(\mathbf{K})\mathbf{I}\preceq\mathrm{Tr}[\mathbf{K}]\mathbf{I}$, where 'Tr' is the trace. Moreover for $a\neq 0$ the following holds $|\mathbf{I}+|a|\ \mathbf{K}|\leq\max\{1,|a|\}^{\mathrm{Rank}[\mathbf{K}]}\ |\mathbf{I}+\mathbf{K}|$.

Inequality (c): since the rank of a matrix is at most the minimum between the number of rows and columns.

Inequality (d): by using Jensen's inequality and because of the input power constraints.

Lower Bound. The NNC lower bound for the general FD network [6] adapted to the HD case [3] gives for each $\mathcal{A}$

$$
\begin{aligned}
&\max\{R\}\geq I(X_1,X_{\mathcal{A}};\widehat{Y}_{\mathcal{A}^c},Y_K|X_{\mathcal{A}^c},X_K,S_{[2:K-1]},S_1{=}1,S_K{=}0)\\
&\quad -I(Y_{\mathcal{A}};\widehat{Y}_{\mathcal{A}}|X_{[1:K]},\widehat{Y}_{\mathcal{A}^c},Y_K,S_{[2:K-1]},S_1{=}1,S_K{=}0)\\
&\overset{(a)}{\geq}\sum_{s\in[1:2^{K-2}]}\lambda_s\log\left|\mathbf{I}_{|\mathcal{A}^c|+1}+\frac{1}{1+\sigma^2}\mathbf{H}_{\mathcal{A},s}\mathbf{H}_{\mathcal{A},s}^H\right|-|\mathcal{A}|\log\left(1+\frac{1}{\sigma^2}\right)\\
&\overset{(b)}{\geq}\sum_{s\in[1:2^{K-2}]}\lambda_s\log\left|\mathbf{I}_{|\mathcal{A}^c|+1}+\mathbf{H}_{\mathcal{A},s}\mathbf{H}_{\mathcal{A},s}^H\right|-|\mathcal{A}|\log\left(1+\frac{1}{\sigma^2}\right)\\
&+\sum_{s\in[1:2^{K-2}]}\lambda_s\mathrm{Rank}[\mathbf{H}_{\mathcal{A},s}]\log\left(\min\left\{1,\frac{1}{1+\sigma^2}\right\}\right)\\
&\overset{(c)}{\geq}\sum_{s\in[1:2^{K-2}]}\lambda_s\ \log\left|\mathbf{I}_{|\mathcal{A}^c|+1}+\mathbf{H}_{\mathcal{A},s}\mathbf{H}_{\mathcal{A},s}^H\right| \qquad (11)\\
&-\min\{1+|\mathcal{A}|,1+|\mathcal{A}^c|\}\log(1+\sigma^2)-|\mathcal{A}|\log\left(1+\frac{1}{\sigma^2}\right) \qquad (12)
\end{aligned}
$$

where the inequalities are due to the following facts:

Inequality (a): in all states, we consider i.i.d. $\mathcal{N}(0,1)$ inputs with $Q{=}S_{[1:K]}$ to allow the relays to coordinate, and with $\widehat{Y}_k{:=}Y_k{+}\widehat{Z}_k$ for $\widehat{Z}_k{\sim}\mathcal{N}(0,\sigma^2)$ independent of everything else.

Inequality (b): for $a\neq 0$ the following holds $|\mathbf{I}+|a|\ \mathbf{K}|\geq\min\{1,|a|\}^{\mathrm{Rank}[\mathbf{K}]}\ |\mathbf{I}+\mathbf{K}|$.

Inequality (c): since the rank of a matrix is the minimum between the number of rows and columns.

Gap. By taking the difference between the upper and the lower bounds and by using the possibly suboptimal value $\sigma^2=1$ we immediately obtain the gap in Theorem 1.

Appendix B
Examples of MWBM problems

Example 1: Case $k=n=2$. Consider

$$
\log\left(\left|\mathbf{I}_2+\mathbf{H}\mathbf{H}^H\right|\right)\overset{\mathsf{SNR}\gg 1}{\doteq}\log\left(\mathsf{SNR}^{\max\{\beta_{31}+\beta_{42},\beta_{32}+\beta_{41}\}}\right)
$$

$$
\mathbf{H}:=\begin{bmatrix}h_{31} & h_{32}\\ h_{41} & h_{42}\end{bmatrix}=\begin{bmatrix}\mathsf{SNR}^{\beta_{31}/2}\ e^{j\theta_{31}} & \mathsf{SNR}^{\beta_{32}/2}\ e^{j\theta_{32}}\\ \mathsf{SNR}^{\beta_{41}/2}\ e^{j\theta_{41}} & \mathsf{SNR}^{\beta_{42}/2}\ e^{j\theta_{42}}\end{bmatrix}.
$$

The corresponding MWBM problem has one set of vertices $\mathcal{A}_1$ consisting of $k{=}|\mathcal{A}_1|{=}2$ nodes (refer to these vertices as nodes 1 and 2 – see second subscript in the channel gains) and the other set of vertices $\mathcal{A}_2$ consisting also of $n=|\mathcal{A}_2|=2$ nodes (refer to these vertices as nodes 3 and 4 – see first subscript in the channel gains). The weights of the edges connecting the vertices in $\mathcal{A}_1$ to the vertices in $\mathcal{A}_2$ can be represented as the non-negative weights $\beta_{ij},\ i=3,4,\ j=1,2$. One possible matching assigns node 3 to node 1 and node 4 to node 2 (total weight $\beta_{31}+\beta_{42}$), while the other assigns node 3 to node 2 and node 4 to node 1 (total weight $\beta_{32}+\beta_{41}$); the best assignment is the one that gives the largest total weight. This MWBM is exactly the pre-log of the log-det formula.

To exclude the case of rank deficient channel matrix in our setting we pose a reasonable distribution, such as for example the i.i.d. uniform distribution, on the phases $\theta_{ij},\ i{=}3,4,\ j{=}1,2$, so that almost surely the channel matrix is full rank.

Example 2: Case $k=2, n=3$. The MWBM allows to find the high-SNR approximation of the capacity for any MIMO system. As an example, consider a full-rank MIMO systems with $n=3$ transmit antennas (nodes 1, 2, 3) and $k=2$ receive antennas (nodes 4 and 5) and with SNR-exponent matrix $\mathbf{B}=\begin{bmatrix}\beta_{41} & \beta_{42} & \beta_{43}\\ \beta_{51} & \beta_{52} & \beta_{53}\end{bmatrix}$. In this case we have

$$
\begin{aligned}
\mathrm{MWBM}(\mathbf{B})=\max\Big\{&\beta_{41}+\beta_{52},\beta_{41}+\beta_{53},\beta_{42}+\beta_{51},\\
&\beta_{42}+\beta_{53},\beta_{43}+\beta_{51},\beta_{43}+\beta_{52}\Big\}
\end{aligned}
$$

which can also be obtained, with enough patience, by computing the limiting value of the corresponding log-det formula.

References

[1] T. Cover and A.E. Gamal, "Capacity theorems for the relay channel," *IEEE Trans. on Info. Theory*, vol. 25, no. 5, pp. 572 – 584, Sep. 1979.

[2] A. Host-Madsen, "On the capacity of wireless relaying," in *Proceedings IEEE VTC*, 2002, vol. 3, pp. 1333 – 1337.

[3] G. Kramer, "Models and theory for relay channels with receive constraints," in *Allerton*, 2004, pp. 1312–1321.

[4] A.S. Avestimehr, *Wireless network information flow: a deterministic approach*, Ph.D. thesis, EECS Department, University of California, Berkeley, Oct 2008.

[5] A. Özgür and S.N. Diggavi, "Approximately achieving gaussian relay network capacity with lattice codes," *arxiv:1005.1284*, 2010.

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

[7] B. Chern and A. Özgür, "Achieving the capacity of the n-relay gaussian diamond network within log(n) bits," in *Proceedings IEEE ITW*, 2012, pp. 377 – 380.

[8] H. Bagheri, A.S. Motahari, and A.K. Khandani, "On the capacity of the half-duplex diamond channel," in *Proceedings IEEE ISIT (see also arxiv:0911.1426)*, 2010, pp. 649 –653.

[9] S. Brahma, A. Ozgur, and C. Fragouli, "Simple schedules for half-duplex networks," in *Proceedings IEEE ISIT*, 2012, pp. 1112 –1116.

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

[11] M. Cardone, D. Tuninetti, R. Knopp, and U. Salim, "Gaussian half-duplex relay networks: improved constant gap and connections with the assignment problem," *arxiv:1304.5790*, 2013.

[12] M. Cardone, D. Tuninetti, R. Knopp, and U. Salim, "The capacity to within a constant gap of the gaussian half-duplex relay channel," *to appear in IEEE ISIT*, July 2013.