On the k-pairs problem

Ali Al-Bashabsheh and Abbas Yongacoglu

School of Information Technology and Engineering

University of Ottawa, Ottawa, Canada

{aalba059, yongacog}@site.uottawa.ca

Abstract— We consider network coding rates for directed and undirected k-pairs networks. For directed networks, meagerness is known to be an upper bound on network coding rates. We show that network coding rate can be $\Theta(|V|)$ multiplicative factor smaller than meagerness. For the undirected case, we show some progress in the direction of the k-pairs conjecture.

I. INTRODUCTION

It is known that the min-cut is a necessary and sufficient condition for achievable throughputs in multicast networks [1]. In general networks, the min-cut is not a sufficient condition and there is no answer yet to what rates are achievable in such networks. In this work we consider the *k-pairs problem* which is also referred to as the *multiple unicast sessions* problem. For directed networks, it is known that network coding may provide higher rates than routing. On the other hand, for undirected k-pairs networks, Li and Li conjectured that network coding can not provide higher rates than fractional routing [2]. This conjecture has been verified to be true for few networks [2] [3] [4] [5].

II. DEFINITIONS AND PROBLEM FORMULATION

A directed graph $G(V, E)$ is specified by a set of nodes V and a set of directed edges E (an incidence function is not necessary since in this work all graphs are assumed to be simple graphs). For any edge $e = (u, v)$ we write $\text{head}(e) = v$ and $\text{tail}(e) = u$. For a node $v \in V$ we denote by $\text{In}(v) = \{e \in E : \text{head}(e) = v\}$ the set of all edges going into v and by $\text{Out}(v) = \{e \in E : \text{tail}(e) = v\}$ the set of all edges departing from v. Moreover, for any $U \subseteq V$ we use $\text{In}(U) = \{e \in E : \text{head}(e) \in U, \text{tail}(e) \notin U\}$ to indicate the set of all edges entering U. Similarly, $\text{Out}(U) = \{e \in E : \text{tail}(e) \in U, \text{head}(e) \notin U\}$ denotes the set of edges outgoing from U.

Similar to directed graphs, an undirected graph $G(V, E)$ is specified by two sets V and E where edges do not have a prespecified direction and can provide a bidirectional transportation of information. At some places, an undirected edge e between nodes u and v might be replaced with two directed edges (u, v) and (v, u) whose capacities sum to the capacity of e. For notational ease, we might drop the parenthesis and use uv and vu to denote the edges directed from u to v and v to u, respectively, while preserving the notation $\{u, v\}$ for the undirected edge between u and v. The set of all directed edges obtained from E will be denoted E_d, i.e., $E_d = \{(u, v) : \{u, v\} \in E\}$. In this work, all edges are assumed to have unit capacity.

A directed (undirected) k-pairs network consists of an underlying directed (undirected) graph, $G(V, E)$, and a set of k source-sink pairs. A source-sink pair uniquely identifies a commodity to be communicated from the source to the sink. Let $\mathcal{I} = \{1, 2, \ldots, k\}$ be the set of commodities, then for any $i \in \mathcal{I}$ we use $s(i) \in V$ and $t(i) \in V$ to denote the nodes which (respectively) generates and demands commodity i. We refer to $s(i)$ as the *source node* and $t(i)$ as the *sink node* of $i \in \mathcal{I}$ and always assume $s(i) \neq t(i)$. Note that a node $v \in V$ can be a source node or a sink node for more than one commodity. We denote by $S(v)$ the set of all commodities for which v is a source node, i.e., $S(v) = \{i \in \mathcal{I} : s(i) = v\}$. Similarly, let $T(v)$ be the set of all commodities for which v is a sink node, i.e., $T(v) = \{i \in \mathcal{I} : t(i) = v\}$. Also for any set of nodes $U \subseteq V$, let $S(U) = \{i \in \mathcal{I} : s(i) \in U\}$ be the set of all commodities whose sources are in U and $T(U) = \{i \in \mathcal{I} : t(i) \in U\}$ be the set of all commodities whose sinks are in U.

Given an undirected network, $\mathcal{N}$, with an underlying graph $G(V, E)$ and a set of commodities $\mathcal{I}$. A set of edges $A \subseteq E$ is said to *separate* commodity $i \in \mathcal{I}$ if every path from $s(i)$ to $t(i)$ contains at least one edge from A. Let $\mathcal{J}$ be the set of commodities separated by A, then *sparsity* [6] [7] of A is defined as $\mathscr{S}(A) = |A|/|\mathcal{J}|$. Moreover, the sparsity of the graph is defined as $\mathscr{S}_G = \min_{A \subseteq E} \mathscr{S}(A)$. It is clear that sparsity is a bottleneck for the communication pairs (Indeed, some authors refer to $\mathscr{S}_G$ as the min-cut bound [6]). Thus, in undirected networks, sparsity is an upper bound on achievable rates with or without network coding. Another bound on routing rates can be defined in terms of the Wiener index. For any pair of nodes $u \neq v \in V$ let $d(u, v)$ be the number of edges in the shortest path between u and v in G. The *wiener index* [8] of a graph G is defined as $D_G = \sum_{\{u,v\} \subseteq V} d(u, v)$ which is a commonly used quantity in chemical literature. We define the wiener index of the network as $D_{\mathcal{N}} = \sum_{i \in \mathcal{I}} d(s(i), t(i))$. Obviously, if their is a commodity between every pair of distinct nodes in the network, i.e. $|\mathcal{I}| = \binom{|V|}{2}$, then the wiener indices of the graph and the network are equal. The *wiener bound* of the network $\mathcal{N}$ is defined as $W_{\mathcal{N}} = |E|/D_{\mathcal{N}}$. Clearly $W_{\mathcal{N}}$ is an upper bound on achievable routing rates in undirected k-pairs networks. This follows since in routing, if an edge is used to transpose a fraction of commodity i, then an equal fraction of the edge capacity is exclusively used by such commodity, i.e, an edge does not carry a combination of messages from different commodities.

For directed networks, sparsity is still an upper bound on routing rates but it is not an upper bound on network coding rates. *Meagerness* was introduced in [7] to bound network coding rates in directed networks. For any set $A \subseteq E$ of edges and a set of commodities $\mathcal{J} \subseteq \mathcal{I}$ we say A *isolates* $\mathcal{J}$ if every path from $s(i)$ to $t(j)$ $\forall i, j \in \mathcal{J}$ contains at least one edge from A. The meagerness of set A is defined as

$$\mathcal{M}(A) = \min_{\mathcal{J}:A \text{ isolates } \mathcal{J}} \frac{|A|}{|\mathcal{J}|}$$

and the meagerness of the network $\mathcal{N}$ is defined as $\mathcal{M}_\mathcal{N} = \min_{A \subseteq E} \mathcal{M}(A)$.

With each commodity $i \in \mathcal{I}$ we associate a R.V. X_i which represents a message generated at $s(i)$ and to be correctly recovered at $t(i)$. For notational convenience we might use set subscript. More specifically, let A be any set, then $X_A = \{X_a : a \in A\}$ (if A is empty we set X_A to be a constant). Also with each directed edge $e = uv$ we associate a R.V. X_{uv} which is a deterministic function of $X_{S(u)}$ and $X_{\text{In}(u)}$. A sink node $t(i)$ must be able to recover its message using only the information available from $\text{In}(t(i))$ and $\{X_j : j \in S(t(i))\}$. In other words, each sink recovers its message by computing a function of $X_{\text{In}(t(i))}$ and $X_{S(t(i))}$. The set of edge functions and sink functions defines a *network code*. Such network code implies an achievable rate tuple $(r_1, \ldots, r_k)$ where $r_i \leq H(X_i)$ is the rate at which the ith commodity is communicated. Obviously, $H(X_e)$ must not exceed the edge capacity $\forall e \in E$. An *achievable symmetric rate* is the rate tuple $(r, r, \ldots, r)$ which can be uniquely identified with the scalar r. The *network coding rate* is defined as the supremum of all achievable symmetric rates with network coding.

The condition that X_{uv} is a function of $X_{S(u)}$ and $X_{\text{In}(u)}$ is equivalent to $H(X_{\text{In}(u)}, X_{S(u)}, X_{uv}) = H(X_{\text{In}(u)}, X_{S(u)})$ since $H(X_{uv}|X_{\text{In}(u)}, X_{S(u)}) = 0$. From monotonicity of entropy, the previous equality can be written as $H(X_{\text{In}(u)}, X_{S(u)}, X_{uv}) \leq H(X_{\text{In}(u)}, X_{S(u)})$. This has motivated the authors in [3] to define the *input-output* inequality which states that for any $U \subseteq V$, $H(X_{\text{In}(U)}, X_{S(U)}, X_{\text{Out}(U)}, X_{T(U)}) \leq H(X_{\text{In}(U)}, X_{S(U)})$. Finally, at some places we use the *submodularity* of entropy which asserts that for any sets A_1 and A_2 we have $H(X_{A_1}) + H(X_{A_2}) \geq H(X_{A_1 \cup A_2}) + H(X_{A_1 \cap A_2})$.

III. DIRECTED NETWORKS

Meagerness was introduced in [7] to bound network coding rates in directed networks. In the same work, the authors provided a network referred to as the split butterfly to illustrate that the meagerness bound might not be tight. In this section we show that network coding rate can be $\Theta(|V|)$ multiplicative factor smaller than meagerness. This shows that for some networks which exhibit some topological asymmetries, meagerness may become too loose and terribly fails to tightly bound such networks' coding rates[1].

---

[1]Recently, it was brought to our attention (see acknowledgment) that a similar result was obtained in [9]. However, a different network topology was used in the proof.

Let $\mathcal{N}_1$ be a directed k-pairs network with a set of commodities $\mathcal{I} = \{1, 2, \ldots, k\}$. The nodes of the underlying graph consist of k source nodes $s(1), \ldots, s(k)$, two intermediate nodes u, v, and k sink nodes $t(1), \ldots, t(k)$. The set of edges can be described as follows: There is an edge from every source node to the intermediate node u and there is an edge from v to every sink node. A single edge connects u to v. Finally, every sink node $t(i)$ has an incoming edge from $s(j)$ $\forall i < j$. Fig.1 shows network $\mathcal{N}_1$.
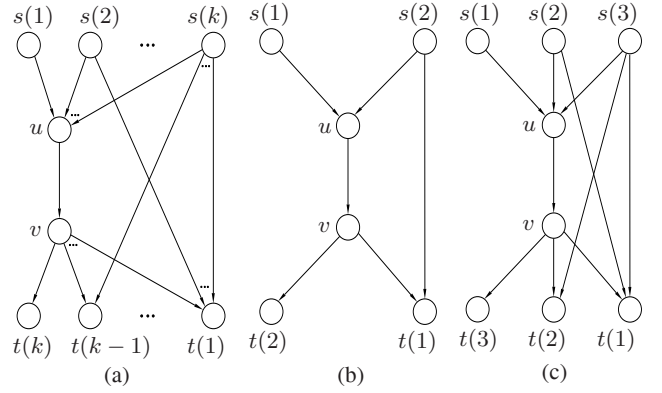

Fig. 1. Network $\mathcal{N}_1$: (a)-The network for any k, (b)- An instance of $\mathcal{N}_1$ with $k = 2$, (c)- An instance of $\mathcal{N}_1$ with $k = 3$

Lemma 1: The value of the most meager cut in $\mathcal{N}_1$ is 1.

Proof: For any set of commodities $\mathcal{J} \subseteq \mathcal{I}$ we determine the meagerness of the most meager set of edges that isolates $\mathcal{J}$. First note that $\forall i \in \mathcal{I}$ there exists a path from $s(i)$ to $t(i)$ passing through the edge $e = (u, v)$. Thus any isolating set A must contain e. Now consider the following two cases:

- If $|\mathcal{J}| = 1$, then $A = \{e\}$ and $\mathcal{M}(A) = 1$.
- If $|\mathcal{J}| \geq 2$. Let $\mathcal{J} = \{i_1, i_2, \ldots, i_{|\mathcal{J}|}\}$ where without loss of generality $1 \leq i_1 < i_2 < \ldots < i_{|\mathcal{J}|} \leq k$. Since a cut A must isolate all the commodities in $\mathcal{J}$, it must isolate $s(i_{|\mathcal{J}|})$ from all sinks $t(i_1), t(i_2), \ldots, t(i_{|\mathcal{J}|})$. From the structure of $\mathcal{N}_1$, there exists an edge from $s(i_{|\mathcal{J}|})$ to every sink node $t(i)$, $\forall i \in \mathcal{J} \backslash \{i_{|\mathcal{J}|}\}$. Let F be the set of such edges, then $|F| = |\mathcal{J}| - 1$ and for any isolating set, A, we must have $\{e\} \cup F \subseteq A$. Therefore, the capacity of any isolating set A is at least $|\mathcal{J}|$. Therefore, $\mathcal{M}(A) \geq 1$.

The lemma follows by noting that $\mathcal{M}_{\mathcal{N}_1} = \min_{A \subseteq E} \mathcal{M}(A) = 1$. ∎

Theorem 1: There exist unit capacity, directed acyclic k-pairs networks where the network coding rate is $\Theta(|V|)$ multiplicative factor smaller than meagerness.

Proof: Consider the network $\mathcal{N}_1$ with k sources as in Fig.1. Note that $t(k)$ must recover the message of $s(k)$, i.e. X_k, from the information carried by e. Similarly $t(k - 1)$ recovers X_{k-1} as a function of X_k and X_e and so on until $t(1)$ where X_1 is computed as a function of $X_2, \ldots, X_k$ and

X_e. Thus we have

$$t(k) \text{ gives: } H(X_k, X_e) \leq H(X_e) \tag{1}$$

$$t(k-1) \; : H(X_{k-1}, X_k, X_e) \leq H(X_k, X_e) \tag{2}$$

$$\vdots$$

$$t(2): H(X_2, X_3, \ldots, X_k, X_e) \leq H(X_3, X_4, \ldots, X_k, X_e) \tag{3}$$

$$t(1): H(X_1, X_2, \ldots, X_k, X_e) \leq H(X_2, X_3, \ldots, X_k, X_e) \tag{4}$$

Applying forward substitution on the previous set of inequalities we obtain

$$
\begin{aligned}
H(X_e) &\geq H(X_1, X_2, \ldots, X_k, X_e) & (5) \\
&\geq H(X_1, X_2, \ldots, X_k) & (6) \\
&= \sum_{i \in \mathcal{I}} H(X_i) & (7) \\
&\geq \sum_{i \in \mathcal{I}} r_i & (8) \\
&= rk & (9)
\end{aligned}
$$

where (6) follows since entropy is non-decreasing and (7) is due to the independence of sources. (8) and (9) follows from the definitions of rate and symmetric rate. Since edge e has unit capacity we have $H(X_e) \leq 1$. Thus, the network coding rate is upper bounded as

$$r \leq \frac{1}{k} \tag{10}$$

The theorem follows from (10) and Lemma 1 by noting that $k = \Theta(|V|)$ for $\mathcal{N}_1$. ∎

IV. UNDIRECTED NETWORKS

Undirected k-pairs networks where considered in [2] where it was conjectured that network coding can not provide any rate improvement over routing. Since sparsity, $\mathscr{S}$, is an upper bound of both routing and network coding rates, the conjecture trivially holds true if the routing rate is equal to $\mathscr{S}$. Hence, to verify the validity of the conjecture, one must consider networks whose routing rate is strictly less than their sparsity. Hereafter we refer to such networks as *gaped* networks. One such network that has been extensively considered is the Okamura-Seymour, OS, network [10]. The OS network is a 4-pairs undirected network with $|V| = 5$, $|E| = 6$ whose Weiner bound $W = 3/4$, sparsity $\mathscr{S} = 1$ and routing rate equals W. It is not hard to verify that the underlying graph, G, of the OS network exhibits the smallest number of vertices among all underlying graphs, non-isomorphic to G, of 4-pairs gaped networks (note that different networks might have the same unlabeled graph as their underlying graph). In [3] [4] [5] it was independently shown that the network coding rate of the OS network is indeed equal to the routing rate. Hence, moving one step toward the k-pairs conjecture.

Another class of networks for which the conjecture has been verified is the set of *special bipartite* networks [4]. A summary of networks for which the conjecture holds true (including the ones obtained in the next two subsections) is listed below. Note

that the classes in the list are not disjoint and might greatly intersect.

- k-pairs networks whose maximum achievable rates are equal to their sparsity. An undirect k-pairs network $\mathcal{N}$ is known to belong to this class of networks if
  - $\mathcal{N}$ has one commodity [11].
  - $\mathcal{N}$ has two commodities, i.e. $k = 2$, [12].
  - $\mathcal{N}$ has an underlying planar graph, G, that can be drawn such that all source and sink nodes lay on the outer face of G [10].
- $\mathcal{N}$ is the Okamura-Seymour network [3] [4] [5].
- $\mathcal{N}$ is a special bipartite network [4], [13].
- $\mathcal{N}$ is the three commodity network in figure 2.
- $\mathcal{N}$ is a Type-I bipartite network, corollary 1.
- $\mathcal{N}$ is a Type-II bipartite network, corollary 2.

A. A Three-Commodity Network

In this subsection we consider a three commodity network $\mathcal{N}_2$, Fig.2. It is known that routing can not achieve the sparsity (min-cut) of this network [12]. To see this, consider all possible cuts in the network. It can be seen that sparsity is $4/3$. But the Wiener bound asserts that the routing rate can not exceed $8/7$. It is easy to advise a routing scheme achieving rate $8/7$ for $\mathcal{N}_2$. In the following we show that network coding does not have any rate advantages over routing and thus confirm the k-pairs conjecture over this network. Network $\mathcal{N}_2$ was considered in [5] when all edges have capacity 2. The authors used an algorithm called *progressive d-separating edge-set* or PdE to show that network coding can not achieve the rate tuple $(r_a, r_b, r_g) = (1, 4, 2)$ in $\mathcal{N}_2$.
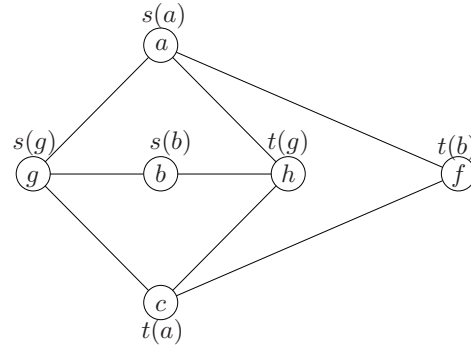

Fig. 2.   Network $\mathcal{N}_2$

Theorem 2: The network coding rate for $\mathcal{N}_2$ is $8/7$.

Proof: Applying the input-output inequality at node g we obtain

$$H(X_g, X_{ag}, X_{bg}, X_{cg}, X_{ga}, X_{gb}, X_{gc}) \leq H(X_g, X_{ag}, X_{bg}, X_{cg}) \tag{11}$$

and at node h we obtain

$$H(X_g, X_{ah}, X_{bh}, X_{ch}, X_{ha}, X_{hb}, X_{hc}) \leq H(X_{ah}, X_{bh}, X_{ch}) \tag{12}$$

1830

adding (11) and (12) and using submodularity in the LHS and the union bound in the RHS we obtain

$$H(X_g, X_{E'}) \le H(X_{ag}, X_{bg}, X_{cg}) + H(X_{ah}, X_{bh}, X_{ch}) \quad (13)$$

where $X_{E'} = X_{E_d} \backslash \{X_{af}, X_{fa}, X_{cf}, X_{fc}\}$ and E_d is the set of directed edges obtained from E. Now consider node b and note that X_b must be delivered through the directed edges bh and bg, i.e., X_b is a function of X_{bg} and X_{bh}. Since $X_{bg}, X_{bh} \in X_{E'}$, then from (13) we can write

$$H(X_b, X_g, X_{E'}) \le H(X_{ag}, X_{bg}, X_{cg}) + H(X_{ah}, X_{bh}, X_{ch}) \quad (14)$$

The input-output inequality at f gives $H(X_b, X_{af}, X_{cf}, X_{fa}, X_{fc}) \le H(X_{af}, X_{cf})$. Adding this to (14) and using submodularity at the LHS we obtain $H(X_b) + H(X_b, X_g, X_E) \le H(X_{af}, X_{cf}) + H(X_{ag}, X_{bg}, X_{cg}) + H(X_{ah}, X_{bh}, X_{ch})$. From node c, we know X_a must be recovered from $X_{gc}, X_{hc}, X_{fc} \in X_E$ and thus we can write

$$\begin{aligned} H(X_b) + H(X_a, X_b, X_g, X_E) &\le H(X_{ag}, X_{bg}, X_{cg}) \\ &+ H(X_{ah}, X_{bh}, X_{ch}) \\ &+ H(X_{af}, X_{cf}) \quad (15) \\ &\le H(X_{ag}) + H(X_{bg}) + H(X_{cg}) \\ &+ H(X_{ah}) + H(X_{bh}) + H(X_{ch}) \\ &+ H(X_{af}) + H(X_{cf}) \quad (16) \end{aligned}$$

Since the sources are independent and $H(X_a, X_b, X_g) \le H(X_a, X_b, X_g, X_E)$, (16) gives

$$\begin{aligned} H(X_a) + 2H(X_b) + H(X_g) &\le H(X_{ag}) + H(X_{bg}) + H(X_{cg}) \\ &+ H(X_{ah}) + H(X_{bh}) + H(X_{ch}) \\ &+ H(X_{af}) + H(X_{cf}) \quad (17) \end{aligned}$$

Now we apply the input-output inequality at node a and write

$$\begin{aligned} H(X_a, X_{ga}, X_{ha}, X_{fa}, X_{ag}, X_{ah}, X_{af}) \\ \le H(X_a, X_{ga}, X_{ha}, X_{fa}) \quad (18) \end{aligned}$$

and at node c we obtain

$$\begin{aligned} H(X_a, X_{gc}, X_{hc}, X_{fc}, X_{cg}, X_{ch}, X_{cf}) \\ \le H(X_{gc}, X_{hc}, X_{fc}) \quad (19) \end{aligned}$$

computing (18) + (19) and using submodularity we get

$$H(X_a, X_{E''}) \le H(X_{ga}, X_{ha}, X_{fa}) + H(X_{gc}, X_{hc}, X_{fc}) \quad (20)$$

where $X_{E''} = X_{E_d} \backslash \{X_{gb}, X_{bg}, X_{bh}, X_{hb}\}$. Since node f demands X_b, X_b must be a function of $X_{af}, X_{cf} \in X_{E''}$. Therefore,

$$H(X_a, X_b, X_{E''}) \le H(X_{ga}, X_{ha}, X_{fa}) + H(X_{gc}, X_{hc}, X_{fc}) \quad (21)$$

Applying input-output inequality at node b gives $H(X_b, X_{gb}, X_{hb}, X_{bg}, X_{bh}) \le H(X_b, X_{gb}, X_{hb})$. Adding this to (21) and using submodularity we get $H(X_a, X_b, X_E) \le H(X_{ga}, X_{ha}, X_{fa}) + H(X_{gc}, X_{hc}, X_{fc}) + H(X_{gb}, X_{hb})$.

Since X_g is recoverable from $X_{bh}, X_{ah}, X_{ch} \in X_E$, we can write

$$\begin{aligned} H(X_a, X_b, X_g, X_E) &\le H(X_{ga}, X_{ha}, X_{fa}) + H(X_{gc}, X_{hc}, X_{fc}) \\ &+ H(X_{gb}, X_{hb}) \quad (22) \end{aligned}$$

Thus, using independence of sources on the LHS and the union bound on the RHS we obtain

$$\begin{aligned} H(X_a) + H(X_b) + H(X_g) &\le H(X_{ga}) + H(X_{ha}) + H(X_{fa}) \\ &+ H(X_{gc}) + H(X_{hc}) + H(X_{fc}) \\ &+ H(X_{gb}) + H(X_{hb}) \quad (23) \end{aligned}$$

Adding (17) and (23) and noting that $H(X_{ij}) + H(X_{ji}) \le 1$ we obtain

$$2r_a + 3r_b + 2r_g \le 2H(X_a) + 3H(X_b) + 2H(X_g) \le 8 \quad (24)$$

The theorem follows by setting $r_a = r_b = r_g = r$ and noting that there exist a fractional routing scheme achieving rate $8/7$. ∎

B. Networks on Bipartite Graphs

In this subsection we consider undirected k-pairs networks with underlying bipartite graphs. Let $\mathcal{N}$ be a k-pairs undirected network with a set of commodities $\mathcal{I}$ and an underlying bipartite graph $G(V \cup W, E)$. This problem was considered in [4] for the case when each commodity $i \in \mathcal{I}$ is such that $s(i)$ and $t(i)$ are located in the same partition V or W. In this section we extend this study to any bipartite k-pairs network. Let $\mathcal{I}_{VV} = \{i \in S(V) : t(i) \in V\}$ be the set of all commodities whose sources and sinks are in V, also let $\mathcal{I}_{VW} = \{i \in S(V) : t(i) \in W\}$ be the set of all commodities whose sources are in V and sinks are in W. On the other hand, let $\mathcal{I}_{WW} = \{i \in S(W) : t(i) \in W\}$ be the set of all commodities whose sources and sinks are in W and $\mathcal{I}_{WV} = \{i \in S(W) : t(i) \in V\}$ be the set of commodities from W to V.

The following is a Lemma required in proving the next theorem. We present the lemma without a proof and refer the interested reader to [4] where a stronger result was proven.

Lemma 2: For any collection of sets $A_1, \ldots, A_n$

$$\sum_{i=1}^{n} H(X_{A_i}) \ge H(X_{\bigcup_{i=1}^{n} A_i}) + H(X_{\bigcup_{1 \le i < j \le n} A_i \cap A_j})$$

Theorem 3: For an undirected k-pairs network on a bipartite graph $G(V \cup W, E)$ with a set of commodities $\mathcal{I} = \mathcal{I}_{VV} \cup \mathcal{I}_{VW} \cup \mathcal{I}_{WV} \cup \mathcal{I}_{WW}$, the network coding rate is bounded as

$$r \le \frac{|E|}{(|\mathcal{I}_{VW}| + |\mathcal{I}_{WV}|) + 2(|\mathcal{I}_{VV}| + |\mathcal{I}_{WW}|)}$$

Proof: Applying the input-output inequality at each node $v \in V$ we can write

$$\begin{aligned} H(X_{S(v)}, X_{\text{In}(v)}, X_{\text{Out}(v)}, X_{T(v)}) &\le H(X_{S(v)}, X_{\text{In}(v)}) \quad (25) \\ &\le H(X_{S(v)}) + H(X_{\text{In}(v)}) \quad (26) \end{aligned}$$

adding (26) over all $v \in V$ we can write

$$H(X_{S(V)}, X_{\text{In}(V)}, X_{\text{Out}(V)}, X_{T(V)}) + H(X_{\mathcal{I}_{VV}})$$

$$\leq \sum_{v \in V} H(X_{S(v)}) + \sum_{v \in V} H(X_{\text{In}(v)}) \quad (27)$$

$$\leq \sum_{v \in V} H(X_{S(v)}) + \sum_{v \in V} \sum_{e \in \text{In}(v)} H(X_e) \quad (28)$$

$$= \sum_{v \in V} H(X_{S(v)}) + \sum_{e \in \text{In}(V)} H(X_e) \quad (29)$$

The LHS of (27) follows from Lemma 2 where the second term is obtained by noting that $S(v) \cap T(u) \subseteq \mathcal{I}_{VV}, \forall u, v \in V$ and $\bigcup_{\{u,v\} \subseteq V} S(u) \cap T(v) = \mathcal{I}_{VV}$. (28) follows from the union bound and (29) follows since a node $v \in V$ has no neighbors in V (from the definition of a bipartite). Since the underlying graph is bipartite we know that $\text{In}(V) \cup \text{Out}(V) = E_d$. Using this in the LHS of (29) and noting that all commodities must be communicated via E_d, we can write

$$H(X_{\mathcal{I}}) + H(X_{\mathcal{I}_{VV}}) \leq \sum_{v \in V} H(X_{S(v)}) + \sum_{e \in \text{In}(V)} H(X_e) \quad (30)$$

Using the same argument at the partition W, we obtain

$$H(X_{\mathcal{I}}) + H(X_{\mathcal{I}_{WW}}) \leq \sum_{w \in W} H(X_{S(w)}) + \sum_{e \in \text{In}(W)} H(X_e) \quad (31)$$

Adding (30) and (31) and noting that $\text{In}(W) = \text{Out}(V)$ we obtain

$$2H(X_{\mathcal{I}}) + H(X_{\mathcal{I}_{VV}}) + H(X_{\mathcal{I}_{WW}}) \leq \sum_{v \in V \cup W} H(X_{S(v)}) + |E| \quad (32)$$

$$= \sum_{i \in \mathcal{I}} H(X_i) + |E| \quad (33)$$

where in (32) we used the fact that $\sum_{e \in \text{In}(W)} H(X_e) + \sum_{e \in \text{Out}(W)} H(X_e) \leq \sum_{e \in E} 1 = |E|$. Using the independence of the sources on the LHS to get

$$\sum_{i \in \mathcal{I}} H(X_i) + \sum_{i \in \mathcal{I}_{VV}} H(X_i) + \sum_{i \in \mathcal{I}_{WW}} H(X_i) \leq |E| \quad (34)$$

Since $r_i \leq H(X_i)$, the previous result become

$$\sum_{i \in \mathcal{I}} r_i + \sum_{i \in \mathcal{I}_{VV}} r_i + \sum_{i \in \mathcal{I}_{WW}} r_i \leq |E| \quad (35)$$

and the symmetric rate can be bounded as

$$r \leq \frac{|E|}{|\mathcal{I}| + |\mathcal{I}_{VV}| + |\mathcal{I}_{WW}|} \quad (36)$$

The theorem follows by noting that $|\mathcal{I}| = |\mathcal{I}_{VV}| + |\mathcal{I}_{VW}| + |\mathcal{I}_{WV}| + |\mathcal{I}_{WW}|$. ∎

In the following we provide two subclasses of k-pairs bipartite networks for which the k-pairs conjecture holds true. Let $\mathcal{N}$ be a k-pairs bipartite network with a set of commodities $\mathcal{I}$ and an underlying *complete* bipartite graph $G(V \cup W, E)$, i.e, $G = K_{|V|, |W|}$. Then $\mathcal{N}$ is a Type-I k-pairs bipartite network if the following three conditions are satisfied

- For every unordered pair of distinct vertices $u, v \in V$ there exists a commodity $i \in \mathcal{I}$ such that $\{s(i), t(i)\} = \{u, v\}$.
- For every unordered pair of distinct vertices $u, v \in W$ there exists a commodity $i \in \mathcal{I}$ such that $\{s(i), u(i)\} = \{u, v\}$.
- There is no commodity $i \in \mathcal{I}$ such that $s(i) \in V$ and $t(i) \in W$ or $s(i) \in W$ and $t(i) \in V$.

On the other hand, $\mathcal{N}$ is a Type-II network if for every unordered pair of distinct vertices $u, v \in V \cup W$ there exists a commodity $i \in \mathcal{I}$ such that $\{s(i), u(i)\} = \{u, v\}$.

Corollary 1: The k-pairs conjecture holds for Type-I networks.

Proof: (sketch) The corollary follows by showing that there exists a fractional routing scheme achieving the rate r in theorem 3 with $|\mathcal{I}_{VW}| = |\mathcal{I}_{WV}| = 0$, $|\mathcal{I}_{VV}| = \binom{|V|}{2}, |\mathcal{I}_{WW}| = \binom{|W|}{2}$ and $|E| = |V|.|W|$. The proof of the existence of such routing scheme is omitted. ∎

Corollary 2: The k-pairs conjecture holds for Type-II networks.

Proof: (sketch) Similar to the proof of the previous corollary where in this case we have $|\mathcal{I}_{VW}| + |\mathcal{I}_{WV}| = |V|.|W|$, $|\mathcal{I}_{VV}| = \binom{|V|}{2}$ and $|\mathcal{I}_{WW}| = \binom{|W|}{2}$. ∎

V. ACKNOWLEDGMENT

REFERENCES

[1] R. Ahlswede, N. Cai, S.-Y. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. on Info. Theory.*, vol. 46, no. 4, pp. 1204–1216, July 2000.
[2] Z. Li and B. Li, "Network coding: The case of multiple unicast sessions," in *Proceedings of the 42nd Allerton Annual Conference on Communication, Control and Computing*, 2004.
[3] K. Jain, V. Vazirani, R. Yeung, and G. Yuval, "On the capacity of multipble unicast sessions in undirected graphs," in *ISIT*, 2005.
[4] N. J. A. Harvey, R. Kleinberg, and A. Rasala-Lehman, "On the capacity of information networks," *IEEE Trans. on Info. Theory*, vol. 52, no. 6, pp. 2345–2364, June 2006.
[5] G. Kramer and S. A. Savari, "Edge-cut bounds on network coding rates," *Journal of Network and Systems Management*, vol. 14, no. 1, pp. 49–67, 2006.
[6] F. Sharokhi and D. W. Matula, "The maximum concurrent flow problem," *Journal of the ACM*, vol. 37, no. 2, pp. 318–334, April 1990.
[7] A. Rasala-Lehman, "Network coding," Ph.D. dissertation, Department of Electrical Eng. and Computer Science, MIT, Cambridge, MA, 2005.
[8] H. Wiener, "Structural determination of paraffin boiling points," *Journal of the American Chemical Society*, vol. 69, no. 1, pp. 17–20, 1947.
[9] N. Harvey and R. Kleinberg, "Tighter cut-based bounds for k-pairs communication problems," in *43rd Allerton Conference on Communication, Control, and Computing*, Monticello, IL, Sept. 2005.
[10] H. Okamura and P. Seymour, "Multicommodity flows in planar graphs," *J. Combinat. Theory, Ser. B*, vol. 31, no. 1, pp. 75–81, 1981.
[11] L. Ford and D. Fulkerson, "Maximal flow through a network," *Canadian Journal of Mathematics*, vol. 8, no. 3, pp. 399–404, 1956.
[12] T. C. Hu, "Multi-commodity network flows," *Operations Research*, vol. 11, no. 3, pp. 344–360, 1963.
[13] F. Shahrokhi and D. Matula, "The maximum concurrent flow problem," *Journal of the Association for Computing Machinery*, vol. 37, no. 2, pp. 318–334, 1990.