MIT Open Access Articles

Network coded gossip with correlated data

Massachusetts Institute of Technology

Network Coded Gossip with Correlated Data

Bernhard Haeupler*
MIT
haeupler@mit.edu

Asaf Cohen*
Ben-Gurion University
coasaf@cse.bgu.ac.il

Chen Avin
Ben-Gurion University
avin@cse.bgu.ac.il

Muriel Médard
MIT
medard@mit.edu

Abstract—**We design and analyze gossip algorithms for networks with correlated data. In these networks, either the data to be distributed, the data already available at the nodes, or both, are correlated. Although coding schemes for correlated data have been studied extensively, the focus has been on characterizing the rate region in static memory-free networks. In a gossip-based scheme, however, nodes communicate among each other by continuously exchanging packets according to some underlying communication model. The main figure of merit in this setting is the stopping time – the time required until nodes can successfully decode. While Gossip schemes are practical, distributed and scalable, they have only been studied for uncorrelated data.**

We wish to close this gap by providing techniques to analyze network coded gossip in (dynamic) networks with correlated data. We give a clean framework for oblivious network models that applies to a multitude of network and communication scenarios, specify a general setting for distributed correlated data, and give tight bounds on the stopping times of network coded protocols in this wide range of scenarios.

I. INTRODUCTION

In this paper, we design and analyze information dissemination algorithms in communication networks with correlated data. In these networks, either the data to be distributed, the data already available at the nodes, or both, are correlated. This problem arises in a many networking applications, such as sensor, peer-to-peer or content distribution networks. One such example is a large set of distributed temperature sensors with a clock at the receiver. Both the temperatures at different sensor locations and the time at which a measurement is taken have high correlations among each other.

While the current information theory literature includes several coding schemes for correlated data, the focus in these works is mainly on characterizing the rate region - the set of achievable rates. On the other hand, recent work in the networking literature offers a multitude of efficient, decentralized and address-oblivious schemes for information dissemination (e.g., randomized gossip). Unfortunately these schemes treat the data as uncorrelated and neglect any available information at the receivers. The focus of this paper is to close this gap and give tools for analyzing gossip-based algorithms in networks with correlated data.

A. Related Work

Distributed Source Coding: Distributed compression has been studied in information theory mainly through small, canonical problems. In [1], Slepian and Wolf considered the

problem of separately encoding two correlated sources and joint decoding. In [2] and [3] the problem of compression with a rate-limited helper is considered. In [4], Ho et al. considered the multicast problem with correlated sources which can be viewed as extending the Slepian-Wolf problem to arbitrary networks through *network coding*. Further extensions appeared in [5] and [6]. In all these studies the goal is to characterize the *rate region* for (fixed) static and memory-free networks, that is, the set of required capacities needed for a multicast.

(Network Coded) Gossip: Gossip schemes were first introduced in [7] as a simple and decentralized way to disseminate a piece of information in a network. A detailed analysis of a class of these algorithms is given in [8]. In these schemes nodes communicate by continuously picking communication partners in a randomized fashion and then forwarding the information. The main figure of merit is the stopping time – the time needed for all nodes to be informed. Such randomized gossip-based protocols are attractive due to their locality, simplicity, and structure-free nature, and have been offered in the literature for various tasks. For the task of distributing multiple messages [9] introduced *algebraic gossip*, a network coding-based gossip protocol in which nodes exchange linear combinations of their available messages. This idea was extended to arbitrary networks in [10] and [11]. Haeupler [12] proved tight bounds for the stopping time of algebraic gossip for various models, including (adversarial) dynamically changing networks [13] and nodes with limited memory [14]. Improved bounds for non-uniform gossip were given in [15]. The projection analysis developed in [12] will play a key role in this paper.

B. Our Contributions

To our knowledge this paper is the first to combine these two strains of research and analyze gossip based protocols in networks with correlated data. Our contributions are manifold: First, we give a clean and general framework for oblivious network models in Section III and define a setting for a correlated environment in Section II. In this general setting we extend the projection analysis of [12] by making a connection between the coefficient vectors a node knows and the amount and type of information it has learned. This results in simple, direct and self-contained proofs of tight bounds on the stopping time in the canonical models of one source and side information at the receivers, as well as two correlated sources. In Section V we then give results for the general scenario of multiple sources and side information. We do this by providing tight bounds

on the time required for any set of (fractional) capacities to be induced by the (random) packet exchanges generated in an oblivious network model. This allows to transform results on the rate region of static memory-free networks (e.g., [4], [6]) into bounds on the stopping times of gossip-based algorithms. These capacity bounds are interesting in their own right and have the potential to be useful in other information dissemination problems.

II. NETWORK AND INFORMATION MODEL

In this section we state the broadcast problem, define the communication model and the information model which and state the information model defining the nature of the source and side information.

1) Network and Communication Model: For simplicity, we will assume that the network consists of a fixed set V of n nodes. Communication takes place in synchronous rounds. In each round, each node v decides on a packet p_v of s bits to be sent out (possibly using randomness). Given the current state of the network, the network model then specifies (possibly using randomness) which packet will be received by which node in the current round. This corresponds to a probability distribution over directed edge sets where a directed edge (v, u) means that the packet p_v is successfully received by node u. Nodes are assumed to have unlimited storage (for schemes with limited buffers see [14]). We denote the set of directed edges chosen for round t with E_t and call it the active edge set for time t.

2) Source and Side Information: We assume k messages are initially distributed over the network. The i-th message constitutes of l i.i.d. samples from the random variable X_i, namely, a vector x_i. The message vectors $x_1, \ldots, x_k$ are initially distributed to nodes (i.e., sources) such that each vector is given to at least one node. We also assume that each node $v \in V$ (or terminal) has some side information. The side information y_v at node $v \in V$ is drawn as l i.i.d. samples from Y_v. Note that the variables $\{X_i\}_{i=1}^k \cup \{Y_v\}_{v \in V}$ are arbitrarily correlated according to some known memoryless joint distribution. We are interested in the time when nodes are able to decode $x_1, \ldots, x_k$ based on their side information and the packets exchanged with other nodes.

3) The Encoding and Decoding Schemes: For a given field size q and slack $\delta > 0$ we assume throughout that nodes employ the following coding scheme: Prior to communication, source nodes perform random binning, that is, for every $1 \leq i \leq k$ each node receiving the message vector x_i applies the same random mapping into $2^{l(H(X_i)+\delta)}$ bins. The resulting bin indices (which are the same for every node initialized with x_i) are interpreted as vectors of length $h = \left\lceil \frac{l}{\log q}(H(X_i) + \delta) \right\rceil$ over the finite field $\mathcal{F}_q$. These vectors are then split into $\frac{h \log q}{s}$ blocks of $\frac{s}{\log q}$ symbols in $\mathcal{F}_q$ each, for a total of s bits per block.

During the communication phase nodes sends out random linear combinations (over F_q) of these blocks as packets. To keep track of the linear combination contained in a packet

one coefficient for each block of each message is introduced and sent in the header of each packet. As in all prior works on distributed network coding (e.g., [9], [10], [11], [12], [14], [15]), we assume that $\frac{s}{\log q}$ is sufficiently large compared to the number of coefficients. This renders the overhead of the header negligible leaving a packet size of s bits as desired.

At each node, independent linear equations on the blocks are collected for decoding. We denote with S_v the subspace spanned by all coefficient vectors received at node v. We also use the following notion of knowledge from [12]:

Definition 1. A node knows a coefficient vector μ iff S_v is not orthogonal to μ, that is, there exists a vector $c \in S_v$ such that $\langle c, \mu \rangle \neq 0$.

Lastly, we will make use of the following lemma on random binning.

Lemma 1 ([6], [3]). *Let $X \in \mathcal{X}$ and $Y \in \mathcal{Y}$ be two arbitrarily correlated random variables and let x, y be two vectors that are created by taking l i.i.d. samples from their joint distribution. Suppose, for some $\epsilon > 0$, all possible sequences in $\mathcal{X}^l$ are randomly and uniformly divided to at least $2^{l(H(X|Y)+\delta)}$ bins. Then joint typicality decoding correctly decodes x with high probability (as $l \to \infty$) using y and any $\lceil (H(X|Y) + \delta)l \rceil$ bits of information on the bin index of the bin in which the true x resides.*

In particular, Lemma 1 asserts that having access to the side information vector y, the message vector x can be decoded with high probability using any $\left\lceil \frac{l}{s}(H(X|Y) + \delta) \right\rceil$ *linearly independent equations* on the blocks describing the bin index of x.

III. OBLIVIOUS NETWORK MODELS

In this section we introduce the definition of an oblivious network model. This gives a clean and very general framework capturing a wide variety of communication and (dynamic) network settings. While this was already somewhat implicit in [12] we restrict ourself to networks without adaptive adversarial behavior. This greatly facilitates the much cleaner framework presented in this section.

Definition 2. A network model is *oblivious* if the active edge set E_t of time t only depends on t, E'_t for any $t' < t$ and some randomness. We call an oblivious network model furthermore i.i.d. if the active edge set E_t is sampled independently for every time t from a distribution of edge sets.

The following are examples of oblivious (and i.i.d.) network models: In the *(Uniform) Gossip Model* [10], [11], [12] one has an underlying (directed) graph G and in each round each node picks a random neighbor as a communication partner. A node then sends (PUSH) or requests (PULL) a packet from its partner or both (EXCHANGE). The *Rumor Mongering* [9] or *Random Phone Calls Model* [8] is a well-studied special case of this in which G is complete, that is, nodes pick a random node as a communication partner. It is easy to include more sophisticated features in an oblivious network model.

Random packet losses in wired networks, or characteristics of radio networks like half-duplex transmission, collisions, packet loss rates depending on SNR and more can be easily modeled by removing edges according to (randomized) rules. An interesting example of non-i.i.d. oblivious network models are *(edge-)markovian evolving graphs* [16].

For any oblivious network model M we can define a random flooding process $F(M, p, S)$. Informally, this process describes which nodes are informed over time if initially only nodes in S are informed and from there on every informed node informs all its communication partners (as specified by M). The only modification to this standard infectious process is the parameter p which adds an independent probability of p for each transmission to be overheard.

Definition 3. Let M be an oblivious network model, p be a probability of fault and $S \subseteq V$ be a starting set of nodes. We define the flooding process $F(M, p, S)$ to be the random process $S_1 \subseteq S_2 \subseteq \ldots$ that is characterized by $S_1 = S$ and for every time t we define S_{t+1} by taking each of the (directed) edges E_t specified by M for time t independently with probability $1 - p$ to obtain E_t' and then we set $S_{t+1} = \{v \in V \mid \exists u \in S_t : (u, v) \in E_t' \ \lor \ v = u\}$.

Note, that Definition 3 is only well defined if M is oblivious. Furthermore, F becomes a time-homogeneous Markov chain if M is an i.i.d. oblivious network model. Also, as long as for every time t the union over the edges in M from t to infinity is almost surely connected then F is absorbing with the unique absorbing state V. We say the flooding process F stops if it reaches this absorbing state and we denote the time this happens with the random variable S_F. The next definition pairs this flooding time with a throughput parameter α that corresponds to the exponent of the flooding process tail probability. The reason for this definition and its connection to the multi-message throughput behavior of network coding becomes apparent in the statement and proof of Theorem 1 below.

Definition 4. We say an oblivious network model M on a node set V floods in time T with throughput α if there exists a prime power q such that for every $v \in V$ and every $k > 0$ we have $P[S_{F(M, 1/q, \{v\})} \geq T + k] < q^{-\alpha k}$.

To give a few illustrating examples of flooding times we note that the random phone call network model on n nodes floods in $\Theta(\log n)$ time with constant throughput. The uniform gossip model on a connected degree bounded graph G floods in time $\Theta(D)$ and with constant throughput where D is the diameter of G. In many oblivious network models it is easy to give tight bounds on the flooding time and throughput.

With this framework for oblivious network models in place we can give a cleaner restatement of Theorem 3 in [12]. We also provide a sketch of the proof since we will later expand on the ideas used therein.

Theorem 1 (Theorem 3 of [12])**.** *Suppose M is an oblivious network model that floods in time T with throughput α. Then,*

for any k, random linear network coding in the network model M spreads k arbitrarily distributed messages to all nodes with probability $1 - \epsilon$ after $T' = T + \frac{1}{\alpha}(k + \log \epsilon^{-1})$ rounds.

Proof: The random linear network coding protocol we analyze will use the same field size that achieves the parameters T and α for M in Definition 4. We fix a coefficient vector $\mu \in F_q^k$. This vector is initially known to a non-empty subset S of nodes. It is easy to check that the probability that a node v does not know μ after it has received a packet from a node that knows μ is at most $1/q$. This implies that knowledge of μ spreads through the network exactly as the flooding process $F_{M, 1/q, S}$. Using the assumption, Definition 4 and the monotonicity of $S_{F(M, 1/q, S)}$ in S we get that the probability that a vector $\mu \in F_q^k$ is not known to all vectors after T' steps is at most $q^{-(k + \log \epsilon^{-1})}$. A union bound over all q^k vectors shows that with probability at least $1 - q^{-\log \epsilon^{-1}} \geq 1 - \epsilon$ all node will know about all vectors and it is easy to see that this implies that all nodes are able to decode the messages. ∎

IV. Simple, Direct Proofs for Tight Stopping Times

In this section we give a simple, direct derivation of tight stopping time bounds for gossip with one source and side information at the nodes and gossip with two correlated sources. Our two main results in this section are the following.

Theorem 2. *Suppose M is an oblivious network model that floods in time T with throughput α. We assume a single message x generated from X and side information y_v generated from Y_v at every node. Fix an error probability $\epsilon > 0$. Then, for any $\delta > 0$, for any large enough block length l and any packet size s, node v will correctly decode x with probability at least $1 - \epsilon$ after $T' = T + \frac{1}{\alpha}\left(\frac{l}{s}(H(X|Y_v) + \delta) + \log \epsilon^{-1} + 3\right)$ rounds.*

Theorem 3. *Suppose M is an oblivious network model that floods in time T with throughput α. We assume two messages x_1, x_2 are generated from X_1, X_2 and nodes have no side information. Fix an error probability $\epsilon > 0$. Then, for any $\delta > 0$ and for large enough l, with probability at least $1 - \epsilon$ every node will correctly decode x_1, x_2 after $T + \frac{1}{\alpha}\left(\left\lceil \frac{l}{s}(H(X_1, X_2) + 2\delta)\right\rceil + \log \epsilon^{-1} + 3\right)$ rounds.*

The idea for proving these theorems is to generalize the observation of [12] that the question of when a node can decode is equivalent to determining when this node knows (see Definition 1) enough coefficient vectors. The proof of Theorem 1 shows that flooding or spreading of knowledge of vectors can be analyzed using a union bound. This implies that only the number of vectors needed is of importance. In the case with *uncorrelated* sources and no side information essentially knowledge of all coefficient vectors is necessary. In the correlated scenario, however, we want to relate the number of vectors a node v needs to know to the conditional entropy $H(X|Y_v)$. Lemma 1 helps in this respect. It asserts that in order to decode, a node with side information Y does not necessarily need $i = \lceil(H(X|Y) + \delta)l\rceil$ specific bits, but rather, assuming joint typicality decoding, it requires only

any sufficient amount of information about the index of the bin in which x resides. This is achieved by *any* i/s packets containing independent equations on the bin index. We can thus focus on the knowledge a node is required to obtain in order for its coefficient matrix to have rank at least i/s.

Unfortunately it is possible that a node knows many vectors without having a large rank. In fact, upon reception of the first packet a node gets to know all but a $1/q$ fraction of all vectors. On the other hand, in order to prove faster stopping times we want to argue that the knowledge of only an exponentially small fraction of all vectors suffices for decoding. This is achieve by the following lemma which shows that indeed only q^l *specific coefficient vectors* suffice to guarantee that at least l independent coefficient vectors were received:

Lemma 2. *Let $\mathcal{V}$ be be a finite dimensional vector space over a finite field $\mathcal{F}_q$. For every $0 \le h < \dim \mathcal{V}$ there exist $w = q^h + 1$ vectors $v_1, \ldots, v_w \in \mathcal{V}$ such that for any (subspace) $K \subset \mathcal{V}$ for which $K^\perp$ does not contain v_i for any i has dimension at least $h + 1$.*

It is now possible to prove the two main results of this section. Their proofs are self-contained and involve only random binning (Lemma 1) and Lemma 2:

Proof Sketch of Theorem 2: We use the field size q that achieves the parameters T and α in Definition 4. We furthermore choose l large enough so that the decoding probability in Lemma 1 is at most $\epsilon/2$. By Lemma 1, any node v with access to the side information vector y_v and $\left\lceil \frac{l}{s}(H(X|Y_v) + \delta) \right\rceil$ independent equations on the blocks describing the bin index of x assigned by the random binning procedure, can decode x with probability at least $1 - \epsilon/2$. It thus remains to show that with probability $1 - \epsilon/2$ we have $\dim S_v \ge \left\lceil \frac{l}{s}(H(X|Y_v) + \delta) \right\rceil$ after T' rounds. To prove this we apply Lemma 2 and get that there exists a set Z of $2^{\left\lceil \frac{l}{s}(H(X|Y_v) + \delta) \right\rceil}$ coefficient vectors such that if v has knowledge of these vectors, it indeed has sufficiently many independent equations. Furthermore, we refer to the proof of Theorem 1 for the fact that knowledge of any coefficient vector (in Z) spreads through the network like a flooding process. As before we thus get the fact that in the assumed network model M the probability that any of the coefficient vectors (in Z) is not known after $T + \frac{1}{\alpha}(k + (\log \epsilon^{-1} + 1))$ rounds is smaller than $\epsilon/2 \cdot q^{-k}$. Setting $k = \log |Z|$ and using a union bound over all coefficient vectors in Z we get as a result that indeed after T' rounds the probability that v has received sufficiently many independent coefficient vectors is at least $1 - \epsilon/2$. ∎

While the proof of Theorem 3 is similar in nature to that of Theorem 2, there is delicate point when considering multiple sources. In a single source scenario, for each terminal there is only one equation governing the rate. That is, $r \ge H(X|Y_v) + \delta$. Using Lemma 2, this rate constraint is translated into a rank constraint, namely, $dim(S_v) \ge \left\lceil \frac{l}{s}(H(X|Y_v) + \delta) \right\rceil$. For more than one source, however, the rate region is given by multiple rate constraints, and one has to make sure all are satisfied. Indeed, for two sources and no side information at the nodes this can be done using a single rank constraint. For more than

two sources, or when additional side information is available, a more refined analysis is required. This is the subject of the next section.

V. CHARACTERIZING CAPACITIES IN OBLIVIOUS NETWORK MODELS

To date, analysis of gossip schemes focused only on the dissemination time - the number of rounds required to gain the complete knowledge in the network. However, especially when dynamic networks are analyzed, it is interesting to gain a more accurate measure of the *actual capacities achievable between sets of nodes*. Namely, to analyze the capacities induced by the packet exchange process in algebraic gossip. This is an interesting question in its own right, and, in particular, can give a "black-box" method to transfer any results of prior works that bound the rates or capacities needed between sources and sinks in the static memory-free setting to stopping times in oblivious network models.

Herein, we develop such a characterization, and apply it to the results of [4] and [6] to obtain stopping times for gossip protocols with an arbitrary number of correlated sources and side-information, generalizing the results from the last section. We first introduce the required notation.

Definition 5. Let $T > 0$, node set V and active edges $E_1, E_2, \ldots, E_T$ be given. We define a *path* P from s to d to be a sequence of nodes $P = (v_0, v_1, \ldots, v_T)$ such that $v_0 = s$, $v_T = d$ and for every $t \le T$ we have $v_{t-1} = v_t$ or $(v_{t-1}, v_t) \in E_t$. We furthermore define a set of m paths $P_1, \ldots, P_M$ with weights $w_1, \ldots, w_m$ to be *valid* if for every $t < T$ and every $(u, v) \in E_t$ the weights of paths using (u, v) sum to at most one. Lastly, we say a set of valid weighted paths *achieves a capacity* of c between two nodes s and d if the weights of paths from s to d sum up to c.

Quite intuitively these paths correspond to an information flow through the network from the sources to the sink. This intuition can be made formal and one can give an explicit equivalence between the algebraic gossip protocol and random linear network coding in the classical memory-free model (e.g., [4]). This was done in [17] which also describes the hypergraph G_{PNC} that corresponds to a sequence of active edges. We omit the details of this equivalence and instead only recall the facts needed in this paper:

Lemma 3. *Let node set V, the active edges $E_1, \ldots, E_T$, destination $d \in V$ and sources $s_1, s_2, \ldots, s_k \in V$ be given. Algebraic gossip on $\{E_i\}_{i=1}^T$ is equivalent to classical random linear network coding in the transformed hypergraph G_{PNC} described in [17]. In particular, if for some integers $c_1, \ldots, c_k$, it is possible for every s_i to transmit c_i packets to d, then there exists a sequence of valid paths of weight one and a rate c_i between s_i and d. Conversely, if for some positive reals $c_1, \ldots, c_k$ there is a set of valid weighted paths that achieve a capacity c_i between s_i and d, then the capacities c_i lie in the min-cut region of G_{PNC}.*

Given this setup we show the first result in this direction:

Lemma 4. *Let M be a network model on a node set V that floods in time T with throughput α. For any T', any $\epsilon > 0$, any destination $d \in V$ and any set of k source nodes $s_1, s_2, \ldots, s_k \in V$ with integral capacities $c_1, c_2, \ldots, c_k \geq 1$ suppose $E_1, \ldots, E_{T'}$ is a sequence of active edges on V sampled from M. If $T' \geq T + \frac{1}{\alpha}(\sum_i c_i + \log \epsilon^{-1})$ then with probability at least $1 - \epsilon$ there exists a selection of valid weighted paths that achieve the capacity c_i between s_i and t for every i.*

Proof: We think of putting c_i messages at node s_i and run the standard algebraic gossip protocol for T' rounds using the field size q that achieves the parameters T and α on M. Theorem 1 now shows that with probability $1 - \epsilon$ the sink t can decode the messages. According to Lemma 3 this shows that there are c_i mutually disjoint paths from s_i to d for every i with weight one which achieve the desired capacities. ∎

Note that the above lemma requires the capacities to be integral and thus essentially asks for the time until a certain number of mutually disjoint paths occur. While this is sufficient and optimal in the *uncorrelated* information spreading setting this requirement can be a severe restriction.

One setting where this makes a drastic difference is when we have k sources and the total capacities needed sum up to less then one. This corresponds to asking for the time until there is a path from each of the sources to the sink – without these paths having to be disjoint. If one considers for example the random phone call model with n nodes and k sources it takes in expectation $\log n + k$ time until a disjoint path between a node and each source appears while merely $\log n + \log k$ rounds are sufficient to get this for non-disjoint paths.

The following lemma generalizes this observation and strengthens Lemma 4 in this direction to give order optimal bounds for any set of fractional capacities:

Lemma 5. *Let M be a network model on a node set V that floods in time T with throughput α. For any T', any $\epsilon > 0$, any sink $d \in V$ and any set of k source nodes $s_1, s_2, \ldots, s_k \in V$ with rates $c_1, c_2, \ldots, c_k > 0$, if $T' \geq T + \frac{1}{\alpha}(\lceil \sum_i c_i \rceil + \log k + \log \epsilon^{-1})$ then with probability at least $1 - \epsilon$ there exists a selection of valid weighted paths that achieve a capacity of c_i between s_i and d for every i.*

Proof: The idea is to combine k applications of Lemma 4 using a union bound and capacity sharing. We set the failure probability to ϵ/k and in the ith application of Lemma 4 we set the c_i to $\lceil \sum_i c_i \rceil$ while setting all other capacities to zero. Out of this we get that for every i with probability $1 - \epsilon/k$ the number of disjoint paths from s_i to d is at least $\lceil \sum_i c_i \rceil$. Via a union bound we get that with probability of $1 - \epsilon$ all these paths are there. Note, that while the paths from each source are disjoint the paths starting at different sources may not be disjoint. We now take the union of these paths while choosing the weight of a paths starting at source s_i to be $\frac{c_i}{\lceil \sum_j c_j \rceil}$. This gives capacity of c_i between s_i and d. Furthermore, since any edge e is used by at most one path going out from each source, we get that the total weight on e summed over all paths is at most $\sum_i \frac{c_i}{\lceil \sum_j c_j \rceil} \leq 1$. ∎

We will use Lemma 5 to prove our main result about information dissemination with correlated data in oblivious networks. To state our result we need the following definition:

Definition 6 (Slepian-Wolf region [1])**.** A capacity vector $c = (c_1, \ldots, c_k)$ is sufficient for $v \in V$ if and only if for every index subset $S \in [k]$ we have $\sum_{i \in S} c_i \geq H(X_S | X_{\overline{S}}, Y_v)$.

Putting together Lemma 5, Lemma 3 and applying the results on network coding with correlated data from [4] and [6] in a black-box manner, we can now directly state our main result which generalizes and encompasses Theorem 3 and Theorem 2:

Theorem 4. *Suppose M is an oblivious network model that allows floods in time T with throughput α. Then, for any $\delta > 0$ and error probability $\epsilon > 0$, there exists an l such that for any joint distribution of $X_1, \ldots, X_k$ and the Y_v's, any packet size $s > 0$, any node v and any capacity vector $(c_1, \ldots, c_k)$ that is sufficient for v, node v will correctly decoding $x_1, \ldots, x_k$ with probability at least $1 - \epsilon$ after $T + \frac{1}{\alpha}(\lceil \frac{l}{s} \sum_{i \in [k]} c_i + \delta \rceil + \log k + \log \epsilon^{-1} + \delta)$ rounds.*

References

[1] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inf. Theory*, vol. 19, no. 4, pp. 471–480, 1973.

[2] R. Ahlswede and J. Körner, "Source coding with side information and a converse for degraded broadcast channel," *IEEE Trans. Inf. Theory*, vol. 21, no. 6, pp. 629–637, 1975.

[3] A. Cohen, S. Avestimehr, and M. Effros, "On networks with side information," in *ISIT*, 2009, pp. 1343–1347.

[4] T. Ho, M. Médard, R. Koetter, D. R. Karger, M. Effros, J. Shi, and B. Leong, "A random linear network coding approach to multicast," *IEEE Trans. Inf. Theory*, vol. 52, no. 10, pp. 4413–4430, 2006.

[5] J. Barros and S. D. Servetto, "Network information flow with correlated sources," *IEEE Trans. Inf. Theory*, vol. 52, no. 1, pp. 155–170, 2006.

[6] M. Bakshi and M. Effros, "On achievable rates for multicast in the presence of side information," in *ISIT*, 2008, pp. 1661–1665.

[7] A. Demers, D. Greene, C. Hauser, W. Irish, J. Larson, S. Shenker, H. Sturgis, D. Swinehart, and D. Terry, "Epidemic algorithms for replicated database maintenance," in *PODC*, 1987, pp. 1–12.

[8] R. Karp, C. Schindelhauer, S. Shenker, and B. Vocking, "Randomized rumor spreading," in *FOCS*, 2000, vol. 41, pp. 565–574.

[9] S. Deb, M. Médard, and C. Choute, "Algebraic gossip: a network coding approach to optimal multiple rumor mongering," *Trans. Networking*, vol. 14, pp. 2486–2507, 2006.

[10] D. Mosk-Aoyamam and D. Shah, "Information dissemination via network coding," in *ISIT*, 2006, pp. 1748–1752.

[11] M. Borokhovich, C. Avin, and Z. Lotker, "Tight bounds for algebraic gossip on graphs," in *ISIT*, 2010, pp. 1758–1762.

[12] B. Haeupler, "Analyzing network coding gossip made easy," in *STOC*, 2011, pp. 293–302.

[13] B. Haeupler and D. Karger, "Faster information dissemination in dynamic networks via network coding," in *PODC*, 2011, pp. 381–390.

[14] B. Haeupler and M. Medard, "One packet suffices - highly efficient packetized network coding with finite memory," in *ISIT*, 2011, pp. 1151–1155.

[15] C. Avin, M. Borokhovich, K. Censor-Hillel, and Z. Lotker, "Order optimal information spreading using algebraic gossip," in *PODC*, 2011, pp. 363–372.

[16] A. Clementi, C. Macci, A. Monti, F. Pasquale, and R. Silvestri, "Flooding time in edge-markovian dynamic graphs," in *PODC*, 2008, pp. 213–222.

[17] B. Haeupler, M. Kim, and M. Medard, "Optimality of Network Coding with Buffers," in *ITW*, 2011, pp. 533–537.