Flow-based XOR Network Coding for Lossy Wireless Networks

Abdallah Khreishah, Issa M. Khalil, Pouya Ostovari, and Jie Wu

***Abstract*—A practical way for maximizing the throughput of a wireless network is to decompose the network into a superposition of small two-hop networks such that network coding can be performed inside these small networks to resolve bottlenecks. We call these networks *2-hop relay networks*. Therefore, studying the capacity of 2-hop relay networks is very important. Most practical network coding protocols that perform the superposition ignore the diversity among the links by turning off coding when the channels are lossy. Other protocols deal with the packets separately – not as members of flows – which makes the network coding problem with lossy links intractable. In this paper, we use a different approach by looking at flows or batches instead of individual packets. We characterize the capacity region of the 2-hop relay network with packet erasure channels when the coding operations are limited to XOR. We derive our results by constructing an upper bound on the capacity region and then providing a coding scheme that can achieve the upper bound. The capacity characterization is in terms of linear equations. We also extend our 2-hop relay networks results to multihop wireless networks by providing a linear program that can perform the superposition optimally. We perform extensive simulations for both the 2-hop relay and large wireless networks and show the superiority of our protocols over the network coding protocols that deal with the packets separately.**

***Index Terms*—Network coding, wireless networks, capacity, fairness, 2-hop relay networks, packet erasure channels.**

I. Introduction

One of the fundamental challenges in wireless network research is characterizing the capacity of such networks. The capacity refers to the set of all possible end-to-end rates that can be achieved by the users simultaneously [2]. Characterizing the capacity for wireless networks is not a straightforward extension from wireline networks. This is due to the unique characteristics of wireless networks, such as the broadcast nature, the interference among the links, the diversity, and the lossy behavior of the wireless links.

Traditionally, the broadcast nature of wireless links is considered a challenge due to the interference effect it creates and the unnecessary multiple copies of the same packet it produces. If we allow intermediate nodes to code the packets, the broadcast nature becomes an opportunity that needs to be exploited. Take Fig. 1 as an example: if the broadcast nature of wireless links is not exploited, and assuming that nodes A and C are out of range of each other, we need four transmissions to exchange two packets between nodes A and C. The relay node B can exploit the broadcast nature of its output links and reduce the number of transmissions to three by XORing the two packets, as shown in the figure. We call this example the Alice and Bob example. Typically, the broadcast nature of wireless networks is considered a disadvantage, due to the interference effect it creates. Network coding exploits the broadcast nature property by coding the packets which results in a reduction in the number of transmissions. In the example in Fig. 1, coding the two packets together allows the relay node B to send one broadcast transmission instead of two. The reason is that the coded packet can be utilized by both nodes A and C to recover their own packets. With network coding, the broadcast nature of wireless links turns into an advantage instead of being a disadvantage.

In *intersession network coding* (IRNC), intermediate relay nodes code packets from different flows at intermediate nodes. IRNC exploits the broadcast nature of wireless links and reduces the number of packets to be sent, as explained in the example in Fig. 1. In general, it is hard to perform IRNC because the problem is NP-hard [3], and linear coding is not sufficient for the problem [4]. However, one can limit coding opportunities to be in the local neighborhood. Empirical studies have shown substantial throughput improvement in wireless networks when IRNC coding is limited to local XOR opportunities, as in COPE [5]. The example in Fig. 1 represents COPE. The local neighborhood structure is termed as a *2-hop relay network*, we will discuss this later. Based on the COPE approach, the problem of coding-aware routing and scheduling was studied in [6]. The formulation in [6] involves linear programming that is computed centrally. The work in [7] studied the fundamental limit of how many sessions can be encoded simultaneously together when COPE is used. The fundamental limit depends on geometry; therefore, the maximum number of sessions that can be coded together under a typical setting is limited to five. The work in [8] considered an algorithm with lower complexity than COPE and designed its optimal scheduler. Our previous work [9] considered *pairwise* IRNC that allows coding over multihops, but limits coding to be among only two original packets. We designed its corresponding optimal scheduler and rate controller.

As can be seen from the previous discussion, IRNC is well suited when the links are not lossy. However, IRNC does not work well when the links have a moderate loss probability of 20%, as the work in [5] turns off coding in this case. In [10], IRNC with lossy links is considered. However, the authors did not optimize overhearing and limited the operations to be

Abdallah Khreishah, Jie Wu and Pouya Ostovari are with the Department of Computer & Information Sciences, Temple University, Philadelphia, PA. E-mail: {akhreish,jiewu,ostovari}@temple.edu. Issa Khalil is with the Faculty of Information Technology, United Arab Emirates University, UAE. E-mail: ikhalil@uaeu.ac.ae. Part of this work appeared in Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM), Houston, USA, Dec 5-9, 2011 [1].

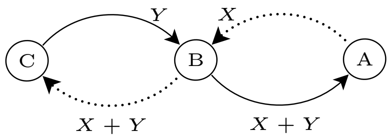

Fig. 1. An example of a network with two flows.

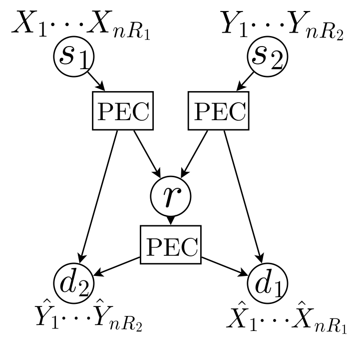

Fig. 2. A 2-hop relay network with two sessions.

only XOR. The optimal solution was found to be #P-complete, and several approximation algorithms were obtained. The work in [11] considered energy efficiency in lossy wireless networks with XOR-based IRNC and provided a heuristic to solve the IRNC problem.

The reason that the optimal solution for lossy 2-hop relay netwoks is #P-complete is that the packets were considered separately, not as members of flows. Several works have looked at the network coding problem with flows, but with different settings than in this paper. In [12], [13], the authors studied the capacity of two-way relay channel. The channel model they adopt is the Gaussian channel model. As we will explain in this paper, the two-way relay channel is a special case of the two-hop relay network, and the packet erasure channel we adopt here is different from the Gaussian channel model and represents realistic scenarios as we will describe in Section II. In [14] a similar problem to ours is studied but for the Gaussian channel model. In [15], [16], the authors studied the reverse carpooling problem with network coding which is similar to the 2-way relay channel that is a special case of our problem. In [17], the authors study two-hop relay networks, but for only two sessions and with a packet-by-packet feedback. As we will see in Section II, we consider arbitrary number of sessions and one feedback message per batch. The most related works to ours are [18], [19] as a 2-hop relay network is studied with packet erasure channels. However, in these works coding is allowed over any finite field size while here coding is limited to XOR, as the nodes in many wireless networks have limited computational power and cannot perform operations over large finite fields.

Our contributions in this paper are as follows:

- We characterize the capacity region of 2-hop relay networks with packets erasure channels using linear equations when only XOR operations are used. This makes it easy to use different objective functions. These objective functions can represent the sum of the throughput, strict fairness, or proportional fairness. The reason we limit the operations to XOR is that XOR is a lightweight operation. Also, most kinds of wireless networks have limited computational power such that they cannot perform operations over large finite fields.
- We provide a coding scheme that can achieve the capacity with very few feedback messages.
- We extend our results to general multihop networks by using the 2-hop relay networks as building blocks.
- We perform different simulations to evaluate our schemes. Our simulation results show that the optimal solution for the capacity region can increase the throughput of 2-hop relay networks by 82% while enhancing the fairness, compared to the state-of-the-art approaches. For the multihop networks our scheme can enhance the capacity by up to 5 times.

The rest of the paper is orginized as follows. In Section II, we describe the network settings and then present the characterization in Section III. We provide an achievable rate region for a general multihop that uses the 2-hop relay networks results as building blocks in Section IV. We then present the simulation results in Section V and conclude the paper in Section VI.

II. The Settings

The two-hop relay network consists of N sessions, where each session i is represented by the source node s_i, the destination node d_i, and the rate R_i that should be supported between s_i and d_i. The destination node d_i can not overhear the source node packets, but can overhear other sources' packets. Therefore, we use the relay node r to code different session packets and to send the coded packets through its outgoing broadcast link so that the overall capacity region can be enhanced. Node r receives a limited number of feedback messages from $d_i, \forall i$ about the overheard packets to help in deciding the coded combination. Fig. 2 represents the 2-hop relay network for two sessions, i.e., $N = 2$. In the figure, PEC stands for *packet erasure channel*. The PEC is a broadcast channel where every sent packet can be received by any subset of the receivers. The reception at the receivers depends on the probability of reception between the source and any individual receiver. We use p_{uv} to denote the reception probability at node v of the packet sent by node u. We assume that the reception processes across the individual links of the PEC are independent.

For example, when $N = 2$, each of s_1, s_2, and r can use the corresponding PEC n times, respectively. Source s_1 would like to send $n \times R_1$ packets $X_1, \cdots, X_{nR_1}$ to destination d_1, and s_2 would like to send $n \times R_2$ packets $Y_1, \cdots, Y_{nR_2}$ to d_2. We are interested in the largest achievable rate pair (R_1, R_2) that guarantees the recoverability of $X_1, \cdots, X_{nR_1}$ from the coded packets $\hat{X}_1, \cdots, \hat{X}_{nR_1}$ at d_1 and the recoverability of $Y_1, \cdots, Y_{nR_2}$ from the coded packets $\widehat{Y}_1, \cdots, \widehat{Y}_{nR_1}$ at d_2 (with a close-to-1 probability for sufficiently large n when node r is limited to perform only XOR operations). We also assume that the destination nodes do not store the XORed packets. They only store the non-coded packets and use them for future decoding.

Symbol	Definition
N	Number of sessions
i, j	Index for a session
R_i	Rate of session i
s_i	Source of session i
d_i	Destination of session i
r	Relay node
X, Y	Symbols to represent the packets
t_r^A	Fraction of time the relay node sends XORed packet formed by packets belonging to all of the sessions in A
t_i	Fraction of time node s_i is scheduled
p_{uv}	Delivery rate between nodes u and v
$R_{i,B}$	The rate of the packets that is sent by node s_i and overheard by only the nodes in $r \bigcup(\bigcup_{j\in B, j\neq i} d_j)$
x_i^A	The achievable rate for session i from the auxiliary session formed by XORing packets from the sessions in the set A
$x_i^{A,B}$	The achievable rate for session i from the auxiliary session formed by XORing packets, from the sessions in the set A, with the constraint that session i packets are used in XORing, are received by exactly all the nodes in $r \bigcup(\bigcup_{j\in B, j\neq i} d_j)$
α	Path loss order
T^*	Decodable SNR threshold
D	The Euclidean distance
Y_i^{AB}	The set of packets for session i that are overheard by nodes d_j, $j \in B, j \neq i$ and to be coded in the auxiliary session A

TABLE I
SUMMARY OF THE SYMBOLS USED IN THIS PAPER

Note that because the 2-hop relay network results from decomposing the big network into smaller 2-hop relay networks, we have the constraints that d_i cannot overhear s_i. This agrees with the practical COPE protocol setting. Also, note that the Alice and Bob example in Fig. 1 is a special case of our settings. This can be done by setting $N = 2$ and $p_{s_1 d_2} = p_{s_2 d_1} = 1$. In the literature this network is also called reverse carpooling or two-way relay channel. Note also that the channel model that we adopt here represents the realistic scenarios. In the realistic scenarios the sent signal is either received entirely by the receiver if the received signal to noise and interference ratio (SINR) is above a threshold or dropped by the receiver otherwise. This is exactly the behavior that the packet erasure channel represents.

We use t_r^A to represent the fraction of time over which the relay node sends XORed packets that were formed by the packets of the sessions in the set A. We also use x_i^A to represent the achievable rate for session i from the auxiliary session formed by XORing packets from the sessions in set A. Symbol x_i^{AB} represents the achievable rate for session i from the auxiliary session formed by XORing packets from the sessions in set A, with the constraint that session i packets used in XORing are received by exactly all of the nodes in $r \bigcup(\bigcup_{j\in B} d_j)$ before being XORed i.e., the received packets by the relay node and the set of receivers for the sessions in the set B. We use $R_{i,B}$ to represent the rate at which packets sent by s_i are overheared by r and exactly all of the nodes d_j, $j \in B, i \neq j$. Throughout the paper, we use the term "auxiliary session" to refer to the session formed by XORing different packets from different sessions.

Table I summarizes the symbols used in the next section.

III. The Capacity Region

A. The Characterization

The following theorem characterizes the capacity region of the 2-hop relay networks when the relay node r is limited to performing XOR operations.

Theorem 1: The capacity region of the 2-hop relay network, when only XOR operations are allowed, can be represented by the following set of equations:

$$R_i \leq \sum_{A:i\in A} x_i^A, \forall i \quad (1)$$

$$x_i^A \leq t_r^A p_{rd_i}, \forall A, i \in A \quad (2)$$

$$x_i^A = \sum_{B:(A\backslash i)\subseteq B} x_i^{AB} \forall A, i \in A \quad (3)$$

$$\sum_{A:(A\backslash i)\subseteq B} x_i^{AB} = t_i R_{i,B}, \forall B, i \notin B \quad (4)$$

Proof: We prove our theorem by showing that the constrains are necessary and sufficient.

Necessity: Using XOR coding, any coded packet is formed by XORing packets of sessions i, $\forall i \in A$, where A is a set of sessions belonging to the power set of all sessions. Constraint (1) states that the total rate of session i is the sum of the achievable rate for session i from all of the auxiliary sessions A, where $i \in A$.

Since t_r^A is the frequency of sending XORed packets by the relay node formed by XORing packets of the sessions in the set A, node d_i will receive XORed packets for the auxiliary session A from the relay node at rate $t_r^A p_{rd_i}$. Therefore, constraint (2) should be satisfied for any achievable XOR-based code.

Note also that (2) does not require the coded packet for the auxiliary session A to be received by all of d_i, $i \in A$; every time it is sent, any one of the d_i that receive this packet can decode it, and it will count as a decodable packet.

For any auxiliary session A, and $i \in A$, the set of the packets for session i that are XORed in this auxiliary session should be received from s_i by all of the nodes in the set $r \bigcup(\bigcup_{j\in A, j\neq i} d_j)$. The reason for that is because r should be able to relay the XORed packets formed in part by these packets, and also because all d_j should have enough remedy packets to remove the components corresponding to these packets from the XORed packets and to recover their respective packets. Also, the set of packets for session i that are received from s_i by any super set of $r \bigcup(\bigcup_{j\in A, j\neq i} d_j)$ can be used in the XORed auxiliary session A because this will guarantee that all of the nodes in the set $r \bigcup(\bigcup_{j\in A, j\neq i} d_j)$ have received these packets. This explains the constraint (3).

The right hand side of (4) $R_{i,B}$ represents the rate of session i packets received by exactly all of the nodes in the set $r \bigcup(\bigcup_{j\in B, j\neq i} d_j)$ after being sent by s_i. These packets can be used by any auxiliary session A such that $(A\backslash i) \subseteq B$. This is because this guarantees that all of the nodes $d_j, j \in A, i \neq j$ will have enough remedy packets to remove session i components in the XORed packets. Therefore, we have constraint (4). We postpone the calculation of a closed form expression for $R_{i,B}$ until the end of this section.

Note that the packets sent by s_i can be divided among all of the auxiliary sessions A, $i \in A$. This is due to the following:

- Because the right hand side of (4) represents the rate at which an exact, specific set of nodes are receiving the packets from s_i. Therefore, every triple (i, A, B) can be assigned an exclusive share of these packets.
- Because each x_i^{AB} appears only once in (3), the packets of session i that are used in the auxiliary session A will be $\bigcup_{B:(A\setminus i)\subseteq B} Y_i^{AB}$, where Y_i^{AB} are the set of packets assigned for the triple (i, A, B).

Sufficiency (an achievable coding scheme):

- Node s_i, $\forall i$ keeps trying to send its nR_i packets one-by-one until all of them are received by the relay node.
- Feedback messages from all d_i about the overheard packets are sent to the relay node r.
- For every set A, the relay node chooses the corresponding feasible $x_i^A, \forall i$ using the linear program that is represented by the constraints ((1)-(4)) and the appropriate objective function. It also assigns nx_i^A packets for every A and i, such that these packets are received by r and all $j \in A, j \neq i$. As explained before, we can assign unique packets for every A.
- For every A, the relay node XORs one packet from each nx_i^A packet for all $i \in A$ and then sends it. If this packet is received by d_j for $j \in A$, this means that the packets belonging to session j in the XORed packets can be recovered by d_j. Therefore, we remove this packet from the set of packets assigned to j and A at the relay node. The relay node keeps performing the XORing and sends until all of the packets assigned for set A at the relay node are sent.

The necessary and sufficient conditions together prove our theorem. ■

Note that the last step in the achievable coding scheme assumes instant feedback. To avoid such an assumption, the relay node can use *fountain codes* [20], [21] and achieves the same rates asymptotically using only XOR operations. The fountain codes can be used as follows:

- The relay node applies a fountain code on every set of packets Y_i^{AB} separately.
- The relay node performs XOR on these packets, as explained before.
- Upon receiving these coded packets, the destination nodes can recover the fountain coded packets, because they overheard the remedy packets.
- The destination nodes apply the inverse of the fountain code to retrieve the original packets.

Note that the use of fountain codes does not increase the complexity of our algorithm. This is due to the use of only XOR operations. Also, fountain codes works on batches which agrees with the settings of our algorithm. Therefore, the number of feedback messages will be very low. Note also that because COPE [5] works on a packet by packet mode it misses a lot of coding opportunities, because coding is performed on the head of queue packets. Also, COPE requires enabling overhearing for any sent packet as a form of feedback while here we need only one feedback message per batch.

B. *Computing* $\mathbf{R_{i,B}}$

In this section, we provide a closed form expression for $R_{i,B}$ when the links are independent. The closed form solution is not straightforward because every packet has to be received by the relay node. Therefore, every s_i has to keep sending a packet until it is received by the relay node. We have:

$$
\begin{aligned}
R_{i,B} &= (\text{delivery rate from } s_i \text{ to } r) \\
&\quad \times (\text{probability that } r \text{ receives a} \\
&\text{symbol and by the time the symbol is received by } r \\
&\quad \text{it is received by only the nodes in } d_j, \\
&\quad j \in B,\ j \neq i) \\
&= p_{s_i r} \sum_{n=1}^{\infty} \text{Probability}\{r \text{ receives the packet on time} \\
&\text{slot } n\} \times \text{Probability}\{\text{only the nodes in } d_j,\ j \in B \\
&\text{receive the packet in time slots } 1, \dots, n\} \\
&= p_{s_i r} \Big[\sum_{n=1}^{\infty} p_{s_i r} (1 - p_{s_i r})^{n-1} \big(\prod_{j \notin B} (1 - p_{s_i d_j})^n\big) \times \\
&\prod_{j \in B} (1 - (1 - p_{s_i d_j})^n) \Big] \\
&= p_{s_i r}^2 \times \sum_{n=1}^{\infty} \prod_{j \notin B} (1 - p_{s_i d_j}) \Big[(1 - p_{s_i r}) \\
&\prod_{j \notin B} (1 - p_{s_i d_j}) \Big]^{n-1} \left[\prod_{j \in B} (1 - (1 - p_{s_i d_j})^n) \right]
\end{aligned}
$$

Therefore, we have:

$$
\begin{aligned}
R_{i,B} &= p_{s_i r}^2 \prod_{j \notin B} (1 - p_{s_i d_j}) \times \sum_{n=0}^{\infty} \Big[(1 - p_{s_i r}) \\
&\prod_{j \notin B} (1 - p_{s_i d_j}) \Big]^n \left[\prod_{j \in B} (1 - (1 - p_{s_i d_j})^{n+1}) \right] \\
&= p_{s_i r}^2 \prod_{j \notin B} (1 - p_{s_i d_j}) \times \sum_{n=0}^{\infty} \Big[[(1 - p_{s_i r}) \\
&\prod_{j \notin B} (1 - p_{s_i d_j})]^n [\sum_{H: H \subseteq B} (-1)^{|H|} \prod_{k \in H} (1 - p_{s_i d_k})^{n+1}] \Big].
\end{aligned}
$$

By Fubini's theorem [22], we have:

$$
\begin{aligned}
R_{i,B} &= p_{s_i r}^2 \prod_{j \notin B} (1 - p_{s_i d_j}) \sum_{H: H \subseteq A} (-1)^{|H|} \prod_{k \in H} (1 - p_{s_i d_k}) \\
&\left[\sum_{n=0}^{\infty} [(1 - p_{s_i r}) \prod_{j \notin B} (1 - p_{s_i d_j}) \prod_{k \in H} (1 - p_{s_i d_k})]^n \right] \\
&= p_{s_i r}^2 \prod_{j \notin B} (1 - p_{s_i d_j}) \sum_{H: H \subseteq B} (-1)^{|H|} \prod_{k \in H} (1 - p_{s_i d_k}) \\
&\left[\frac{1}{1 - [(1 - p_{s_i r}) \prod_{j \notin B} (1 - p_{s_i d_j}) \prod_{k \in H} (1 - p_{s_i d_k})]} \right]
\end{aligned}
$$

Note that our results can be extended to the case of flexible scheduling, such that every source node s_i is scheduled for

Symbol	Definition
N	Number of sessions
i, j	Index for a session
R_i	Rate of session i
s_i	Source of session i
d_i	Destination of session i
r	Relay node
$\mathcal{P}_i$	Path for session i
p_{uv}	Delivery rate between nodes u and v
$R_{i,B}(u)$	The rate of session i packets that are sent by $\phi(u,i)$ overheard by only the nodes in $u \bigcup (\bigcup_{j\in B, j\neq i} \gamma(u,j))$
$x_i^A(u,v)$	The achievable rate through link (u,v) for session i from the auxiliary session formed by XORing packets from the sessions in the set A
$x_i^{A,B}(u,v)$	The achievable rate for session i through link (u,v) from the auxiliary session formed by XORing packets, from the sessions in the set A, with the constraint that session i packets are used in XORing, are received by exactly all the nodes in $r \bigcup (\bigcup_{j\in B, j\neq i} d_j)$
$\mathcal{C}(u)$	The set of sessions such that there respective paths use node u as an intermediate node
$\phi(u,i)$	Previous-hop node of u on $\mathcal{P}_i$
$\gamma(u,i)$	Next-hop node of u on $\mathcal{P}_i$
t_u^A	The fraction of the time node u is scheduled to send packets from the auxiliary session A.

TABLE II
SUMMARY OF THE SYMBOLS USED IN THIS PAPER

a t_i fraction of the time. This can be done by multiplying the closed form for $R_{i,B}$ by t_i. Using our approach, we can maximize or minimize any objective function. This makes our approach more flexible as we will see in our simulation results.

IV. Extension to the Multihop Case

In this section, we use our 2-hop relay network results as building blocks in large lossy multihop networks. We assume that there are N sessions in the network. The i-th session has a source node s_i, a destination node d_i, and a rate R_i that should be supported between the source and the destination. We use $\mathbb{P}_i$ to represent a path between s_i and d_i. As can be seen, our results can be extended to the case of multiple paths between s_i and d_i. We use $\phi(u,i), (\gamma(u,i))$ to represent the previous-hop (next-hop) node on $\mathbb{P}_i$ for node u. We also use $\mathcal{C}(u)$ to represent the set of sessions that use node u as an intermediate node. Similar to the 2-hop relay network results, symbol $X_i^A(u,v)$ is used to refer to the rate of packets sent through link (u,v) for session i from the auxiliary session formed by XORing packets from the sessions in set A. Also, symbol $X_i^{AB}(u,v)$ is used to represent the rate of packets sent through link (u,v) for session i from the auxiliary session formed by XORing packets from the sessions in set A, with the constraint that session i packets that are used in XORing are received by exactly all of the nodes in $r \bigcup \big(\bigcup_{j\in B, j\neq i} \gamma(u,j)\big)$. We also use t_u to represent the fraction of time that node u is scheduled for, and t_u^A is used to represent the fraction of time that node u is scheduled to send Xored packets from set A. Table II represents the symbols used in the multihop case.

The following set of constraints represents an achievable capacity region using the optimal local XOR coding for the 2-hop relay network as building blocks.

$$\sum_{A:i\in A} X_i^A(u,v) - \sum_{A:i\in A} X_i^A(v,w) \leq 0, \quad \forall i, v, \quad s.t. \quad (u,v)\in \mathbb{P}(i), (v,w) \in \mathbb{P}(i) \tag{5}$$

$$R_i - X_i^{\{i\}}(s(i),u) \leq 0, \forall, \quad \forall i \quad s.t. \quad (s(i),u) \in \mathbb{P}(i) \tag{6}$$

$$X_i^A(u,v) \leq t_u^A p_{u,v}, \forall u, i \in \mathcal{C}(u), A \subseteq \mathcal{C}(u), (u,v) \in \mathbb{P}(i) \tag{7}$$

$$X_i^A = \sum_{B:(A\setminus i)\subseteq B} x_i^{AB} \forall (u,v) \in \mathbb{P}(i), A, B \subseteq \mathcal{C}(u) \tag{8}$$

$$\sum_{A:(A\setminus i)\subseteq B} x_i^{AB} = t_{\phi(u,i)}^{\{i\}} R_{i,B}, \forall (u,v) \in \mathbb{P}(i), A, B \subseteq \mathcal{C}(u) \tag{9}$$

$$[\sum_u t_u^A] \in \mathcal{CO}(\vec{\alpha}). \tag{10}$$

We can compute $R_{i,B}(u)$ using (11).

Constraint (5) represents balance equations; the total rate of the sent packets by a node should be equal to the total rate of the received ones. Constraint (6) is for s_i, as it can not encode the packets for this session because coding will not provide a gain in this case. Constraints (7)-(11) are obtained by modeling each node as a relay node for a local 2-hop relay network.

Let $\{\vec{\alpha}\}$ represent the set of all possible link scheduling possibilities. Then, the fraction of time that each node u is scheduled ($\sum_A t_u^A$) can be represented by the convex hull of $\{\vec{\alpha}\}$, which depends on the interference model. In this paper we assume the use of IEEE 802.11, while other models can be used under our settings. In the simulations, we construct the conflict graph of our network in a similar way to [23], [6]. Then, similar to [23], [6], we use the independent set constraints to come up with a lower bound on the rate region and the clique constraints to come up with an upper bound. We report the results only when the upper and lower bound meet. It should be noted that both the independent set and clique constraints can be expressed using linear constraints. Note that the use of the independent sets and the cliques is just to for evaluation purposes run the linear program. This is equivalent to running the IEEE 802.11 protocol in reality which allows us to make fair comparison with COPE that uses the IEEE 802.11.

V. Simulations

In this section, we present simulation results to show the effectiveness of our flow-based scheme over the schemes that deal with packets separately. We first present simulations for the 2-hop relay networks case, and then we present the large wireless networks case. We conducted the simulations using CPLEX and MATLAB which are standard tools for performing linear programming. Other similar works [6], [23] have used these standard software which make reproducing the results easy.

$$R_{i,B}(u) = p^2_{\phi(u,i)u} \prod_{j \notin B} (1 - p_{\phi(u,i)\gamma(u,j)}) \sum_{H: H \subseteq B} (-1)^{|H|} \prod_{k \in H} (1 - p_{\phi(u,i)\gamma(u,k)})$$
$$\left[\frac{1}{1 - [(1 - p_{\phi(u,i)u}) \prod_{j \notin B} (1 - p_{\phi(u,i)\gamma(u,j)}) \prod_{k \in H} (1 - p_{\phi(u,i)\gamma(u,k)})]} \right] \tag{11}$$

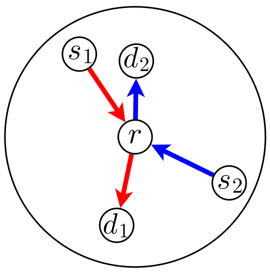

Fig. 3. A figure representing the settings of the simulations.

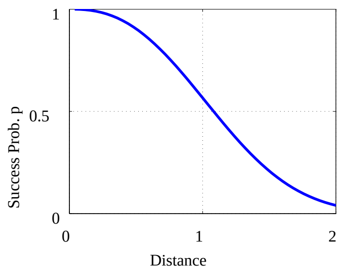

Fig. 4. The relationship between the distance and the signal strength for the Rayleigh fading channel.

A. 2-hop Relay Networks Results

We construct a unit circle with the relay r placed at the center. We then place N source nodes s_i and N destination nodes d_i in the circle at random (see Fig. 3). The only condition we impose is that for each (s_i, d_i) pair, d_i must be in the 90-degree pie area opposite to s_i (see Fig. 3). The reason for this assumption is that the final objective of studying 2-hop relay networks is to use them in larger networks as building blocks. Therefore, d_i should be reachable from s_i through r. Otherwise, if in a large multihop network d_i is directly reacheable from s_i, we should call d_i as r and then try to find another node for d_i to form a 2-hop relay network. For each randomly constructed network, we use the Euclidean distance between each node to determine the overhearing probability. More explicitly, for any two nodes separated by distance D, we use the Rayleigh fading model to decide the overhearing probability:

$$p = \int_{T^*}^{\infty} \frac{2x}{\sigma^2} e^{-\frac{x^2}{\sigma^2}} dx,$$

where we choose:

$$\sigma^2 \triangleq \frac{1}{(4\pi)^2 D^\alpha}$$

the path loss order $\alpha = 2.5$, and the decodable SNR threshold $T^* = 0.06$. Fig. 4 respresents the relationship between the overhearing probability p and the distance D. We assume that the overhearing event among different receivers is independent.

For each randomly generated network, we compute the overhearing probabilities and use the corresponding linear constraints on the time-sharing variables' ts and the rate variables' Rs to compute the achievable rate of each scheme.

Given a randomly generated network, the achievable sum rates are computed for all of the schemes. We then repeat this computation for 1,000 randomly generated networks. Let $\zeta^*_{\text{scheme},k}$ denote the achievable sum rate of the given scheme for the k-th randomly chosen topology. We are interested in the following two performance metrics: The average sum rate over 1000 topologies, $\frac{1}{1000}\sum_{k=1}^{1000} \zeta^*_{\text{scheme},k}$ and per topology improvement $\triangleq \frac{\zeta^*_{\text{scheme},k} - \zeta^*_{\text{baseline},k}}{\zeta^*_{\text{baseline},k}}$.

Fig. 5 represents the average sum rate over the 1000 topologies for different values of N and for different schemes. The simulated schemes are: (1) COPE, from [5], which is the basic XOR-based coding scheme; (2) CLONE [10], which is the state-of-the-art, loss-aware coding scheme that deals with the packets separately, not as members of flows. Two versions of CLONE are simulated. These are CLONE-binary and CLONE-multi. The details of the two CLONE schemes are described in [10]. It's worth noting that CLONE-multi has a very large complexity, which makes it difficult to report the results for $N = 6$; (3) Our optimal scheme. Since our optimal scheme can be cast with different objective functions, we simulate three objective functions. These are maximizing the total throughput "Cap-Sum", achieving strict fairness "Cap-Strictf", and achieving proportional fairness "Cap-PrFair".

As can be seen from the figure, COPE performs poorly under the lossy links environment. Also, the average throughput using COPE decreases as the number of sessions increases. CLONE-binary and CLONE-multi perform better than COPE, but the average throughput does not increase as the number of sessions increases. Our optimal scheme outperforms all of the other schemes. When the objective is to maximize the total throughput, our scheme enhances the average throughput by $1.8 - 3.7$ fold compared to COPE, depending on the number of sessions. Our scheme also enhances the average throughput over CLONE-multi by $1.5 - 1.8$ fold and about $1.2 - 1.45$ fold over CLONE-binary, depending on the number of sessions. Even when the objective function is strict fairness or proportional fairness, our scheme enhances the throughput over the best state-of-the-art scheme by around 20%. This shows that our scheme outperforms all of the previous ones in both fairness and throughput. Fig. 6 represents the CDF function of the per topology throughput for different schemes when $N = 6$. The results in the figure confirm our results.

Fig. 7 represents the CDF for the per topology percentage gain that can be obtained by our schemes compared to CLONE-binary. As can be seen from the figure, for some topologies, the gain of our scheme, when the objective is to maximize the throughput, is about 82%. Also, for 20% of the topologies, the gain is above 58%. This means that we can find topologies where our scheme can almost double the capacity of the network over the state-of-the-art schemes with lower complexity. Fig. 7 also shows that when the objective is to achieve strict or proportional fairness, there are topologies

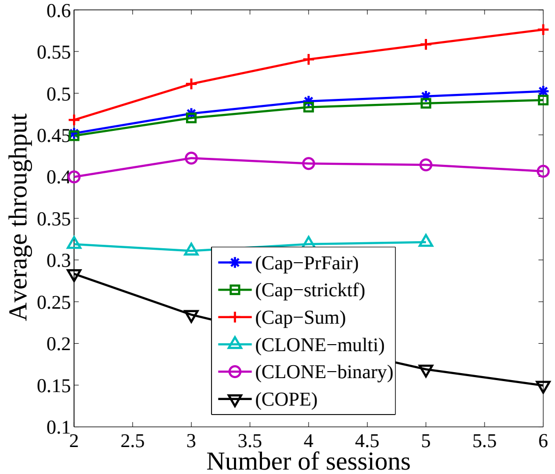

Fig. 5. The average throughput for 1,000 topologies with different values of N.

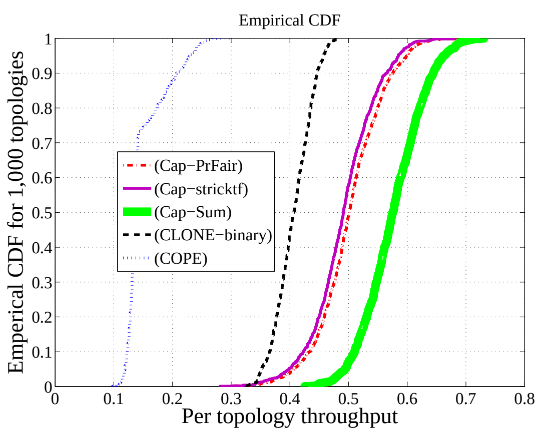

Fig. 6. The CDF of the total achievable rate for the 1,000 topologies when $N = 6$.

that our scheme can do while increasing the throughput by 60%. These results show that our schemes can achieve fairness and maximize the throughput by a moderate amount simultaneously. This joint objective has been targeted by many works [24], [25], [26], but none have been able to get moderate improvement in both directions. It is worth mentioning that for only less than 2% of the simulated topologies, our schemes reduced the throughput compared to CLONE-binary in order to achieve the fairness objective.

B. Large Wireless Networks Results

We generate 100 different random topologies inside a square of length 15. Each generated topology has 20 nodes, and we vary the number of sessions from 4 to 10. We adopt the same channel model with the same parameters as in the 2-hop relay networks' simulations. We simulate 5 different schemes. These are: (1) COPE from [5]; (2) the scheme in Section IV with the objective of maximizing the total throughput, we refer to this scheme as XOR; (3) the scheme in Section IV with the objective of achieving strict fairness among the flows, we refer

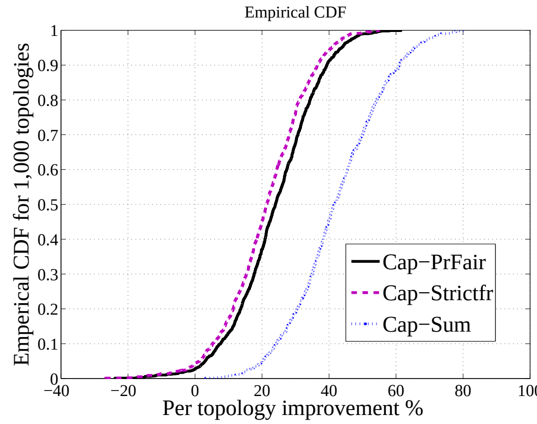

Fig. 7. The CDF of the per topology rate improvement compared to CLONE-binary for the 1,000 topologies when $N = 6$.

to this scheme as XOR-Fair; (4) the scheme in Section IV with multipaths and the objective of maximizing the total throughput, we refer to this scheme as XOR-Multi; (5) the scheme in Section IV with multipaths and the objective of achieving strict fairness among the flows, we refer to this scheme as XOR-Multi-Fair. We do not simulate CLONE, as it is not designed for multihop networks.

Fig. 8 represents the average throughput of the different schemes. As can be seen from the figure, when the objective is to maximize the throughput our scheme achieves 4-4.5 times that of COPE. Even when the objective is to achieve strict fairness, our scheme still enhances the throughput by 15-20% over COPE. Fig. 9 represents the emperical CDF function of the per topology improvement of our schemes with respect to COPE. When the objective is to maximize the throughput, our schemes always outperform COPE. This improvement varies from 20% to about 35 fold. In about 90% of the topologies, our schemes achieves more than double the achievable throughput by COPE, and in 10% of the topologies, the improvement is above 10 fold. When the objective is to achieve fairness among the flows our scheme reduces the throughput in about 1% of the topologies. For some topologies, the throughput improvement is 17 fold while the strict fairness is guaranteed. We compare the average fairness index used in [27] that is achieved by the different schemes. For a given topology, the fairness index is computed as $\frac{(\sum_{i=1}^{N} R_i)^2}{N\sum_{i=1}^{N} R_i^2}$. The fairness index varies from 0 to 1 and bigger values mean a better fairness performance. As can be seen from Fig. 10, our schemes improve the fairness index by 2-4 fold. Also, when multipaths are used with more coding opportunities, the fairness index increases.

VI. Conclusion

In this work, we took a different look at the local inter-session network coding problem in lossy wireless networks. We considered the case where the coding operations at the relay node are limited to XOR operations. We also considered flows instead of individual packets and characterized the corresponding capacity region. Our characterization turned out to be

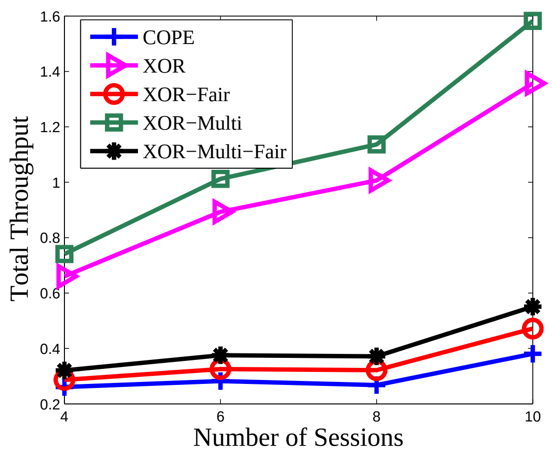

Fig. 8. The average sum throughput for the 100 large topologies with different number of sessions.

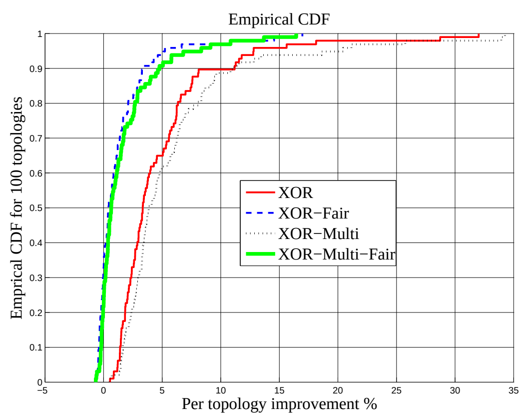

Fig. 9. The CDF of the per topology rate improvement compared to COPE for the 100 large topologies each with 8 sessions.

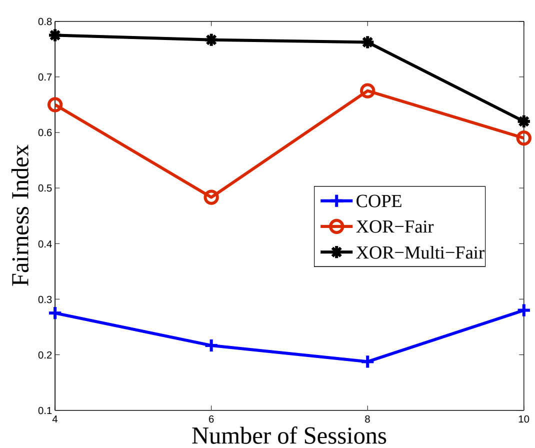

Fig. 10. The average fairness index for the 100 large topologies with different number of sessions.

in terms of linear constraints, which is tractable compared to the characterization without flows. We also provided a coding scheme that achieves the capacity. We then used the local coding results as building blocks in large wireless networks and represent the corresponding achievable rate region using linear constraints. Our simulation results showed the superiority of our scheme in terms of throughput and fairness. Our future work will be to develop distributed algorithms for the large wireless multihop networks case.

VII. Acknowledgment

This research was supported in part by NSF grants ECCS 1128209, CNS 1065444, CCF 1028167, CNS 0948184, and CCF 0830289.

References

[1] A. Khreishah, J. Wu, P. Ostovari, and I. Khalil, "Flow based xor network coding for lossy wireless networks," in *IEEE Global Telecommunications Conference (GLOBECOM), Houston, TX, USA*, Dec 2011.

[2] P. Gupta and P. Kumar, "The capacity of wireless networks," *IEEE Trans. on Info. Theory*, vol. 46, no. 2, pp. 388–404, 2000.

[3] A. Lehman and E. Lehman, "Complexity classification of network information flow problems," in *Proc. of ACM-SIAM SODA*, Jan 2004.

[4] R. Dougherty, C. Freiling, and K. Zeger, "Insufficiency of linear coding in network information flow," *IEEE Trans. on Information Theory*, vol. 51, no. 8, pp. 2745–2759, 2005.

[5] S. Katti, H. Rahul, W. Hu, D. Katabi, M. Medard, and J. Crowcroft, "XORs in the air: Practical wireless network coding," in *Proc. of ACM Special Interest Group on Data Commun. (SIGCOMM), Pisa, Italy*, Sept 2006.

[6] S. Sengupta, S. Rayanchu, and S. Banerjee, "An analysis of wireless network coding for unicast sessions: The case for coding-aware routing," in *Proc. of IEEE Conf. on Computer Communications (INFOCOM), Anocharage, AK*, May 2007.

[7] J. Le, J. Lui, and D. Chiu, "How many packets can we encode? — an analysis of practical wireless network coding,," in *Proc. of 27th IEEE Conf. on Computer Communications (INFOCOM). Phoenix, USA,*, April 2008.

[8] P. Chaporkar and A. Proutiere, "Adaptive network coding and scheduling for maximizing throughput in wireless networks," in *Proc. of ACM Int'l. Conf. on Mobile Computing & Networking (MobiCom) Montréal, Canada,*, Sep 2007.

[9] A. Khreishah, C.-C. Wang, and N. Shroff, "Cross-layer optimization for wireless multihop networks with pairwise intersession network coding," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 5, pp. 606–621, 2009.

[10] S. Rayanchu, S. Sen, J. Wu, S. Banerjee, and S. Sengupta, "Loss-aware network coding for unicast wireless sessions: Design, implementation, and performance evaluation." in *Proc. of ACM Sigmetrics, Anapolis, MD*, June 2008.

[11] T. Cui, L. Chen, and T. Ho, "Energy efficient opportunistic network coding for wireless networks,," in *Proc. of 27th IEEE Conf. on Computer Communications (INFOCOM). Phoenix, USA,*, April 2008.

[12] A. Avestimehr, A. Sezgin, and D. Tse, "Approximate capacity of the two-way relay channel: A deterministic approach," in *46th Annual Allerton Conference on Communication, Control, and Computing, Monticello, IL, USA*, Sep 2008.

[13] W. Nam, S.-Y. Chung, and Y. Lee, "Capacity of the gaussian two-way relay channel to within 1/2 bit," *IEEE Transactions on Information Theory*, vol. 56, no. 11, pp. 5488–5494, 2010.

[14] Y. Wu, "Broadcasting when receivers know some messages a priori,," in *Proc. IEEE Intl Symp. Inform. Theory. Nice, France,*, June 2007.

[15] J. Marden and M. Effros, "The price of selfishness in network coding," in *Proc. of Workshop on Network Coding (NETCOD)*, June 2009.

[16] A. Parandehgheibi, A. Ozdaglar, M. Effros, and M. Medard, "Optimal reverse carpooling over wireless networks: A distributed optimization approach," in *Proc. of CISS, Princeton, NJ*, March 2010.

[17] L. Georgiadis and L. Tassiulas, "Broadcast erasure channel with feedback – capacity and algorithms,," in *Proc. 5th Workshop on Network Coding, Theory, & Applications (NetCod). Lausanne, Switzerland,*, June 2009.

[18] C.-C. Wang, “On the capacity region of 1-hop intersession network coding – a broadcast packet erasure channel analysis with general message side information,,” in *Proc. of IEEE ISIT*, June 2010.

[19] C.-C. Wang, A. Khreishah, and N. Shroff, “On cross-layer optimizations for intersession network coding on practical 2-hop relay networks,,” in *Asilomar Conference*, November 2009.

[20] M. Luby, “LT codes,” in *Proc. of IEEE Symposium on the Foundations of Computer Science (FOCS)*, Nov. 2002.

[21] A. Shokrollahi, “Raptor codes,” *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2551–2567, 2006.

[22] W. Rudin, *Real and complex analysis*. McGraw-Hill, 1987.

[23] K. Jain, J. Padhye, V. Padmanabhan, and L. Qiu, “Impact of interference on multi-hop wireless network performance,,” in *ACM MobiCom. San Diego, CA,*, Sept 2003.

[24] T. Bonald and L. Massoulie, “Impact of fairness on Internet performance,” in *Proc. of ACM Joint Int’l Conf. on Measurement and Modeling of Computer Systems (Sigmetrics), Cambridge, MA*, June 2001.

[25] J. Mo and J. Walrand, “Fair end-to-end window-based congestion control,” *IEEE/ACM Trans. on Networking*, vol. 8, no. 5, pp. 556–567, 2000.

[26] F. P. Kelly, A. Maulloo, and D. Tan, “Rate control in communication networks: Shadow prices, proportional fairness and stability,” *Journal of the Operational Research Society*, vol. 49, pp. 237–252, 1998.

[27] R. Jain, D. Chiu, and W. Hawe, “A quantitative measure of fairness and discrimination for resource allocation in shared computer systems,” Digital Equipment Corporation, Tech. Rep., Tech. Rep., September 1984.

Abdallah Khreishah received his Ph.D and M.S. degrees in Electrical and Computer Engineering from Purdue University in 2010 and 2006, respectively. Prior to that, he received his B.S. degree with honors from Jordan University of Science & Technology in 2004. Immediately after finishing his Ph.D, Abdallah joined the Department of Computer & Information Sciences at Temple University as an Assitant Professor. During the last year of his Ph.D, Abdallah worked with NEESCOM. Abdallah is an active researcher. His research spans the areas of network coding, wireless networks, congestion control, cloud computing, network security, and database systems for large projects. His research projects are funded by the National Science Foundation and the UAE Research Foundation. Abdallah is also an active teacher at Temple University where he taught several graduate and undergraduate courses.

Issa Khalil received the B.Sc. and the M.S. degrees from Jordan University of Science and Technology in 1994 and 1996 and the PhD degree from Purdue University, USA in 2006, all in Computer Engineering. Immediately thereafter he worked as a post-doctoral researcher at the Dependable Computing Systems Lab of Purdue University until he joined the Faculty of Information Technology (FIT) of the United Arab Emirates University (UAEU) in August 2007 where he now an associate professor. Khalils research interests span the areas of wireless and wireline communication networks. He is especially interested in security, routing, and performance of wireless Sensor, Ad Hoc and Mesh networks. His current research is funded by grants from National Research Foundation, Emirates Foundation, and United Arab Emirates University. Dr. Khalil served as the technical program co-chair of the 6th International Conference on Innovations in Information Technology and was appointed as a Technical Program Committee member and reviewer for many international conferences and journals. In June 2011 Khalil has been granted the FIT outstanding professor award for outstanding performance in research, teaching and service.

Pouya Ostovari received the B.S degree in computer engineering from Shahid Beheshti Univeristy, Tehran, Iran, in 2006, and the M.S degree from the Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran, in 2009. He is currently working toward the Ph.D. Degree at The Department of Computer and Information Sciences, Temple University. He is a research assistant at the Temple University. His research interests include wireless networks, sensor networks, and distributed systems.

Jie Wu is chair and Laura H. Carnell professor in the Department of Computer and Information Sciences at Temple University. Prior to joining Temple University, he was a program director at the National Science Foundation. His research interests include: wireless networks, mobile computing, routing protocols, fault-tolerant computing, and interconnection networks. He has published more than 550 papers in various journals and conference proceedings. He serves in the editorial boards of IEEE Transactions on Computers and Journal of Parallel and Distributed Computing. Dr. Wu was also general co-chair for IEEE MASS 2006, IEEE IPDPS 2008, and DCOSS 2009 and is the program co-chair for IEEE INFOCOM 2011. He has served as an IEEE Computer Society distinguished visitor. Currently, he is the chair of the IEEE Technical Committee on Distributed Processing (TCDP), an ACM distinguished speaker and a Fellow of the IEEE. Dr. Wu is the recipient of 2011 China Computer Federation (CCF) Overseas Outstanding Achievement Award.