# IS RATE ADAPTATION BENEFICIAL FOR INTER-SESSION NETWORK CODING?

Yuchul Kim and Gustavo de Veciana
Wireless Networking and Communications Group
Departement of Electrical and Computer Engineering
The University of Texas at Austin
Email: {yuchul.kim@mail,gustavo@ece}.utexas.edu

***Abstract*—This paper considers the interplay between rate adaptation and inter-session network coding gains in wireless ad hoc or mesh networks. Inter-session network coding opportunities at relay nodes depend on packets being overheard by surrounding nodes – the more packets nodes overhear, the more opportunities relays have to combine packets, resulting in a potential increase in network throughput. Thus, by adapting its transmission rate, a node can increase the range over which its packets are overheard, enabling additional opportunities for coding and increased overall throughput. This paper considers inter-session coding, restricted to a single relay (bottleneck) node, or star topology. Even for such simple topologies the optimal joint rate adaptation and network coding policy is known to be NP-hard, so we consider an optimal pairwise coding policy which can be formulated as a linear programming and provide a simple heuristic for rate adaptation for network coding. We evaluate the averaged throughput in two different scenarios, in which relays have different access opportunities, giving some intuition on the impact of rate adaption in lightly and heavily loaded systems. The gains of joint rate adaptation and network coding are marginal when relay has higher access opportunity than other nodes, or when MAC operates ideally. The gain ranges from 9% to 19% compared to a network without network coding and is around 4% over a network using regular network coding. While, when the relay has equal access opportunity as other source nodes, which is more typical of todays MAC protocols under heavy loads, the gain ranges from 40% to 62% as compared to standard relaying case and is upto around 20% as compared to a network with regular pairwise network coding. And we further increase this gain from 40% to 120% by replacing pairwise coding with sub-optimal general network coding scheme.**

## I. Introduction

Since Ahlswede's seminal work [1], network coding has received significant attention. It has since been shown that network coding achieves maximum capacity for multicast sessions in wired networks while increasing the reliability of lossy networks [2]–[6]. Most work to date has focused on, *intra-session* network coding, where only packets in the same session are encoded together, e.g., [4]–[6]. However, the work of Katti and Katabi which proposed the scheme called COPE, does allow coding across different sessions or *inter-session* network coding [7]. They observe that *overheard* packets, in the context of broadcast wireless media, can be effectively exploited to enable network coding resulting in further throughput improvements. This was followed by [8]–[10] where efforts were made to quantify the possible gains in general wireless networks; they show the expected gains are at most a constant factor, e.g., up to around 2 or 3. However, these bounds were found in an analytical framework which is quite idealized, and it is not yet known how much we can actually expect from network coding in practical settings.

This work of Yuchul Kim is supported by a grant of the CISCO research foundation.
978-1-4244-2677-5/08/$25.00 ©2008 IEEE

In this paper, we explore the potential gains of joint rate adaptation and network coding on random star network topologies. The star topology network with single relay and multiple source and destination nodes can be viewed as a basic building block for general wireless networks. The key idea is to transmit at a reduced rate so as to allow additional neighboring nodes to overhear transmissions creating additional opportunities for network coding at a relay node. To this end, we propose a polynomial time heuristic to jointly determine sub-optimal Tx rates and complementary inter-session coding assuming only pairwise coding. We study the performance of the proposed scheme in two regimes with different assumptions, on MAC contention and in particular on the relay node's access to the medium. It turns out that joint rate adaptation and network coding is effective when the relay has a equal access opportunity to the medium as other nodes, or when the relay is congested due to its limited access opportunity.

The rest of this paper is organized as follows. In Section II, we develop the key intuition for coding aware rate adaptation and its potential gain in terms of increased coverage area and rate region. We then formally describe a system and formulate the optimal pariwise coding problem in Section III. In Section IV, we propose a polynomial time heuristic algorithm for sub-optimal rate adaptation. The proposed algorithm is evaluated in Section V. And, we conclude in Section VI.

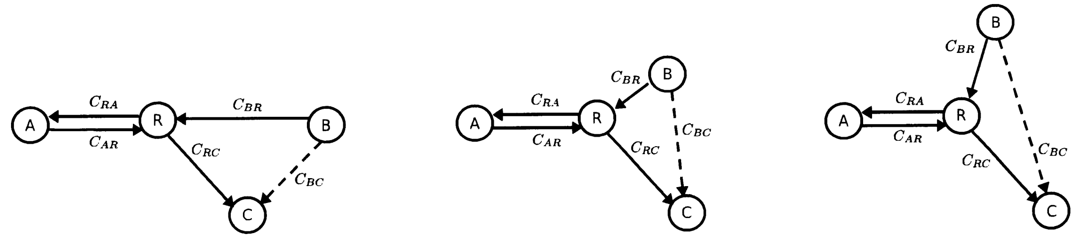
## II. Coding Aware Rate Adaptation

### A. Inter-session network coding

We begin by briefly summarizing the COPE approach [7]. In a network with COPE, destination nodes overhear packet transmissions from neighboring nodes and store them. Information about overheard packets is subsequently sent to neighboring relay nodes. Using this information a relay node combines packets such that the intended destination nodes can decode them. When destination nodes receive a combined packet, they can extract the desired packets from the combined packet using the packets they have previously overheard. This approach is effective at increasing throughput and alleviating congestion in relay nodes acting as traffic 'hubs'.

### B. Intuition underlying inter-session network coding in wireless environments

Note that inter-session coding opportunities at relay nodes depend on the packets overheard by surrounding nodes. That




(a) $C_{BR} < C_{BC}$ (b) $C_{BR} > C_{BC}$ (c) $C_{BR} \gg C_{BC}$

Fig. 1: Three simple toy networks where $A$ sends a packet to $C$ and $B$ sends a packet to $A$ through $R$:(a) $C$'s overhearing $B$ comes for free, (b) capacity can be increased with rate reduction from $C_{BR}$ to $C_{BC}$, and (c) rate reduction degrades network performance.

is, if nodes overhear more packets, a relay may have more opportunities for combining packets, potentially increasing network performance. Until now, researchers have assumed that only nodes in the transmission (Tx) range of the transmitter could overhear packets. So, node placement and the Tx range determined the possibility of such overhearing. However, we argue that one can increase coding opportunities by dynamically increasing the Tx range so that more nodes can overhear a transmission. This can be achieved by increasing the Tx power or reducing the Tx rate. Increasing the Tx power[1] may result in a reduction of spatial reuse, so it is not clear if one can realize increased performance. However by decreasing the Tx rate, one can increase the Tx range while keeping interference power at the same level as before. So the key insight in this paper can be summarized as follows:

*A node reducing its instantaneous Tx rate can increase its overhearing range, leading to possible increase in network coding opportunities and thus in network throughput.*

This statement appears counter intuitive since reducing Tx rate would typically decrease network throughput. Yet, interestingly, it turns out that one can, not only increase network throughput, but also increase the average individual throughputs of nodes. This is best explained through the simple example networks shown in Fig. 1. Using these examples, we will also show that the Tx rate needs to be reduced carefully, otherwise, it can lead to a deterioration of network performance. From now on, we will denote a MAC protocol which uses joint rate adaptation and network coding by RANC.

### C. To code or not to code

Consider the network configuration in Fig. 1 which includes three nodes and one relay. Node $A$ transmits packets to node $C$ and node $B$ transmits packets to node $A$. Both transmissions are relayed through node $R$. Let $C_{XY}$ denote the link capacity between node $X$ and $Y$, and let $C_{AR} = C_{RA} = C_{RC} =$ 1bps. Suppose node $A$ and $B$ each transmit one packet to relay $R$ respectively. Then, node $R$ relays the two packets to node $C$ and $A$ using either network coding or simple relaying. For each case, we will calculate *the network throughput* defined as the total number of bits transported divided by the total amount of time to transport the bits. For simplicity, we assume the instantaneous Tx rate on a link is given by its link capacity. We also assume that if node $B$ transmits at rate $C_{BC}$ to node $C$, then any node $X$ with link capacity $C_{BX} > C_{BC}$ can overhear the transmission.

In Case (a) shown in Fig. 1 where $C_{BR} = 0.5$ and $C_{BC} = 1$, node $B$ transmits with rate $C_{BR}$ and node $C$ can overhear the transmission since node $B$ is closer to node $C$ than to node $R$. In this case, node $R$ can perform a bit-by-bit XOR over the two packets from node $A$ and $B$ and broadcast. In turn, node $C$ receives the coded packet and can decode it since it has the overheard the packet from node $B$. The network throughput for this case is $T_{NC}^{(a)} = \frac{2b}{\frac{b}{1}+\frac{b}{0.5}+\frac{b}{1}} = \frac{1}{2}$, where the $2b$ in numerator is the number of bits transported in the network and $\frac{b}{1} + \frac{b}{0.5}$ in denominator is a time required to send two packets from source nodes $A$ and $B$ to relay $R$. The additional term, $\frac{b}{1}$ in the denominator is the time required for the relay to broadcast the combined packet. In a similar way, the network throughput when only relaying is used can be calculated and is $T_R^{(a)} = \frac{2b}{\frac{b}{1}+\frac{b}{0.5}+\frac{b}{1}+\frac{b}{1}} = \frac{2}{5}$. Clearly when overhearing occurs naturally, inter-session coding increases network throughput.

In Case (b) where $C_{BR} > C_{BC}$, overhearing no longer comes for free. Unless node $B$ decreases its instantaneous Tx rate, it can not ensure node $C$ overhears its transmissions. Once it reduces its instantaneous Tx rate (e.g, by reducing the modulation order or decreasing channel coding rate) node $C$ will overhear the transmission and node $R$ can perform network coding. Suppose, $C_{BR} = 1$ and $C_{BC} = 0.8$. Then, if $B$ reduces its rate to $R$ from 1 down to 0.8, we have $T_{NC}^{(b)} = \frac{8}{13}$, otherwise we have $T_R^{(b)} = \frac{1}{2}$. The benefit of rate reduction should be clear in this case. Hence, for both Case (a) and (b), network coding is beneficial.

Our last Case (c) is different. In this case, even though we can ensure node $C$ overhears by reducing node $B$'s instantaneous Tx rate, this will not increase the network throughput since $B$'s transmission would be excessively slow. In this case, relaying gives higher network throughput: $T_{NC}^{(c)} = \frac{2}{5} < T_R^{(c)} = \frac{1}{2}$. These examples show that the TX rate should be carefully determined.

In Fig. 2, we illustrate the potential locations where instantaneous rate adaptation is beneficial for the simple network topology discussed above. To generate these figures we assumed, a static free space channel model with attenuation factor of 3.5 without fading and shadowing were considered. A transmit power of 1W and noise power spectral density of -174dBm were used. The minimum required SNR for decoding was set to as 3dB. The Tx rate of any link was set to the Shannon capacity of the link. Fig. 2a and Fig. 2b show what is the optimal relaying strategy for all possible locations for node $B$. Here, nodes $A$, $R$ and $C$ are fixed and we move node $B$ to all possible locations in a two

[1]We consider fixed Tx power.

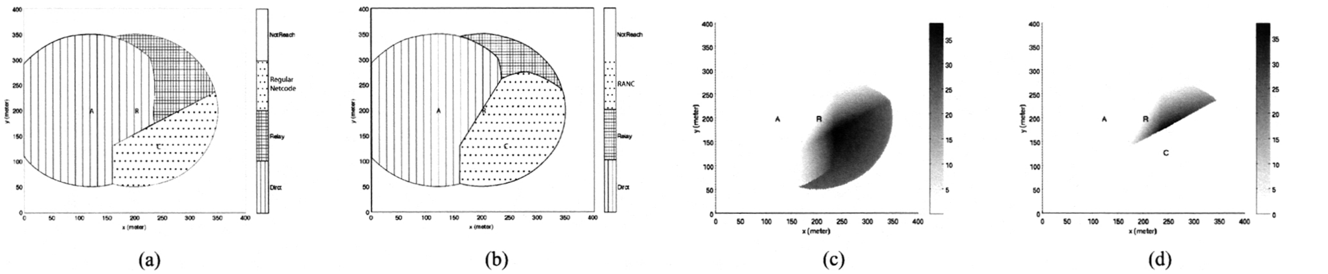



Fig. 2: Optimal relaying strategy for the four nodes network in Fig. 1 as the position of node B is varied: (a) Optimal relaying method in a network with regular network coding, (b) Optimal relaying method in a network with network coding with rate adaptation, (c) Throughput gain (%) of rate adaptive network coding compared to without network coding, (d) Throughput gain (%) of rate adaptive network coding compared to regular network coding.

dimensional square of size $400 \times 400\text{m}^2$. The distance between nodes $A$ and $R$ is equal to that of node $R$ and $C$ - 80 meters. For any given location for node $B$, we first check whether node $B$ can communicate with either node $R$ or $A$. If node $B$ can not communicate with either of them, then, we declare it "not reachable." If node $B$ can communicate with either of them, then, three possible relaying strategies were compared: direct delivery, simple relaying, and network coding. The throughput is computed as previously. The best relaying method, i.e., giving the highest throughput, was selected. One can visualize the expanded region where joint rate adaptation and network coding are used by comparing Fig. 2b to Fig. 2a.

Fig. 2c exhibits throughput gain of rate adaptive network coding compared to direct relaying method. Clearly the gain depends on node $B$'s location. We get the highest gain when node $B$ is at the midpoint of the line connecting nodes $R$ and $C$, since at that location node $B$ can send at its highest rate without rate reduction. Note that for network coding, node $B$ should make sure that both nodes $R$ and $C$ hear the transmission. Fig. 2d exhibits the gain of network coding with rate adaptation versus regular network coding without rate adaptation. One can see that the region where the gain is positive is the same as the expanded region in Fig. 2b. Depending on the location of node $B$, one can expect various gains upto 33%. Note that this is an additional gain from rate adaptation for network coding. More generally one might expect this result to hold for network topologies with more nodes around relay hubs where there are more coding opportunities. In other words, it appears that rate adaptation should reduce the area of region where relaying and direct delivery are better than network coding.

## III. System Model and Problem Formulation

### A. Network Model

We consider a fixed wireless network having a relay $z$ via which a set of source nodes $S$ communicate with a set of destination nodes $D$. Each source node $x \in S$ has a corresponding distinct destination node $y \in D$. A pair of source and destination node specify a session. Let $s(y)$ denote the corresponding source node of a node $y \in D$. Each source node has only one destination node and vice versa[2]. We assume an ideal MAC without contention, i.e., each source nodes take its turns to transmit its packet. The number of bits per packet is fixed.

[2] This assumption can be justified by separating a physical source(destination) node with two different destination(source) nodes into logical two source(destination) nodes.

### B. Transmission rate vector

We define $\mathcal{R} = (r^i, 1 \le i \le |\mathcal{R}|)$ to be an ordered set of discrete rates supported by nodes, where $|\mathcal{R}| < \infty$ denotes the cardinality of the set $\mathcal{R}$. So, $r^1$ and $r^{|\mathcal{R}|}$ are the lowest and the highest supported rates, respectively. The assumption that $\mathcal{R}$ is a finite set is intended to model today's systems, where the number of Tx rates supported is limited due to modulation, slot length, coding rate, etc [11]. Note that even though we have a rate set $\mathcal{R}$, the actual Tx rate at any link is determined by the transmit power, distance, thermal noise, bandwidth, interference, etc. The maximum Tx rate between node $x$ and $y$ is formally defined as

$$r_{xy}^m = \max\left\{r \in \mathcal{R} | r \le \frac{w}{2}\log(1 + \frac{\rho\|x-y\|^{-\alpha}}{\eta + I})\right\},$$

where $\rho$ is the Tx power of a node, $\eta$ is noise power spectral density, $w$ is system bandwidth under consideration, $\alpha$ is the attenuation factor, $\|x-y\|$ is distance between two nodes $x$ and $y$, and $I$ is the amount of external interference power. Losses from imperfect measurements of interference power and thus incorrect rate adaptation will be reflected as a packet error probability[3]. So, the effective Tx rate is given by $r_{xy}^m$ times the packet delivery probability between $x$ and $y$. For $x \in S$, let

$$\mathcal{R}_x = \{\min(r_{xz}^m, r_{xy}^m) | y \in D \cup \{z\}\}$$

be the set of the highest rates from node $x$ to relay $z$ that allow at least one node $y \in D \cup \{z\}$ including $z$ to hear node $x$'s transmission. This can be understood as the set of node $x$'s possible highest Tx rates to relay $z$ which allow overhearing by others. The objective is to determine a vector $\mathbf{r} = (r_{xz} : x \in S)$ of Tx rates that each node will use to transmit to the relay. We have at most $\prod_{x \in S} |\mathcal{R}_x|$ such vectors since each node $x \in S$ has $|\mathcal{R}_x|$ different potential Tx rates.

### C. Overhearing Graph and Clique Partition

For each Tx rate vector $\mathbf{r}$, a set of destination nodes that overhear the transmission can be determined, which then also determines which packets the relay node can combine. Based on

[3] In the evaluation conducted Section V, we ignore interference $I$ since we consider a star topology network without contention.

the overhearing status of each destination node, we construct a graph $G(\mathbf{r})$ such that any two destination nodes which overhear each other's source node are connected by an undirected edge. We identify sets of packets that can be coded together with valid partitions $\mathbf{V}(\mathbf{r})$ of the graph. These are explained below.

The graph $G(\mathbf{r}) = G(D, E_{\mathbf{r}})$ is composed of a set of destination nodes $D$ and a set of undirected edges $E_{\mathbf{r}}$. The link $(i,j) \in D^2$ is in $E_{\mathbf{r}}$ if and only if node $i$ overhears node $s(j)$'s transmission and node $j$ overhears node $s(i)$'s transmission. We say node $i$ overhears node $s(j)$ if node $i$ is in the transmission range of node $s(j)$, or equivalently if $r^m_{s(j)i} \geq r_{s(j)z}$. So given a Tx rate vector $\mathbf{r}$ uniquely determines an overhearing graph. The set of links $E_{\mathbf{r}}$ is formally defined as

$$E_{\mathbf{r}} = \left\{(i,j) \in D^2 \mid r^m_{s(i)j} \geq r_{s(i)z}, r^m_{s(j)i} \geq r_{s(j)z}\right\}.$$

We use this overhearing graph to find optimal sessions to combine. Note that, by construction, sessions corresponding to nodes in any valid clique in $G(\mathbf{r})$ can be combined because each destination node in the valid clique can overhear other destination nodes' source nodes. So, finding a set of sessions to combine corresponds to finding a clique and finding a family of sets of sessions to combine, thus, corresponds to finding a partition $\mathcal{V} = (C_1, C_2, \cdots)$ of $G(\mathbf{r})$ such that each set $C_i$ of the partition is a clique in $G(\mathbf{r})$. We shall refer to such partition as a *clique-partition*. Note that there may be one or more valid clique-partitions in the graph. We let $\mathbf{V}(\mathbf{r})$ be the set of all valid clique-partitions of $G(\mathbf{r})$.

### D. Cost Function and Formulation

For a given $\mathbf{r}$ and $\mathcal{V} \in \mathbf{V}(\mathbf{r})$, we define the *uplink cost* as a total time required for all source nodes to transmit their packets to the relay, i.e., $u(\mathbf{r}) \equiv \sum_{x\in S} \frac{1}{r_{xz}}$. The uplink cost is a function of rate vector $\mathbf{r}$. Similarly, the *downlink cost* is a total time required for relay to send the received packets to corresponding destination nodes, i.e., $d(\mathcal{V}) \equiv \sum_{C\in\mathcal{V}} \frac{1}{\min_{y\in C} r_{zy}}$, where for each $C \in \mathcal{V}$ we take the minimum rate of $r_{zy}$ for $y \in C$ since we want to ensure all destination nodes of the combined packets receive their associated packets. Note that downlink cost depends on the selected combination of sessions or clique partion $\mathcal{V} \in \mathbf{V}(\mathbf{r})$. Our objective is to find an optimal rate vector $\mathbf{r}^*$ and an associated clique partition of $G(\mathbf{r})$ which minimize the sum of the uplink and the downlink cost, i.e., maximize a throughput. This optimization problem is formally stated as follows:

$$\min_{\substack{\mathbf{r}\in\mathbf{R}\\ \mathcal{V}\in\mathbf{V}(\mathbf{r})}} \sum_{x\in S} \frac{1}{r_{xz}} + \sum_{C\in\mathcal{V}} \frac{1}{\min_{y\in C} r_{zy}}. \qquad (1)$$

Let $\mathbf{r}^*$ denote a solution of the above problem. This combinatorial optimization problem can be decomposed as follows. For a given $\mathbf{r} \in \mathbf{R}$, we first need to find the minimum downlink cost in order to evaluate the total cost for the $\mathbf{r}$. Let $\mathcal{V}^*_{\mathbf{r}}$ be the optimal clique partition giving the minimum cost of the downlink under the rate vector $\mathbf{r}$. Then, our original problem can be rewritten as follows:

$$\min_{\mathbf{r}\in\mathbf{R}} u(\mathbf{r}) + d(\mathcal{V}^*_{\mathbf{r}}), \qquad (2)$$

where $\mathcal{V}^*_{\mathbf{r}}$ is given as

$$\mathcal{V}^*_{\mathbf{r}} \in \arg \min_{\mathcal{V}\in\mathbf{V}(\mathbf{r})} d(\mathcal{V}). \qquad (3)$$

Finding the optimal partition $\mathcal{V}^*_{\mathbf{r}}$ is a variant of well known clique partition problem (CPP). CPP partitions a graph $G$ into disjoint cliques with a minimum number of cliques. Determining the minimum clique partition is known to be a NP-hard problem. Indeed problem (3) is reduced to the classical CPP if cost for each clique is equal to 1, see [12], [13]

*Remark:* Note that the decreasing the Tx rate for uplink transmission makes it easier for each node to overhear other transmissions, which makes the overhearing graph more connected. This in principle results in an increase of the possible coded transmissions, which may ends up decreaseing the downlink delays. So, we see the tradeoff between the uplink and the downlink delay. Our objective is to find optimal point where sum of those two delays are minimized.

### E. Pairwise Coding

For simplicity, suppose we restrict the space of clique partitions $\mathbf{V}(\mathbf{r})$ to partitions including cliques of size less than or equal to two, i.e., we determine only an approximate clique partition, $\mathcal{V}^{p*}_{\mathbf{r}}$. In this case, (3) can be rewritten as the following binary integer program:

$$\begin{aligned}
&\min_{a_l, l\in E_{\mathbf{r}}} \sum_{y\in D} \frac{1}{r_{zy}} + \sum_{l\in E_{\mathbf{r}}} a_l c_l \\
&\quad \text{s.t.} \sum_{l\in\{(y,i)\in E_{\mathbf{r}} \mid i\in D\}} a_l \leq 1, \ \forall y \in D \\
&\quad\quad a_l \in \{0,1\}, \quad \forall l \in E_{\mathbf{r}} \\
&c_l = \frac{1}{\min\{r_{zy_1(l)}, r_{zy_2(l)}\}} - \frac{1}{r_{zy_1(l)}} - \frac{1}{r_{zy_2(l)}} \ \forall l \in E_{\mathbf{r}},
\end{aligned} \qquad (4)$$

where $y_1(l)$ and $y_2(l)$ are the two nodes at two end points of undirected link $l$. Also $a_l$ is binary variable for link $l \in E_{\mathbf{r}}$, if packets destined to $y_1(l)$ and $y_2(l)$ then $a_l = 1$, otherwise $a_l = 0$. Our restriction makes the sub-problem (3) easier. In fact the maximization version of (4) corresponds to the well known maximum weighted matching problem (MWMP). MWMP finds a maching $M \subset E$ of graph $G = (V, E)$ such that the sum of weight of $l \in M$ is maximized. The MWMP is the first "true" binary integer programming problem with polynomial time algorithm[4], see [14]. The linear programming relaxation of (4) with additional constraints is given as follows:

$$\begin{aligned}
&\min_{a_l, l\in E_{\mathbf{r}}} \sum_{y\in D} \frac{1}{r_{zy}} + \sum_{l\in E_{\mathbf{r}}} a_l c_l \\
&\quad \text{s.t.} \sum_{l\in\{(y,i)\in E_{\mathbf{r}} \mid i\in D\}} a_l \leq 1, \ \forall y \in D, \\
&\quad \sum_{l\in E(H)} a_l \leq \left\lfloor \frac{|H|}{2} \right\rfloor \ \forall \text{ odd sets } H \subseteq D \\
&\quad\quad a_l \in \mathbb{R}_+, \quad \forall l \in E_{\mathbf{r}},
\end{aligned} \qquad (5)$$

[4]It is "true" binary integer problem in the sense that it can not be solved mearly by linear programming relaxation.

where $c_l$ is defined in (4), $E(H)$ is the set of edges with both ends in $H$. Note the role of additional constraints called odd set constraints, odd set is a set with $k$ nodes, for $k = 3, 5, 7, \cdots$. The constraints restrict the number of pairs in any odd set $H \subseteq D$ be less than or equal to $\lfloor |H|/2 \rfloor$, i.e., it prevents odd cycles. The LP-relaxation without this restiction may not result in integer solution.

## IV. Heuristic Algorithm For Rate Selection

### A. Assumption

We assume that any source node in this network knows the average link rate between itself and other reachable destination nodes including the relay. These might be estimated by observing RTS and CTS signals or pilot signals. This information is shared by each source node with the relay node, which in turn runs following algorithms to perform joint rate adaptation and network coding.

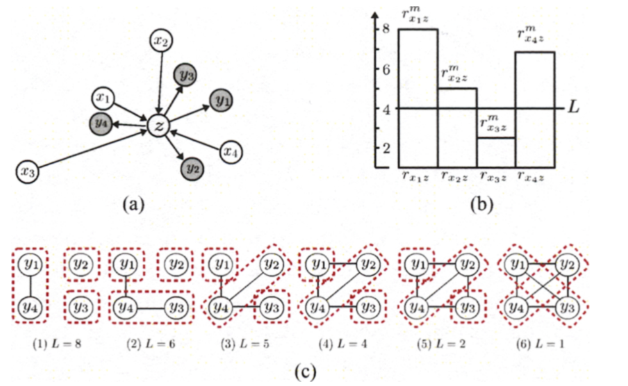


Fig. 3: Tx rate bar lowering and corresponding transformation of overhearing graph in the network with four sessions: (a) Network with four sessions, (b) Max Tx rates of four source nodes and Tx rate bar $L$. The actual Tx rate of node $x_i$ is determined as $\min(r^m_{x_i z}, L)$. (c) As Tx rate bar $L$ decreases from top to bottom overhearing graph (4 nodes black edges between them) changes from (1) to (6) in order. Corresponding optimal pairwise cliques (in dotted red) also change.

### B. Suboptimal Rate Selection Algorithm

In this section, we provide a heuristic algorithm to find sub-optimal rate vector $\hat{\mathbf{r}}^*$, for problem (2). Note that there are $\prod_{x \in S} |\mathcal{R}_x|$ possible Tx rates vectors. Rather than doing an exhaustive search, we shall perform our search as follows. First, note the relation between the Tx rate vector and the overhearing graph in a given network. The maximum Tx rate vector $\mathbf{r}^m = (r^m_{xz} | x \in S)$ results in an overhearing graph $G(\mathbf{r}^m) = G(D, E_{\mathbf{r}^m})$, with a minimum number of edges. Note that this is a subgraph of all possible overhearing graphs that can be generated for all feasible Tx rate vectors. At the other extreme if $\mathbf{r}^1 = (r_{xz} = r^1 | x \in S)$, then the corresponding graph $G(\mathbf{r}^1) = G(D, E_{\mathbf{r}^1})$ is the supergraph of all possible overhearing graphs in this network. So, these two graphs are at extreme ends. For $\mathbf{r} \neq \mathbf{r}^m, \mathbf{r}^1$, the set of edges $E_{\mathbf{r}}$ satisfies $E_{\mathbf{r}^m} \subseteq E_{\mathbf{r}} \subseteq E_{\mathbf{r}^1}$. Our basic strategy is to consider sequence of vectors $\mathbf{r_1}, \mathbf{r_2}, \cdots$, such that $E_{\mathbf{r}^m} \subseteq E_{\mathbf{r_1}} \subseteq E_{\mathbf{r_2}} \subseteq \cdots \subseteq E_{\mathbf{r}^1}$. We shall do this by gradually decreasing the Tx rates by limiting individual Tx rate with Tx *rate bar* $L$. Initially, $L$ is set to the highest rate $r^{|\mathcal{R}|}$, and gradually decreased. For a given $L$, the transmist rate from source to relay may not be feasible, so, the Tx rate of a source node $x \in S$ is set to $\min\{r^m_{xz}, L\}$. With this choice of rates we can quickly estimate the value cost function.

*1) **Clique Partition and Cost Evaluation**:* Every time we lower Tx rate bar $L$, a new graph which possibly includes additional edges is produced. Then, a pairwise clique partition of the graph is found by solving (5). Based on this clique partition, we can evaluate the corresponding downlink cost. To evaluate uplink, we need to only optimize over Tx rates from the sources $S$ to relay that result in the same overhearing graph and thus the same downlink cost. If a clique is of size two, e.g., $\{y_1, y_2\}$, then, $r_{s(y_1)z}$ is be choosen as high as possible, i.e., $\min\{r^m_{s(y_1)z}, r_{s(y_1)y_2}\}$. Note that if $r_{s(y_1)z}$ is increased more than that, then, one can not ensure $y_2$ will overhear its transmission, and the clique is broken. The same applies to $y_2$. If a clique size is one, e.g., $\{y\}$, $r_{s(y)z}$, is reverted to its original max rate $r^m_{s(y)z}$, which may be higher than $L$. This procedure determines the minimum uplink cost.

*2) **Choosing Sub-optimal Rate Vector** $\hat{\mathbf{r}}^*$:* Every time the Tx rate bar decreases a step, a total cost is evaluated. This continues until the bar hits the bottom of the supported rate set $\mathcal{R}$. At this point the rate vector value resulting in the minimum cost is selected as the approximation of the optimal Tx rate vector $\hat{\mathbf{r}}^*$. This algorithm is formally described in Alg.1, where FindCliquePartition is a function which takes the rate vector $\mathbf{r}$ as an input and produces optimal pairwise clique partition $\mathcal{V}^{p*}_{\mathbf{r}}$ as its output, i.e., it solves problem (5), and $K$ is parameter depending on relay's access opportunity.

*3) **Example**:* The algorithm is explained with the example network shown in Fig. 3a. We have four source sessions and corresponding maximum Tx rates from source nodes to relay $z$: $r^m_{x_3 z} < r^m_{x_2 z} < r^m_{x_4 z} < r^m_{x_1 z}$. The Tx rates from the relay to the destination nodes are given as $r_{zy_4} < r_{zy_2} < r_{zy_3} < r_{zy_1}$. The set of supported Tx rates $\mathcal{R}$ is defined as $\{1, 2, \cdots 8\}$.

To check whether there exists any better choice for the Tx rate vector, we first set the Tx bar $L$ to 8 and compute overhearing graph. This graph is shown as (1) in Fig. 3c. Without rate adaptation, node $y_1$ and $y_4$ are overhearing with each other's source nodes so they can be combined. While, node $y_2$ and $y_3$ can not be grouped. If $L$ is lowered below 6, so that the Tx rate of node $x_1$ and $x_4$ are limited by $L$, then, we can make sure that $y_3$ overhears $x_4$'s transmission. This introduces a new link between $y_3$ and $y_4$. When the clique partition is found over this graph, $y_4$ is paired with $y_3$ rather than $y_1$ since $r_{zy_3} < r_{zy_1}$. If we further decrease $L$ to 4, so the Tx rate from $x_2$ is limited by $L$, then, $y_4$ overhears $x_2$'s transmission, which adds a new link between $y_4$ and $y_2$ as shown in (3) of Fig. 3c. Again we can perform clique partitioning over the graph and evaluate minimum cost. In a similar way, if we continue to lower $L$ to 1, then, we obtain the complete graph (6) in Fig. 3c and corresponding clique partition. Among all Tx rates evaluated, the one giving minimum cost is selected. Note that the $L$ is the control parameter for the tradeoff between the two costs.

**Algorithm 1** Sup-optimal Rate Selection Algorithm

```
i ⇐ |R|
while i > 0 do
    L ⇐ r^i
    r_xz ⇐ min{r^m_xz, L}, ∀x ∈ S
    V_r^{p*} ⇐ FindCliquePartition(r)
    r_s(j)z ⇐ max_{y∈C∪{z}\{j}} r^m_s(j)y, ∀j ∈ C, ∀C ∈ V_r^{p*}
    r^i ⇐ (r_xz | x ∈ S)
    E(r^i) ⇐ Σ_{x∈S} K/r_xz + Σ_{W∈V̂} 1/(min_{y∈W} r_zy)
    i ⇐ i − 1
end while
r̂* ⇐ arg min_{1≤i≤|R|} E(r^i)
```

## V. Performance Evaluation

### A. Simulation Environment

In this section, we evaluate the performance of a network using a joint rate adaption and network coding scheme. In particular, we consider a star topology network where a relay node receives packets from source nodes and transmits them to destination nodes. Source and destination nodes are randomly placed within the Tx range of the relay such that the source and destination node can not directly communicate with each other. Each node can support 12 Tx modes, with different modulation and coding rates, transmitter chooses the highest Tx rate supportable based on the received average SNR at the receiver. We say a Tx mode is "supportable" if a desired target PER is achieved under AWGN channel for given average SNR. Static AWGN channel with a path loss attenuation factor of 3.5 was considered. We consider two extreme scenarios, in which relays have different MAC opportunities. In the first scenario, we allow the node acting as relay have higher access opportunity than neighboring nodes acting as data sources. This policy allows us to evaluate the pure gain of network coding, which is not affected by MAC. While, in the second scenario, we assume all nodes including the relay have equal access to the medium. This case shows how network performance is affected by joint rate adaptation and network coding and how these interact with the MAC. Note that, in both cases, we give access opportunity to each source node under max-min fair policy in long term average Tx rate. The only difference is the relay's access opportunity. As a performance metric, network throughput is calculated as the number of bits transported divided by the amount of time spent on uplink and downlink packet transmissions.

### B. Case 1: Relay with more access opportunity than other nodes

In this case, each node takes its turn to transmit a packet (uplink) to the relay and then the relay consumes as many Tx opportunities as it needs to serve the received packets on downlink to destination. The uplink and downlink transmissions form one cycle, which repeats. It keeps max-min fairness on long term average rate across source nodes ($K = 1$). Note the predetermined Tx orders removes contention, and accordingly there is no throughput loss from the contention for the medium. This allows to study the network which achieves its maximum throughput. In this sense, the gains of joint rate adaptation and network coding over simple network coding or no coding at all, can be viewed as pure gains.

Fig. 4 and 5 show an average throughput and an average gain of the rate adaptive network coding (RANC2) and regular network coding (RNC2) both using only pairwise coding policy over random star network topologies. As the number of sessions, $|N|$, increases, the coding opportunities at the relay increase resulting in a throughput increase. The throughput gain of RANC2 compared to no network coding case ranges from 9% to 19% and the gain compared to RNC2 is around 4% as $|N|$ ranges from 2 to 8. Note the gains of baseline or pairwise network coding, looks marginal. This is due to the throughput averaging over random node placement. So, the average gain reduces from the maximum gain of 33% to 6-15% depending on $|N|$. In other words, network coding is not helpful for substantial node placements. In fact, we observed 40% of node placements has no coding opportunity under RNC2 when $|N| = 2$. For RANC2, these cases reduces, but the gain looks still marginal.

### C. Case 2: Relay with equal access opportunity as other nodes

In our second scenario, we assumed that all nodes including the relay have equal access to the medium. That is, the relay, as other nodes, has only one Tx opportunity per cycle. So, the relay behaves like a bottleneck node, in which unsent packets are dropped immediately. So, the relay is supposed to choose one coded or non-coded packet each turn that it gets so as to keep fairness in the long run Tx rates across sessions ($K = |\hat{\mathcal{V}}_{\mathbf{r}}^{\surd *}|$). This scenario gives an idea on the performance of joint rate adaptation and network coding in a heavily loaded networks. Note that all nodes in such a network are assumed to be backlogged and have equal access opportunities.

Fig. 6 and 7 shows the throughput and the thgoughput gain of RNC2, RANC2, RNC and RANC under this scenario, where RNC and RANC are regular and rate adaptive network coding without a limit on the number of packets combined. We first observe that the throughput decreases as $|N|$ increases. It's the natural result of our assumption. As $|N|$ increases, the number of dropped packets also increases and the throughput decreases. The gain of the regular pairwise network coding RNC2 to no coding at all ranges from 20-42%. Once a rate adaptation is applied, the gain increases from 40-62%. However, the relative gain (RANC2 to RNC2) gradually decreases. This is because as the the number of nodes increases, packets are easily paired with another packet even without rate adaptation. Even though pairwise coding looks quite effective as compared to previous scenario, it still does not fully resolve the congestion at the relay.

This motivates us to introduce a more general clique partitioning scheme. For this purpose, a new simple sub-optimal clique partitioning method is introduced at Appendix. Note that the coding with high degree resolves the congestion, and increases throughput. The gain (RANC to noNC) ranges 40% to 120%. And, the relative gain to RNC2 also increases as the number of nodes increases. Note that this implies that increasing the allowable coding degree effectively decreases the number of packets dropped, which is directly translating to throughput increases.

## VI. Conclusion

In this paper, we studied the impact of rate adaptation for inter-session network coding for star topology wireless network. We showed that rate adaptation can be effective at increasing network coding gain and in resolving severe congestion at relay, in particular, in networks where standard MAC protocols are used in which all nodes have equal access opportunities.

## References


[1] R. Ahlswede, N. Cai, S. Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Infomation Theory*, vol. 46, no. 4, pp. 1204–1216, 2000.

[2] S. Y. R. Li, R. W. Yeung, and N. Cai, "Linear network coding," *IEEE Trans. Infomation Theory*, vol. 49, no. 2, pp. 371–381, 2003.

[3] R. Koetter and M. Medard, "An algebraic approach to network coding," *IEEE/ACM Trans. Networking.*, vol. 11, no. 5, pp. 782–795, 2003.

[4] D. S. Lun, M. Medard, and R. Koetter, "Efficient operation of wireless packet networks using network coding," *IWCT*, 2005.

[5] D. S. Lun, M. Medard, and M. Effros, "On coding for reliable communication over packet networks," *Proc. of Allerton Conf. on Comm. Control and Comp.*, Sep 2004.

[6] P. Chou, Y. Wu, and K. Jain, "Practical network coding," *Proc. of Allerton Conf. on Comm. Control and Comp.*, 2003.

[7] Sachin Katti, Hariharan Rahul, Wenjun Hu, Dina Katabi, Muriel Medard, and Jon Crowcroft, "Xor in the air: practical wireless network coding," *ACM SIGCOMM*, 2006.

[8] J. Liu, D. Goeckel, and D. Towsley, "Bounds on the gain of network coding and broadcasting in wireless networks," *Proc. IEEE INFOCOM*, 2007.

[9] J. Le, J. C.S. Lui, and D.M. Chiu, "How many packets can we encode? - an analysis of practical wireless network coding," *Proc. IEEE INFOCOM*, 2008.

[10] A. K. Haddad and R. Riedi, "Bounds on the benefit of network coding: throughput and energy saving in wireless networks," *Proc. IEEE INFOCOM*, 2008.

[11] "Ieee 802.11a standard," .

[12] Dion Gijswijt, Vincent Jost, and Maurice Queyranne, "Clique partitioning of interval graphs with submodular costs on the cliques," Tech. Rep. TR-2006-14, Egerváry Research Group, Budapest, 2006, www.cs.elte.hu/egres.

[13] Saul G. de Amorim, Jean-Pierre Barthelemy, and Celso C. Ribeiro, "Clustering and clique partitioning: Simulated annealing and tabu search approaches," *Journal of Classification*, 1992.

[14] George L. Nemhauser and Laurence A. Wolsey, *Integer and Combinatorial Optimzation*, WILEY, 1988.

[15] Jong Tae Kim and Dong Ryeol Shin, "New efficient clique partitioning:algorithms for register-transfer synthesis of data paths," *Journal of the Korean Physical Society*, vol. 40, no. 4, pp. 754–758, April 2002.


## Appendix

A heuristic clique partitioning algorithm based on [15].

**Algorithm 2** Sup-optimal Clique Partition Algorithm

1: $N \Leftarrow D$
2: **while** $N \neq \phi$ **do**
3: Select $p \in N$ with minimum degree. Tie breaking: select $p$ with minimum tx rate from relay to $p$.
4: Select a node $q$, a neighbor of $p$, with the maximum common edges with $p$. Tie breaking: select $q$ with minimum Tx rate from relay to $q$
5: Delete edges from $p$ and $q$ that do not connect to their common neighbor.
6: Merge $p$ and $q$ and rename it as $p$ with transmission rate $\min\{r_{zp}, r_{zq}\}$.
7: If $p$ has any remaining edge goto step 4 otherwise $p$ become a new clique. Remove $p$ from $N$.
8: **end while**

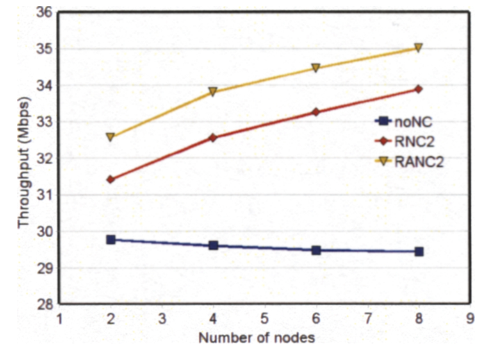



Fig. 4: The average throughput when the relay has higher access priority

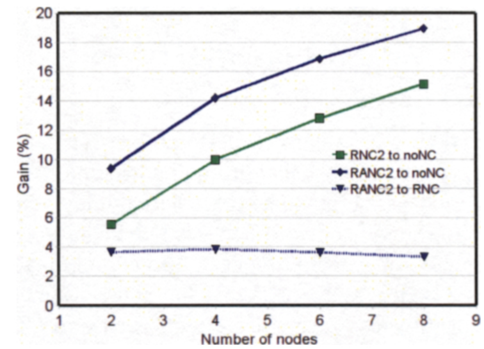



Fig. 5: The gain average throughput when relay has higher access priority

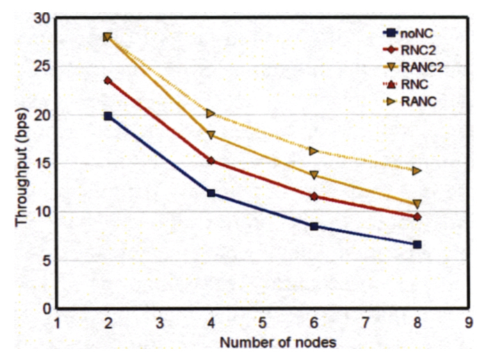



Fig. 6: The average throughput when the relay have equal access priority

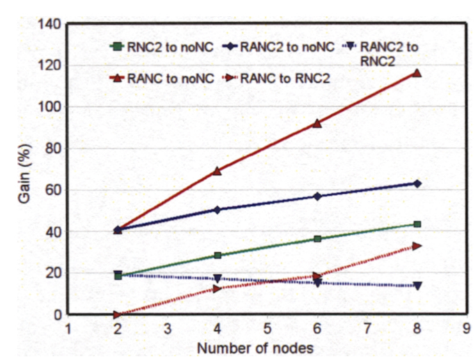



Fig. 7: The gain of average throughput when the relay has equal access priority