Algebraic gossip on Arbitrary Networks

Dinkar Vasudevan and Shrinivas Kudekar
School of Computer and Communication Sciences, EPFL, Lausanne.
Email: \{dinkar.vasudevan, shrinivas.kudekar\}@epfl.ch

Abstract—Consider a network of nodes where each node has a message to communicate to all other nodes. For this communication problem, we analyze a gossip based protocol where coded messages are exchanged. This problem was studied in [4] where a bound to the dissemination time based on the spectral properties of the underlying communication graph is provided. Our contribution is a uniform bound that holds for arbitrary networks.

I. INTRODUCTION

There are \(n\) nodes in a network, each having an information. The goal is to disseminate the information at every node to every other node in the network. We assume that the network is composed of noiseless links and that the nodes are restricted to communicating with their neighbours defined by the network. Our goal is to characterize the dissemination time for protocols based on pairwise communication between nodes.

The problem formulation is motivated by potential applications in sensor networks and databases and is not new. Deb et al [3] considered the problem where \(k\) messages are disseminated in a network of \(N\) nodes. However, their analysis is restricted to complete graphs. The problem when all the nodes have information to communicate has been considered by Aoyama and Shah [4] for arbitrary networks. They refer to this problem as the information dissemination problem and report bounds to the dissemination time based on the spectral properties of the underlying graph. We study the same problem and report uniform bounds to the dissemination time, which depend only on the number of nodes and not on the topology of the network. (See discussion in Section IV for a comparison of results)

The paper is organized as follows – The Section II goes through notation and preliminaries, Section III provides the main result and Section IV puts the result in the context of existing results on the problem.

II. PRELIMINARIES

The network is represented by graph \(G = (V, E)\) where \(V\) represents the set of nodes and \(E\) represents the set of edges in the network. Let \(N\) be the number of nodes in the graph. The only restriction on the graph \(G\) is that we require it to be connected.

A. Asynchronous Time model

We consider an asynchronous time model where we assume there is a global clock that ticks according to a Poisson process with mean \(N\). At each global clock tick, a node \(i\) in \(G\) is picked uniformly at random for communication. Then the node \(i\) picks a neighbour \(j\) independently and uniformly at random. The probability of selecting the node \(j\) is denoted by \(P_{ij} = \frac{1}{N}\), where \(d_i\) is the degree of node \(i\). The pair of nodes then exchange messages (described in the following). Thus, we consider here a push and pull based protocol.

B. Data Transmission – Random Linear Network Coding

The messages exchanged during the communication are linear network codewords (see [11]). Each message is a vector of \(r\) symbols with each symbol an element of a finite field \(\mathbb{F}_q\) of size \(q\). Let \(\mathbb{F}_q^r\) denote the vector space of messages. Let us denote the initial messages at each node by \(m_i \in \mathbb{F}_q^r\). Assume that all the initial messages are linearly independent vectors in \(\mathbb{F}_q^r\). At any time instant \(t\), each node \(i\) has a set of coded messages

\[
\{f_i| f_i = \sum_{k=1}^{N} \alpha_{i{k}} m_k\}
\]

along with the set of coefficients \(\{\alpha_{i{k}}\}, 1 \leq k \leq N\). The set of coded messages at a node span a subspace of the vector space \(\mathbb{F}_q^r\). For node \(i\) at a time \(t\), we denote this subspace by \(S(i, t)\). When the node \(i\) is picked for communication, it randomly picks a vector from its subspace \(S(i, t)\) and transmits the generated vector along with its coefficients (with respect to the basis \(\{m_k\}_{k=1}^{N}\) to node \(j\). More precisely let,

\[
g_{ij} = \sum_{l} \beta_l f_l = \sum_{k=1}^{N} m_k \left(\sum_{l} \alpha_{i{k}} \beta_l\right)
\]

Then node \(i\) transmits to node \(j\) the vector \(g_{ij}\) along with the set of coefficients \(\left(\sum \alpha_{i{k}} \beta_l\right) \in \mathbb{F}_q\) for \(1 \leq k \leq N\). On receiving this information from \(i\), the node \(j\) does likewise. It sends to node \(i\) a randomly generated vector in its span \(S(j, t)\) along with the coefficients. We note here that the overhead of communicating coefficients, which is \(N\) symbols in \(\mathbb{F}_q\) is amortized by choosing \(r\) large enough.

C. Notation

We define the various quantities which appear in our result.

Definition 1 (Stopping Time): The stopping time \(T\) is defined as

\[
T = \min \{\dim(S(i, t)) = N \text{ for all } i \in G\}
\]

At time \(T\), each node \(i\) has message space of dimension \(N\) and hence can recover all the initial messages \(m_k, 1 \leq k \leq N\).
**Definition 2 (δ-information spreading time):** The δ-information spreading time $T_{\delta}$ is defined as the minimum time $t$ such that

$$\Pr\{\dim(S(i,t)) = N \text{ for all } i \in G\} \geq 1 - \delta$$

We introduce the notion of the predecessor graph $W(i,t)$ of node $i$ at time $t$. This graph contains nodes whose messages potentially influence the messages received at node $i$ by time $t$. The graph is constructed in levels – the first level contains the nodes which exchange messages directly with $i$, the second level contains nodes which exchange messages directly with the level one nodes and so on. A precise definition now follows (it is not required for understanding the proof and can be skipped on a first read.)

**Definition 3 (Predecessor Graph):** Define $W^{(0)}(i,t)$ as the node $i$ itself. Let $W^{(1)}(i,t)$ denote the set of nodes which have contacted or been contacted by $i$ until the time $t$. The superscript (1) indicates that these are nodes at “level one”. Let $\tau^{(1)}(l)$ denote the most recent time (preceding $t$) at which the node $i \in W^{(1)}(i,t)$ contacts $i$. We make similar definitions for $W^{(D)}(i,t)$ and $\tau^{(D)}(l)$ for $D \geq 2$, i.e., for level two nodes, level three nodes etc. in the following. For $D \geq 2$ define

$$W^{(D)}(i,t) = \bigcup_{l \in W^{(D-1)}(i,t)} W^{(1)}(l, \tau^{(D-1)}(l))$$

and for $k \in W^{(D)}(i,t)$,

$$\tau^{(D)}(k) = \max_{u \in W^{(D-1)}(i,t)} \begin{cases} \text{Most recent time preceding} \\ \tau^{(D-1)}(u) \text{ when } k \text{ is in contact with } u. \end{cases}$$

We finally define the set of nodes in the predecessor graph $W(i,t)$ as

$$W(i,t) = \bigcup_{D \geq 0} W^{(D)}(i,t)$$

We now define quantities related to the predecessor graph which we use in the proof. Let $R(i,t)$ be a set of all distinct nodes in $W(i,t)$ and $S(R(i,t))$ denote the space spanned by the messages of all the nodes in $R(i,t)$. It holds that $S(i,t) \subseteq S(R(i,t))$.

**III. MAIN RESULT**

**Theorem 1:** If $q > N^5$, the expected stopping time $E[T]$ for the information dissemination problem for any network satisfies

$$E[T] \leq \frac{q}{q - 1} 29N \log N \tag{1}$$

**Sketch of Proof:** We show that after $24N^2 \log N$ global clock ticks (on an average), if nodes $i$ and $j$ have the same message space, then with high probability both the nodes have recovered all messages. We then show that the probability of dimension increase beyond this time is lower bound by the fraction of nodes with message space dimension less than $N$. The expected number of global clock ticks to completion beyond $24N^2 \log N$ is upper bound by $\frac{q}{q - 1} 5N^2 \log N$.

**Proof:** Let $p(t)$ be the probability of strict increase of the total dimension ($\sum_{i \in G} \dim(S(i,t))$) at clock tick $t$.

$$p(t) \geq \frac{1}{N} \sum_{i} \sum_{j} \Pr(S(i,t) \neq S(j,t))(1 - \frac{1}{q})$$

$$= \frac{1}{N} \sum_{i,j} (1 - \Pr(S(i,t) = S(j,t))(1 - \frac{1}{q}) \tag{2}$$

The above follows since the dimension strictly increases if two nodes with nonidentical subspaces exchange messages and at least one of the messages does not lie in $S(i) \cap S(j)$. The latter probability is lower bound by the factor $(1 - \frac{1}{t})$.

We provide an upper bound [4] to $\Pr(S(i,t) = S(j,t))$. This is shown through the following two lemmas. In Lemma [1] we show that if $S(i,t) = S(j,t)$ then it holds with high probability (provided that the field size is large enough) that $S(i,t) = S(R(i,t))$, i.e., the node $i$ can decode messages of all the nodes in its predecessor graph.

Let us define the event $A \triangleq \bigcap_{u \in R(i,t)} \{S(j,t) \geq m_u\}$. In words, $A$ denotes the event that the message of every node in $R(i,t)$ is contained in $S(j,t)$.

**Lemma 1:**

$$\Pr(S(i,t) = S(j,t)) \leq \Pr(S(i,t) = S(j,t), A) + \frac{N^2}{q}$$

The proof is in the Appendix. The second lemma uses results from [5] which show that for $t$ large enough, the predecessor graph of node $i$ contains all nodes in $G$ with high probability. Combined with the previous lemma, this implies that for such a $t$ if $S(i,t) = S(j,t)$, then node $i$ can decode w.h.p the messages of all the nodes in $G$. Let $\gamma(t)$ be the fraction of nodes in the graph which have decoded all the initial messages after $t$ number of clock ticks.

**Lemma 2:** For $t > 24N^2 \log N$,

$$\Pr(S(i,t) = S(j,t), A) \leq \gamma(t) + \frac{1}{N^3}$$

where $\gamma(t)$ is the fraction of nodes $i$ such that $\dim(S(i,t)) = n$.

The proof is in the Appendix. From Lemmas [1] and [2] it follows that for $t > 24N^2 \log N$

$$\Pr(S(i,t) = S(j,t)) \leq \gamma(t) + \frac{1}{N^3} + \frac{N^2}{q} \tag{3}$$

Let $\tau_k$ denote the number of clock ticks till the total dimension increases by $k(k - 1)$. Clearly, $\tau_N$ corresponds to the number of clock ticks when all nodes can decode all the messages. Let $t \in (\tau_m, \tau_{m+1})$. For all such $t$ we have $\gamma(t) \leq \gamma(\tau_{m+1})$. Since the dimension increase till $\tau_{m+1}$ is, by definition, given by $(m + 1)m$, we have

$$\gamma(t) \leq \gamma(\tau_{m+1}) = \frac{(m + 1)m + N}{N^2} \leq \frac{(m + 1)^2}{N^2} + \frac{1}{N}$$

Therefore

$$\Pr(S(i,t) = S(j,t)) \leq \frac{(m + 1)^2}{N^2} + \frac{1}{N} + \frac{1}{N^3} + \frac{N^2}{q}$$
In the Appendix we compute \( Z \) in terms of number of clock ticks. Let \( \tau \) be described by

\[
\tau = \sum_{m=1}^{k-1} \sum_{i=1}^{2m} \chi_{im}
\]

where \( \chi_{im} \) corresponds to the number of clock ticks required to increase the total dimension by 1 in the interval \( (\tau_m, \tau_{m+1}] \).

In the Appendix we compute \( E[\tau_k] \), under the assumption that \( p(t) \) is lower bound as in [3] (which is true for \( t > 24 N^2 \log N \)), and show it is upper bound by \( \frac{q}{N-1} 5N^2 \log N \). This implies that \( E[T_{\text{ticks}}] \leq 24 N^2 \log N + \frac{q}{N-1} 5N^2 \log N \leq \frac{q}{N-1} 29N^2 \log N \), where \( T_{\text{ticks}} \) is the expected stopping time in terms of number of clock ticks. Let \( Z_i \) denote the random time between the \( i \)th and \( i+1 \)th global clock ticks. The \( Z_i \)’s are i.i.d exponentials of rate \( N \). It follows that

\[
E[T] = E \left[ \sum_{i=1}^{T_{\text{ticks}}} Z_i \right]
= E[T_{\text{ticks}}] \sum_{i=1}^{T_{\text{ticks}}} E[Z_i]
= E[T_{\text{ticks}}] \frac{q}{N-1} 29N \log N
\]

\( \blacksquare \)

**Corollary 1:** The \( \delta \)-information spreading time is given by

\[
T_\delta = \frac{q}{N-1} 58N \log N \log_2 \frac{1}{\delta}
\]

By Markov’s inequality, it follows that

\[
\Pr(T \geq \frac{q}{N-1} 58N \log N) \leq \frac{1}{2}
\]

Over \( \log_2 \frac{1}{\delta} \) epochs, each of time \( \frac{q}{N-1} 58N \log N \), it follows that the probability of failure of information dissemination is bounded to within \( \delta \). If \( \delta = \frac{1}{N^2} \), then we have \( T_{\frac{1}{N^2}} = O(N^2 \log^2 N) \).

### IV. Discussion

Deb et. al [3] considered the information dissemination problem (where there are, in general, \( k \leq N \) messages to be disseminated) on a complete graph. They show that the \( T_{\frac{1}{N^2}} \)-information spreading time \( T_{\frac{1}{N^2}} \) is \( O(N) \). This is a stronger result than what we obtain, which is \( O(N \log^2 N) \). However, their analysis is specific to complete graphs and is not extensible to arbitrary graphs.

In [4], the authors compute the expected stopping time \( E[T] \) and the \( T_{\frac{1}{N^2}} \)-information spreading time \( T_{\frac{1}{N^2}} \) for a pull based protocol and show it to be bounded as \( O(\frac{N}{k}) \) where

\[
\hat{\mu} = \sum_{k=1}^{N-1} \frac{k}{\Phi^k}
\]

where

\[
\Phi^k = \min_{S \subseteq V, |S| \leq k} \frac{\sum_{i \in S, j \in S^c} P_{ij}}{|S|}
\]

The quantity \( \Phi^k \) is the \( k \)-conductance of the network. Thus their result is \( E[T] \leq O(\frac{N}{k}) \), \( T_{\frac{1}{N^2}} \leq O(\frac{N}{k}) \). We refer to this bound as the conductance bound. For a complete graph and a ring graph, the conductance bound computes to \( O(N \log N) \) and \( O(N^2) \) respectively [4]. We compute the conductance bound for yet another class of graphs.

Consider a Lollipop-network (see Figure. 1), which consists of a complete graph on \( k \) nodes, denoted by \( K_k \), which is attached a string, denoted by \( C \), of \( 2N \) nodes.

For this graph and for any \( k \) it holds that

\[
\Phi^k \leq \frac{1}{k}
\]

The above holds since for any \( k \), the set \( S \) with minimum conductance can be constructed so that there is only one outgoing edge. Therefore

\[
\hat{\mu} = \sum_{k=1}^{N-1} \frac{k}{\Phi^k} = \frac{N-1}{k^2} = O(N^3)
\]

Thus for the lollipop graph, the conductance bound is \( O(N^2) \).

In contrast to conductance bounds which vary depending on the network, our result states that the expected stopping time \( T \) and the \( T_{\frac{1}{N^2}} \)-information dissemination time \( T_{\frac{1}{N^2}} \) for any network is bounded as \( O(N \log N) \) and \( O(N \log^2 N) \) respectively. A comparison of the bounds with [4] is illustrated in Table I. Our bounds for \( E[T] \) scales the same or better than the conductance bound. Our bounds for \( T_{\frac{1}{N^2}} \) scales worse for the complete graph, but scales better for the ring and lollipop graphs as compared to the conductance bound. We note here

**Table I**

| Graph       | \( E[T] \) | \( T_{\frac{1}{N^2}} \) [4] | \( E[T] \) [4] | \( T_{\frac{1}{N^2}} \) [4] |
|-------------|------------|----------------------------|-----------------|-----------------|
| Complete    | \( O(N \log N) \) | \( O(N \log^2 N) \) | \( O(N \log N) \) | \( O(N \log N) \) |
| Ring        | \( O(N \log N) \) | \( O(N \log^2 N) \) | \( O(N^2) \) | \( O(N^2) \) |
| Lollipop    | \( O(N \log N) \) | \( O(N \log^2 N) \) | \( O(N^2) \) | \( O(N^2) \) |

**Fig. 1.** Lollipop network

\[
\frac{m + 1}{N^2} + \frac{1}{N} + \frac{2}{N^3}
\]

where the last inequality follows by choosing \( q > N^5 \).

Substituting in we have that for \( t \in (\tau_m, \tau_{m+1}] \),

\[
p(t) = \left( 1 - \frac{(m + 1)^2}{N^2} - \frac{1}{N} - \frac{2}{N^3} \right) (1 - \frac{1}{q})
\]

\( \blacksquare \)
that worsening of our bound for the complete graph is because we are not able to prove a concentration style result (which says that $T_{\uparrow}$ is of the same order as $E[T]$), and hence incur the penalty of $\log N$.

REFERENCES

[1] R. W. Yeung, S.-Y. R. Li, N. Cai, and Z. Zhang, "Network Coding Theory", now Publishers, 2005.

[2] R. Koetter and M. Medard, "An Algebraic approach to Network Coding", IEEE/ACM Transactions on Networking, Vol 11, No. 5, pp. 782-795, 2003.

[3] S. Deb, M. Medard and C. Choute, "Algebraic gossip: A network coding approach to optimal multiple rumor mongering", in Proceedings of the IEEE International Symposium on Information Theory, Vol. 52, No. 6, June 2006.

[4] D. Mosk-Aoyama and D. Shah, "Information Dissemination via Network Coding", in Proceedings of the IEEE International Symposium on Information Theory (ISIT), Seattle, July 2006.

[5] U. Fiege, D. Peleg, P. Raghavan, E. Upfal, "Randomized Broadcast in Networks", in Random Structures and Algorithms, Vol. 1, No. 4, 1990

APPENDIX

A. Proof of Lemma 1

Recall that $A \triangleq \bigcap_{u \in R(i, t)} \{S(j, t) \geq m_u\}$.

$$\Pr(S(i, t) = S(j, t)) = \Pr(S(i, t) = S(j, t), A) + \Pr(S(i, t) = S(j, t), A^c)$$

(6)

where $A^c$ is the complement event. In the following we show that

$$\Pr(S(i, t) = S(j, t), A^c)$$

is small if the field size is large. We have,

$$\Pr(S(i, t) = S(j, t), A^c) = \Pr\left(S(i, t) = S(j, t), \bigcup_{u \in R(i, t)} S(j, t) \not\in m_u\right) \leq \sum_{u \in R(i, t)} \Pr\left(S(i, t) = S(j, t), S(j, t) \not\in m_u\right) \leq \sum_{u \in R(i, t)} \Pr\left(S(i, t) = S(j, t)|S(j, t) \not\in m_u\right)$$

(7)

The message percolates from the node $u$ to $i$ along possibly many paths in the predecessor graph (enumerated by all possible paths from $u$ to $i$ in the predecessor graph). Consider any such path and remove the cycles in the path. Now along the path every node is distinct.

Let $u = l_0, l_1, \ldots, l_k = i$ denote the distinct vertices of the path and let $\mu_i$ denote the message sent from $l_{i-1}$ to $l_i$ for $1 \leq i \leq k$. Also define $B \triangleq S(j, t) \not\in m_u$. We have

$$\Pr(S(i, t) = S(j, t)|B) \leq \Pr(\mu_k \in S(j, t)|B) = \Pr(\mu_k \in S(j, t), \mu_{k-1} \not\in S(j, t)|B) + \Pr(\mu_k \in S(j, t), \mu_{k-1} \in S(j, t)|B) \leq \Pr(\mu_k \in S(j, t)|\mu_{k-1} \not\in S(j, t), B) + \Pr(\mu_{k-1} \in S(j, t)|B)$$

Recursively using the above relation, we get

$$\Pr(S(i, t) = S(j, t)|B) \leq \sum_{b=2}^{k} \Pr(\mu_b(t) \in S(j, t)|\mu_{b-1}(t) \not\in S(j, t), B) + \Pr(\mu_1(t) \in S(j, t)|B) \leq \frac{k}{q} \leq \frac{N}{q}$$

(8)

The inequality (a) follows since $\Pr(\mu_b(t) \in S(j, t)|\mu_{b-1}(t) \not\in S(j, t), B)$ implies that the node $l_{b-1}$ weights the message $\mu_{b-1}(t)$ with a specific symbol in $\mathbb{F}_q$. (otherwise $\mu_{b-1}(t)$ would be in $S(j, t)$) The probability of this event is upper bound by $\frac{1}{q}$. The last inequality follows since the path has distinct nodes and hence the length is bounded by $N$.

Substituting (8) in (7) gives

$$\Pr(S(i, t) = S(j, t), A^c) \leq \sum_{u \in R(i, t)} \frac{N}{q} \leq \frac{N^2}{q}$$

(9)

Substituting in (6), we get

$$\Pr(S(i, t) = S(j, t)) \leq \Pr(S(i, t) = S(j, t), A) + \frac{N^2}{q}$$

B. Proof of Lemma 2

Recall that $A \triangleq \bigcap_{u \in R(i, t)} \{S(j, t) \geq m_u\}$. We have,

$$\Pr\left(S(i, t) = S(j, t), A\right)$$

$$= \Pr\left(S(i, t) = S(j, t), A, \dim(S(i, t)) = n\right) + \Pr\left(S(i, t) = S(j, t), A, \dim(S(i, t)) < n\right)$$

$$\leq \Pr\left(\dim(S(i, t)) = n\right) + \Pr\left(S(i, t) = S(j, t), A, \dim(S(i, t)) < n\right)$$

$$\leq \gamma(t) + \Pr\left(S(i, t) = S(j, t), A, \dim(S(i, t)) < n\right)$$

(10)

where we used the fact that at any time instant, a node is picked independently, uniformly at random from the entire network, hence $\Pr(\dim(S(i, t)) = n) = \gamma(t)$.

Note that $\{S(i, t) = S(j, t), A\}$ is the event that the message of every distinct node in the predecessor graph of node $i$ is contained in $S(i, t)$. Thus, the node $i$ can decode the initial message of all the distinct nodes in the predecessor graph. But since $\dim(S(i, t)) < n$, there must exist a node which is not present in the predecessor graph. Thus we have,

$$\Pr\left(S(i, t) = S(j, t), A, \dim(S(i, t)) < n\right)$$

$$\leq \Pr(\text{There exists a node } u \text{ not present in the predecessor graph } W(i, t))$$

$$\leq \sum_{u \in G} \Pr(u \text{ has not branched to } i \text{ within } t)$$

(11)

The argument (a) follows from the union bound and invokes a new term – “branching” defined in [5]. In [5], the authors consider a randomized protocol for broadcasting a message over an arbitrary network under the synchronous model and compute an upper bound to the expected time by which a
message originating at a node spreads all over the network with high probability. The meaning of the term branching is the following – a node $u$ branches to $i$ within time $t$ if and only if it is present in the predecessor graph of $i$ at time $t$. Note that the branching process itself is defined on the original graph $G$ and not on the predecessor graph. In the following, we modify the arguments in [5] to upper bound the probability in \( \Omega \).

$$\Pr(u \text{ has not branched to } i \text{ within } t) \leq \Pr(u \text{ has not branched to } i \text{ within } t)$$
onumber

along the shortest path from $u$ to $i$.

Let $u = x_0, x_1, \ldots, x_m = i$ denote the shortest path from $u$ to $i$ in $G$. A vertex $w$ not on the path cannot be connected to more than three vertices on the path. (Otherwise, the shortest path would be via $u$.) Thus it follows that $\sum_{r=0}^{m} \deg(x_r) \leq 3N$.

Let $\tilde{t}_{r+1}$ denote the time to transmit a message from $x_r$ to $x_{r+1}$. Clearly, $E(\tilde{t}_{r+1}) = \frac{N}{\deg(x_r)} = N \deg(x_r)$. The expected time, $E(\tilde{t}_{u,i})$ to transmit the message along the shortest path from $u$ to $i$ is therefore upper bound by $3N^2$.

From the Markov inequality, it follows that $\Pr(\tilde{t}_{u,i} > 6N^2) < \frac{1}{2}$. Therefore, after $4 \log N$ epochs of duration $6N^2$ each,

$$\Pr(\tilde{t}_{u,i} > 24N^2 \log N) < 2^{-4 \log N} = \frac{1}{N^4}.$$  

If $t > 24N^2 \log N$, we continue \( \Omega \) as

$$\Pr(S(j,t) = S(i,t), A, \dim(S(i,t)) < n) \leq \sum_{u \in G} \Pr(\tilde{t}_{u,i} > t)$$

$$\leq \sum_{u \in G} \Pr(\tilde{t}_{u,i} > 24N^2 \log N) \leq \frac{1}{N^3}.$$ (12)

Substituting the upper bound to the probability in \( \Omega \),

$$\Pr(S(j,t) = S(i,t), A) \leq \gamma(t) + \frac{1}{N^3}.$$ (13)

C. Computation of $E[\tau_N]$

The computation is done in two parts. We first compute $E[\tau_{N-1}]$ and then $E[\tau_N - \tau_{N-1}]$. We omit the prefactor $\frac{q}{4^m}$ in the computations and append it directly in the result.

$$E[\tau_{N-1}] = \sum_{m=1}^{N-2} \sum_{l=1}^{2m} E[X_{m+1}^{l}]$$

\[
\leq \sum_{m=1}^{N-2} 2m \sum_{l=1}^{2m} \frac{1}{p^*(\tau_{m+1})} = \sum_{m=1}^{N-2} \frac{2m}{p^*(\tau_{m+1})}
\]

\[
\leq \sum_{m=1}^{N-2} \frac{2m}{1 - \frac{(m+1)^2}{N^2} - \frac{1}{N} - \frac{2}{N^4}}
\]

\[
\leq \sum_{m=1}^{N-2} \frac{2(m + 1)}{(1 - \frac{(m+1)^2}{N^2} - \frac{1}{N} - \frac{2}{N^4})}
\]

\[
= N \sum_{m=2}^{N-1} \frac{2m/N}{(1 - \frac{m^2}{N^2} - \frac{1}{N} - \frac{2}{N^4})}
\]

$$\leq (a) \sum_{\theta=2/N}^{1 - 1/N} \frac{2\theta}{(1 - \theta^2 - \frac{1}{N} - \frac{2}{N^4})}d\theta + N \frac{2(N-1)/N}{(1 - (N-1)/N^2 - 1/N - 2/N^4)}$$

\[
= -N^2 \left( \ln \left( 1 - (1 - 1/N)^2 \frac{1}{N} - \frac{2}{N^4} \right) \right) + \frac{2(N-1)}{\left( 1 - \frac{1}{N^4} - \frac{3}{N^3} \right)}
\]

$$\leq N^2 \ln N + \frac{2(N-1)}{\left( \frac{1}{N} - \frac{1}{N^3} - \frac{3}{N^2} \right)^3} < 3N^2 \ln N.$$ (14)

In the above, (a) follows by bounding the summation (excluding the term for $m = N - 1$) by the integral.

We now compute the expected time from $\tau_{N-1}$ to $\tau_N$. Let us define $\tau_{N-1}^{(m)}$ as the time in the interval $(\tau_{N-1}, \tau_N)$ when the cluster of nodes which have decoded all messages is of size $m$. Let $m_0$ be the size of the cluster at $\tau_{N-1}$. Thus $m \geq m_0$.

Since $t > 24N^2 \log N$ we use Lemma 1 and Lemma 2 and thus for $t \in (\tau_{N-1}^{(m)}, \tau_N^{(m+1)})$ we have

$$\Pr(S(i,t) = S(j,t)) \leq \gamma(t) + \frac{2}{N^3} \leq \frac{m}{N} + \frac{2}{N^3}.$$ (15)

Thus probability of dimension increase in the interval $(\tau_{N-1}^{(m)}, \tau_N^{(m+1)})$ is lower bound by

$$p(\tau) \geq \left( 1 - \frac{m}{N} - \frac{2}{N^3} \right).$$ (16)

Since the interval $(\tau_{N-1}, \tau_N)$ contains $N(N-1) - (N-1)^2 = 2(N-1)$ dimension increments, it follows that

$$E[\tau_{N-1}^{(m+1)} - \tau_{N-1}^{(m)}] \leq \frac{2(N-2)}{(1 - \frac{m}{N} - \frac{2}{N^3})}.$$ (17)

Summing the above over $m$ going from $m_0$ to $N - 1$, we get

$$E[\tau_N - \tau_{N-1}] \leq (2N - 2) \sum_{m_0}^{N-1} \frac{1}{(1 - \frac{m}{N} - \frac{2}{N^3})}$$

\[
\leq N(2N - 2) \sum_{m_0}^{N-1} \frac{1}{(N - m - \frac{2}{N^3})}
\]

$$< 2N^2 \ln N.$$ (17)

Combining (14) and (17) we get

$$E[\tau_N] \leq \frac{q}{q-1} 5N^2 \ln N.$$