Perfect Pipelining for Streaming Large File in P2P Networks

Fei Chen[*]   Xiaowei Wu[*]

Abstract

We study the efficiency of large file streaming in a peer-to-peer network in which a large file is cut into many pieces of equal size, and initially all pieces are known only by one source node. We analyze the number of rounds required, called the *finishing time*, for all nodes in the network to collect all pieces in the default order.

Based on the basic PUSH-PULL protocol, we design the Constant Out-degree Protocol (COP). At the beginning of the protocol, each node selects a constant number of neighbors, with only whom communication will be initiated. We focus our analysis on the performance of COP on preferential attachment graphs, which are believed to model peer-to-peer networks well. We show that a tight bound of $\Theta(B + \log n)$ rounds can be achieved with high probability for streaming B pieces in preferential attachment graphs with n nodes. Moreover, we show that there is a dichotomy in the results depending on how neighbors are contacted in each round; specifically, when each node avoids repeating initiation with neighbors in the previous $M \geq 2$ rounds, then the finishing time is improved to $\Theta(B + \frac{\log n}{\log \log n})$ with high probability.

For lower bounds, we show that there is a class of regular graphs in which perfect pipelining is impossible for any PUSH-PULL protocols using random neighbor selection.

1 Introduction

The problem of information disseminating in a distributed network has been extensively studied. It is assumed that this information, usually called a rumor, was known by only one node in a large connected network. The goal is to disseminate this rumor to all nodes in the network, with bounded bandwidth and number of communications. Many protocols have been proposed based on various network topologies [7, 14, 15]. Among these protocols, distributed algorithms have gained more attention since each node does not require global information about the network.

A gossip based protocol was first introduced by Demers et al. [7] for maintaining replicated databases. In this protocol, each node is only allowed to communicate with its own neighbors in each round. A similar distributed protocol called PUSH-PULL protocol was introduced later. In this protocol, each node initiates communication with one of its neighbors per round; the node *pushes* the rumor to the neighbor if it is informed of the rumor, or *pulls* the rumor from the neighbor if the neighbor is informed. A common performance measure of protocols is the number of rounds required to have all nodes informed, which is called the *finishing time*.

We consider the rumor spreading problem when the rumor is not a tiny piece of information, but a large file which can not be transmitted from one node to another in one round. In this case, we divide the file into B small pieces so that each piece can be transmitted in one round. This problem can also be considered as the multiple-rumor version of the rumor spreading problem with streaming requirement. We analyse the finishing time for all nodes in the network to collect all B

---

[*]Department of Computer Science, the University of Hong Kong. {`fchen,xwwu`}`@cs.hku.hk`

pieces in the correct order. For instance, in peer-to-peer video streaming, a node needs to receive earlier portions first and the next portion needs to arrive before all currently received portions have been played. We consider two properties of a protocol that are important for streaming multiple pieces: *perfect pipelining* and *ordered pieces transmission*, which we explain below.

1.1 Our Contribution and Results

We study the large file streaming problem in which B ordered pieces of information known to one source node are to be spread in a graph with n nodes. We consider protocols that have the following properties.

1. Loosely speaking, a (randomized) protocol for multiple pieces spreading achieves *perfect pipelining* on $G(V, E)$ if the following holds: if the finishing time for spreading one piece is T (with high probability), then the finishing time for spreading B pieces is $O(B+T)$ (with high probability) in G.

2. A protocol employs *ordered pieces transmission* if the pieces are transmitted and received in the correct order.

Our results and contribution are summarized as follows.

Achieving perfect pipelining with COP. In Section 4, we modify the basic PUSH-PULL protocol and design the Constant Out-degree Protocol (COP). At the beginning of the protocol, each node selects a constant number of neighbors, with only whom communications will be initiated.

We show that using COP, steaming B pieces in preferential attachment graphs requires $O(B+\log n)$ rounds, which achieves perfect pipelining. We also prove a tight lower bound $\Omega(B + \log n)$ for the finishing time. As considered by Doerr et al. [8], if each node does not repeat neighbor initiation from the previous 2 rounds, COP can achieve finishing time $O(B + \frac{\log n}{\log \log n})$, which is also tight. This theoretical gap in the finishing time between the two cases suggests that after the random neighbor selection phase to identify out-going neighbors, deterministic round-robin for the pieces streaming phase is optimal.

We briefly outline our technical contributions. According to Arthur and Panigrahy [1], spreading B pieces in a graph with diameter d and maximum degree D using random PUSH requires $O(D(B + \max\{d, \log n\}))$ rounds with high probability (w.h.p.), no matter how the pieces are chosen for transmission. For preferential attachment graphs, the resulting upper bound $O(\sqrt{n}(B + \log n))$ is loose because the maximum degree is as large as $\Omega(\sqrt{n})$. However, the large degree nodes are crucial for single piece spreading [8] since they are highly connected. To overcome this technical hurdle, our protocol (COP) is based on the following two observations: (1) for single piece spreading, even if each node is "active" for only a constant number of rounds, w.h.p., the piece can still reach all nodes via paths of length $O(\frac{\log n}{\log \log n})$; (2) we can direct the edges to achieve constant out-degree, even when the underlying undirected graph has a high maximum degree; this allows us to achieve the same result as in [1] if we use both PUSH and PULL.

Implication on the performance of streaming. We show that our protocol achieves tight upper bound $\Theta(B + \log n)$ when the *ordered pieces transmission* is adopted. In this case, all nodes will receive all pieces of the file in the correct order and perfect pipelining is achieved. As far as we know, this is the first time perfect pipelining is proved for streaming large file in peer-to-peer network.

Impossibility of perfect pipelining in some cases. In Section 3 we consider the basic PUSH-PULL protocol for streaming B pieces in complete graphs. Unfortunately, perfect pipelining cannot be achieved with this simple protocol. In Section 5, we give a lower bound for the finishing time in a particular class of D-regular graphs, called *Necklace Graphs*. We show that spreading B pieces under any PUSH-PULL protocol in which communication partners are chosen uniformly at random in each round (no matter how those pieces are selected for transmission) in Necklace Graph requires $\Omega(D(B + d))$ rounds, where d is the diameter, for any $\omega(1) \leq D \leq o(n)$.

1.2   Related Work

The single piece (usually called a rumor) spreading problem has been extensively studied. Using the PUSH-PULL protocol, the finishing time (also called stopping time) is $\Theta(\log n)$ for both complete graphs [16, 13] and random graphs [9, 10]. Rumor spreading in preferential attachment graphs (PA-graphs) was also studied and two tight upper bounds $O(\log n)$ and $O(\frac{\log n}{\log \log n})$ for different protocols were shown by Doerr et al. [8]. Given a graph with conductance ϕ, an $O(\phi^{-1} \log n)$ upper bound was proved by Giakkoupis [11], which is tight. For rumor spreading in general graphs, Censor-Hillel et al. [4] gave an algorithm which constructs subgraphs along with the rumor spreading process and showed that the finishing time is at most $O(d + \text{poly} \log n)$, where d is the diameter of the underlying graph, with no dependence on the conductance.

Although observed by experiments [5] and extensively applied on the Internet, the high efficiency of multiple-rumor spreading is hardly analyzed. Deb et al. [6] presented a protocol based on random linear network coding and showed that in the complete graph spreading $B = \Theta(n)$ rumors requires at most $O(n)$ rounds, which is asymptotically optimal. Also using network coding, Haeupler [12] gave a protocol that achieves perfect pipelining in many other graphs. In their protocol, communication partners exchange a random linear combination of rumors and they show that after receiving enough combinations, all rumors can be decoded locally.

Given a graph with diameter $d = \Omega(\log n)$ and maximum out-degree D, an $O(D(B + d))$ upper bound on the finishing time was proved by Arthur and Panigrahy [1] for spreading B rumors with PUSH. Since $\Omega(B+d)$ is a trivial lower bound, this upper bound is tight when D is a constant. This result actually inspired the construction of our protocol COP, which attempts to find a subgraph that has constant out-degree and small diameter in the first step.

2   Preliminaries

Consider a graph G with n nodes and B pieces indexed by $\{1, 2, \ldots, B\}$; for streaming application, we assume that smaller index means higher priority. Throughout this paper we suppose that initially a single *source node* $u_0 \in V$ has all B pieces while other nodes have no pieces. We say a node is *waiting* if it has no pieces, *downloading* if it has at least one piece, and *finished* if it has all the pieces. We use log to denote the logarithm with base 2.

Pieces are transmitted in a round-by-round manner. During a round of the pieces streaming process, we say there is a piece *transmission* between two nodes if some piece is transmitted from one node to the other. The *finishing time* is the number of rounds required until all nodes in the graph become finished.

The Basic PUSH-PULL Protocol. For each node $u \in V$ and integer $t \geq 1$, let p_t^u be the current downloading position of u at time t. Note that $p_0^u = B$ if $u = u_0$, and $p_0^u = 0$ otherwise. In each round each node independently selects a neighbor uniformly at random, and initiates communication with that neighbor. In round $t \geq 1$, for an instance of communication initiated by

u between u and v, the transmission follows the following rules.

1. If $p_{t-1}^u = p_{t-1}^v$, then no transmission is made;

2. If $p_{t-1}^v > p_{t-1}^u$, then node u *pulls* from node v the piece with index $p_{t-1}^u + 1$;

3. If $p_{t-1}^v < p_{t-1}^u$, then node u *pushes* to node v the piece with index $p_{t-1}^v + 1$.

Note that each node initiates communication with exactly one neighbor in each round and at most one piece is transmitted. Also note that all B pieces are initiated by the source node at the beginning and hence each node in the graph will receive all the pieces in the fixed order $1, 2, \ldots, B$.

In Section 3, we show that the basic PUSH-PULL protocol cannot achieve perfect pipelining on the complete graph; in Section 4, we modify the basic PUSH-PULL protocol and design the Constant Out-degree Protocol, which achieves perfect pipelining for directed random graphs and preferential attachment graphs.

3 Low Efficiency of Basic PUSH-PULL

To motivate why we consider COP, we analyze the performance of the basic PUSH-PULL protocol on complete graph K_n with source node u_0 having all B pieces. We show that perfect pipelining can not be achieved since (on average) each node has to wait for $\Omega(\log n)$ rounds to get the next piece.

Theorem 3.1 *The finishing time for streaming B pieces from one source node in a complete graph with n nodes using the basic PUSH-PULL protocol is $\Omega(B \log n)$ with probability $1 - \exp(-\Theta(B))$.*

Note that by experiments, Theorem 3.1 is not true if the pieces are spread randomly. We conjecture that the finishing time for spreading B pieces with random pieces selection is $O(B + \log n)$. To prove Theorem 3.1, we recall some known results for spreading a single rumor in the complete graphs.

Fact 3.1 *The finishing time for spreading one rumor from one source node in the complete graph K_n using the basic PUSH-PULL protocol is*

(1) ([9, Theorem 4.1]) $\Omega(\log n)$ with probability $1 - o(1)$.

(2) ([13, Theorem 2.1]) $O(\log n)$ with probability $1 - O(\frac{1}{n^3})$.

Proof of Theorem 3.1: Let T be the finishing time. Observe that every node will receive the pieces in the order $1, 2, \ldots, B$. For each $i \in \{1, 2, \ldots, B\}$, let T_i be the first round when piece i is obtained by at least 2 nodes and $t_i = T_i - T_{i-1}$, where we define $T_0 = 0$. Note that $t_1 = T_1 = 1$ and $T \geq T_B = \sum_{i=1}^{B} t_i$. Since we are proving a lower bound, we make a speed up assumption that once piece i is received by at least 2 nodes, all nodes in V immediately have all pieces j if $j < i$. Note that this assumption will only make the pieces streaming process faster and under this assumption, all t_is are independent. Under the speed up assumption, we fix one $i \in \{2, \ldots, B\}$ and analyze t_i.

As defined above, t_i is the number of rounds needed from the time piece $i - 1$ leaves the source to the time piece i leaves the source, which happens when the source node u_0 contacts a node v that has piece $i - 1$ or such a node v contacts u_0. Suppose in a round between T_{i-1} and T_i, at most ϵn nodes has piece $i - 1$, where $\epsilon \in (0, \frac{1}{2})$, then the probability for a communication between u_0 and a

4

node having piece $i-1$ to appear is at most $\epsilon + 1 - (1 - \frac{1}{n})^{\epsilon n} = \epsilon + 1 - (1 - \frac{\epsilon n}{n} + \frac{\epsilon n(\epsilon n - 1)}{2n^2} - \ldots) \leq (1 + \epsilon) - (1 - \epsilon) = 2\epsilon$.

Let t be the number of rounds needed until $\sqrt{n}$ nodes have piece $i-1$ after T_{i-1}. Then with probability $1 - O(\frac{1}{n^3})$, $t = O(\log n)$ and in each round, the probability for a communication between u_0 and a node having piece $i-1$ to appear is at most $2\epsilon = \frac{2\sqrt{n}}{n}$. Hence we have

$$\Pr(t_i \geq t) \geq (1 - \frac{2\sqrt{n}}{n})^t \geq 1 - \frac{2t\sqrt{n}}{n} \geq \frac{1}{2}.$$

Hence with probability at least $\frac{1}{2}$, $t_i \geq t$. Since by part 2 of the proof of [13, Theorem 2.1], with probability $1 - o(1)$, $t = \log_3 \sqrt{n} + O(\ln \ln n) = \Omega(\log(\sqrt{n})) = \Omega(\log n)$, we have $t_i = \Omega(\log n)$ with probability at least $\frac{1}{2}(1 - o(1)) \geq \frac{1}{3}$.

By Chernoff bound, we conclude that with probability at least $1 - \exp(-\frac{B-1}{48}) = 1 - \exp(-\Theta(B))$, there are at least $\frac{B-1}{4}$ of t_i's satisfying $t_i = \Omega(\log n)$ and hence the finishing time $T \geq T_B = \sum_{i=1}^{B} t_i \geq \frac{(B-1)}{4}\Omega(\log n) = \Omega(B \log n)$. ∎

However, if we can guarantee that at least a constant fraction of communications are useful, then after $O(B)$ rounds, all nodes will become finished. Following this idea, we define in Section 4 a protocol that constructs a constant maximum out-degree subgraph before applying the basic PUSH-PULL protocol. We show that in several cases, the subgraph constructed by our protocol has small diameter and using basic PUSH-PULL protocol, perfect pipelining can be achieved.

4 Constant Out-degree Protocol

In this section, we propose a simple but efficient randomized protocol called Constant Out-degree Protocol (COP).

Constant Out-degree Protocol. Let D be a fixed constant parameter used in the protocol. Suppose we are given an input graph with one source node.

(1) **Out-going Edge Selection Phase.** A directed subgraph H is formed, where each node u chooses independently its *outgoing edges* and *outgoing neighbors*, with only whom node u will initiate communications in the next phase. We consider all input graphs as multi-graphs. Each node u chooses $\min\{D, deg(u)\}$ (out-going) edges incident to u uniformly at random without replacement. The chosen edges are the outgoing edges and the incident nodes (excluding u) are the outgoing neighbors of u in H.

(2) **Pieces Streaming Phase.** In each round, each node chooses one of its sampled outgoing edges uniformly at random to initiate communication. In round $t \geq 1$, for an instance of communication initiated by u between nodes u and v, the transmission follows as the basic PUSH-PULL protocol.

In other words, we construct a directed subgraph H and apply the basic PUSH-PULL protocol to H, where the directions of initiation coincide with the directions of edges. Since the only difference between COP and the basic PUSH-PULL protocol is an extra neighbor selection process, we know that if perfect pipelining can be achieved in the (random) subgraph H using basic PUSH-PULL, then perfect pipelining is achieved by COP in the input graph. We show in this section that perfect pipelining can be achieved by applying COP to stream multiple pieces in directed random graphs and preferential attachment graphs. Unless otherwise specified, the *diameter* of a graph refers to that of the underlying induced **undirected** graph.

Deterministic Rumor Spreading Phase. Suppose we consider the case that after the random subgraph H is constructed in the first phase each node initiates communication with its outgoing

neighbors in a round-robin way. Recall that since the out-degree of the subgraph H is at most D, there is communication between two neighboring nodes at most every D rounds. It follows that if the diameter of H is at most d, all nodes will become finished after $D(B + d)$ rounds.

Fact 4.1 *Given any input graph with n nodes, if the underlying undirected graph of the subgraph constructed by COP with constant parameter D has diameter d and each node contacts its neighbors in a round-robin way, then for all $B \geq 1$, the finishing time for streaming B pieces starting from one source node u_0 in the input graph using COP with parameter D is at most $D(B + d) = O(B + d)$.*

Note that Fact 4.1 holds with probability 1 since it is deterministic. If the diameter of the graph is $\Omega(\log n)$, the random neighbor chosen process achieves a similar upper bound w.h.p.. We use the following fact [1, Proposition 4.1] on streaming multiple pieces along a path.

Fact 4.2 (Streaming Multiple Pieces on a Path) *Suppose there exists a path $P(v_0, v_s) = \{v_0, v_1, \ldots, v_s\}$ of length s between two nodes v_0 and v_s such that in each round, communication between any two adjacent nodes v_i and v_{i+1} is performed with probability at least p independently for each pair of adjacent nodes. Assume only v_0 has all B pieces and all other nodes are waiting. Then if $s' \geq s$, with probability $1 - 2\exp(-\frac{s'}{2})$, node v_s will become finished after $\frac{4(B+4s')}{p}$ rounds.*

The fact above implies that by using COP protocol, perfect pipelining can be achieved in any graph if the subgraph H constructed has a small diameter d.

Theorem 4.1 *Suppose the subgraph constructed by COP has diameter d, then for all $B \geq 1$, the finishing time for streaming B pieces starting from one source node u_0 is at most $O(B + \max\{d, \log n\})$ with probability $1 - O(\frac{1}{n})$.*

Proof: We denote the subgraph constructed by COP with constant parameter D by H, which has diameter d. Hence for any node v, there exists a path $P = (u_0, u_1, \ldots, u_{s-1}, u_s = v)$ in the undirected graph induced by H between u_0 and v, where $s \leq d$ is the length of the path.

Hence, in the directed graph H, for any $i = \{0, 1, \ldots, s - 1\}$, there is a directed edge between nodes u_i and u_{i+1}. Since both nodes u_i and u_{i+1} have out-degree at most D, which is a constant, communication between u_i and u_{i+1} will be performed with probability at least $\frac{1}{D}$ in each round. However, if there exists two directed edges (u_i, u_{i+1}) and (u_i, u_{i-1}) that start at u_i, then the communications between u_i and u_{i+1}, u_i and u_{i-1} are not independent (only one of them can happen in each round). To make the communications independent, we define *phases* such that each phase consist of two rounds. In each phase, for each $i = \{1, \ldots, s - 1\}$, we assume that the communication between u_i and u_{i+1} is invalid in the first round and the communication between u_i and u_{i-1} is invalid in the second round. Note that the assumption will only make the downloading process slower. Under this assumption, we know that for each $i = \{0, 1, \ldots, s - 1\}$, in each phase, communication between u_i and u_{i+1} will be performed with probability at least $\frac{1}{D}$, independently.

Hence, for each node v in H, there is a path of length $s \leq d$ between u_0 and v such that in each phase, communications between any two adjacent nodes will be performed with probability at least $\frac{1}{D}$, independently for each pair of adjacent nodes. Hence by Fact 4.2, if $s' \geq s$, with probability $1 - 2\exp(-\frac{s'}{2})$, node v will become finished after $4D(B + 4s')$ phases. If $s \geq 4\log n$, then set $s' = s$ and node v will become finished after $4D(B + 4s) = O(B + d)$ phases with probability $1 - O(\frac{1}{n^2})$. If $s < 4\log n$, then set $s' = 4\log n$ and node v will become finished after $4D(B + 16\log n) = O(B + \log n)$ phases with probability $1 - O(\frac{1}{n^2})$. Therefore, with probability $1 - O(\frac{1}{n^2})$, node v will become finished after $O(B + \max\{d, \log n\})$ phases, which means $O(B + \max\{d, \log n\})$ rounds. A union bound on $n - 1$ nodes in H yields the result. ∎

By Theorem 4.1, we know that the finishing time of COP is highly related to the diameter d of the subgraph constructed in the first step. If $d = \Omega(\log n)$, then using round-robin in the pieces streaming phase has the same bound as random neighbor selection. However, as we shall later see, the PA-graphs have diameter $O(\frac{\log n}{\log \log n})$ w.h.p., and indeed we show that there is a dichotomy in the finishing times between deterministic and random pieces streaming phases.

4.1 COP on Directed Random Graphs

We apply COP with $D = 8$ to stream B pieces in directed random graphs $G_{n,p}$ with $p \geq \frac{6 \log n}{n}$. Observe that the neighbor selection phase actually constructs a random D-out-degree subgraph.

Theorem 4.2 (Finishing Time of COP in $G_{n,p}$) *For all $B \geq 1$, the finishing time for streaming B pieces starting from one source node in directed random graphs $G_{n,p}$ with n nodes and $p \geq \frac{6 \log n}{n}$ using the constant out-degree protocol with $D = 8$ is at most $O(B + \log n)$ with probability $1 - O(\frac{1}{n})$.*

Note that for large enough n and p, all nodes in $G_{n,p}$ have out-degrees at least D w.h.p..

Proposition 4.1 (Directed $G_{n,p}$ Has Large Out-degree) *Consider a directed $G_{n,p}$ graph with $p \geq \frac{6 \log n}{n}$. Then, with probability $1 - O(\frac{1}{n})$ the out-degree $d_{out}(u) \geq D = 8$ for every node u.*

Proof: Fix a node u. Since each edge with tail at u appears independently with probability p, we have $\mathbf{E}[d_{out}(u)] = np$. Since $p \geq \frac{6 \log n}{n} \geq \frac{6D}{n}$, by Chernoff bound we have $\Pr(d_{out}(u) < D) \leq \Pr(d_{out}(u) < \frac{1}{6}\mathbf{E}[d_{out}(u)]) \leq \exp(-\frac{1}{2}(\frac{5}{6})^2 \cdot 6 \log n) = O(\frac{1}{n^2})$. That is, we have $d_{out} \geq D$ with failure probability at most $O(\frac{1}{n^2})$. Applying a union bound over all nodes in $G_{n,p}$ yields the result. ∎

In the rest of this section we assume that the directed $G_{n,p}$ graph under consideration has the property as stated in Proposition 4.1, i.e., the out-degree $d_{out}(u) \geq D$ for each node u. This immediately implies that in the subgraph H each node has an out-degree of exactly D. Next we show some properties of the subgraph H constructed by COP and then use these properties to give an upper bound of the finishing time T.

The following easy observation implies that given the event of Proposition 4.1, the subgraph H can be constructed in an alternative procedure, which will be employed in our proofs.

Lemma 4.1 *Let H be a directed subgraph of $G_{n,p}$ which is constructed by sampling a uniform set of D outgoing edges independently for each node. Then for each (ordered) pair of nodes u and v, the edge (u, v) appears in H with probability $\frac{D}{n-1}$.*

Proof: The independence follows from independence of the sampling procedures for constructing $G_{n,p}$ and the subgraph H. Consider a pair of nodes u and v. Since in $G_{n,p}$ the out-degree $d_{out}(u)$ of u is at least D, we know that the probability of (u, v) appears in H equals to $\sum_{d=D}^{n-1} \Pr((u, v) \in E \mid d_{out}(u) = d) \cdot \Pr(d_{out}(u) = d) = \sum_{d=D}^{n-1} \frac{\binom{n-2}{d-1}}{\binom{n-1}{d}} \cdot \frac{\binom{d-1}{D-1}}{\binom{d}{D}} \cdot \Pr(d_{out}(u) = d) = \frac{D}{n-1}$. ∎

We define some useful notations to analyze the diameter of the subgraph H.

Definition 4.1 (Directed Distance) *For nodes $u, v \in V$, the directed distance $d(u, v)$ from node u to node v is the length of a shortest directed path from u to v. In particular, let $d(u, v) = 0$ when $u = v$. If there is no directed path from u to v, then $d(u, v) = \infty$.*

Definition 4.2 (Level and Neighborhood) *For a node $u \in V$ and integer $k \geq 0$, a node v is at level k with respect to u if the distance from u to v is $d(u,v) = k$. Let $L_k^u := \{v \in V : d(u,v) = k\}$ be the set of nodes at level k with respect to u. In particular, let $L_0^u = \{u\}$. The k-neighborhood $H_k^u := \cup_{i \leq k} L_i^u$ of u is the set of nodes at levels at most k with respect to u. Also define $\ell_k^u := |L_k^u|$ and $h_k^u := |H_k^u|$.*

Equivalent Construction of H. Consider a *directed* complete graph K_n with n nodes. Each node u independently samples a set of D edges from all outgoing edges of u (i.e., edges with tail at u) uniformly at random, and include these D edges in H.

According to the Equivalent Construction of H, for $u \in V$ and $k \geq 0$, the set L_{k+1}^u can be constructed in the following manner. During the sampling procedure for outgoing neighbors, if a node $v \in L_k^u$ selects a neighbor w and if $w \notin H_k^u$, then w is included in L_{k+1}^u. Let γ_u be the smallest integer such that $h_{\gamma_u}^u \geq \frac{n}{8}$. Next we prove some connectivity properties of H in Lemmas 4.2 and 4.3, which imply that the graph has $O(\log n)$ diameter.

Lemma 4.2 (Nodes in $\log n$-Neighborhood) $\forall u \in V$, $h_{\log n}^u \geq \frac{n}{8}$ *with probability* $1 - O(\frac{1}{n^2})$.

Proof: For convenience we drop the superscript u in this proof. We first show that for each $i \in \{0, 1, \ldots, \gamma - 1\}$, $h_{i+1} \geq 2h_i$ with probability at least $1 - O(\frac{1}{n^3})$. For $i = 0$, note that $h_0 = 1$ and $\ell_1 = D$. Therefore $h_1 = D + 1 \geq 2h_0$ with probability 1.

For $i \geq 1$, we show that given $h_{i'+1} \geq 2h_{i'}$ for all $i' \leq i-1$, then $\Pr(h_{i+1} < 2h_i) \leq O(\frac{1}{n^4})$. To see this, observe that $h_{i+1} < 2h_i$ implies the following event: the set of nodes selected by nodes in L_i contains at most h_i nodes in H_i and at most h_i nodes in $V \setminus H_i$. The probability that a node in L_i selects its D neighbors within a particular set with size $2h_i$ is $\binom{2h_i}{D}/\binom{n-1}{D} = \frac{2h_i}{n-1} \cdot \frac{2h_i-1}{n-2} \cdots \frac{2h_i-D+1}{n-D} \leq \left(\frac{2h_i}{n-1}\right)^D$. Note that $\ell_i = h_i - h_{i-1} \geq \frac{h_i}{2}$. Also recall that $D = 8$ and $h_i \leq \frac{n}{D}$. Conditioned on the inequalities $h_{i'+1} \geq 2h_{i'}$, we have

$$\Pr(h_{i+1} < 2h_i) \leq \binom{n-h_i}{h_i}\left(\frac{2h_i}{n-1}\right)^{D\ell_i} \leq \binom{n}{2h_i}\left(\frac{2h_i}{n-1}\right)^{4h_i} \leq \left(\frac{en}{2h_i}\right)^{2h_i}\left(\frac{2h_i}{n-1}\right)^{4h_i} \leq \left(\frac{6h_i}{n}\right)^{2h_i}$$

Note that $h_i \in [2, \frac{n}{8})$. If $h_i = \Theta(n)$, then $\Pr(h_{i+1} < 2h_i) \leq \left(\frac{3}{4}\right)^{\Theta(n)} = O(\frac{1}{n^4})$; if $h_i = o(n)$, then $\Pr(h_{i+1} < 2h_i) \leq \left(\frac{12}{n}\right)^4 = O(\frac{1}{n^4})$. Hence $\Pr(h_{i+1} < 2h_i) = O(\frac{1}{n^4})$. By induction we conclude that for any $i \in \{0, 1, \ldots, \gamma - 1\}$, with failure probability at most $(i+1) \cdot O(\frac{1}{n^4}) = O(\frac{1}{n^3})$, we have $h_{i+1} \geq 2h_i$.

Recall that γ is the smallest integer such that $h_\gamma \geq \frac{n}{8}$. Applying a union bound we have $h_{i+1} \geq 2h_i$ for all $i \in \{0, \ldots, \gamma-1\}$ with failure probability at most $O(\frac{1}{n^2})$. Therefore with probability $1 - O(\frac{1}{n^2})$ we have $\gamma \leq \log n$, which implies $h_{\log n} \geq h_\gamma \geq \frac{n}{8}$. ∎

Lemma 4.3 (Reachability of $\log n$-Neighborhood) *Fix a node $u \in V$. Then with probability $1 - O(\frac{1}{n})$, the following event happens: for each $w \in V \setminus H_{\log n}^u$, there exists a node $v \in H_{\log n}^u$ such that $d(w,v) \leq \log n + 1$.*

Proof: Consider a fixed node w. To find a node $v \in H_{\log n}^u$ with $d(w,v) \leq \log n + 1$, it suffices to show that the intersection $H_{\log n+1}^w \cap H_{\log n}^u \neq \emptyset$. Recall that L_k^w is the set of nodes to which the distance from w is exactly k and H_k^w is the set of nodes to which the distances from w are at most k. Then $h_k^w = h_{k-1}^w + \ell_k^w$ and $h_k^w \geq 2h_{k-1}^w$ implies $\ell_k^w \geq \frac{1}{2}h_k^w$ for $k \geq 1$. Note that each node samples its outgoing neighbors independently. By Lemma 4.2, with probability $1 - O(\frac{1}{n^2})$, the following events happen: (1) $h_{\log n}^u \geq \frac{n}{8}$, and (2) there exists $\gamma_w \leq \log n$ such that $\ell_{\gamma_w}^w \geq \frac{n}{16}$. Assuming (1) and (2), the probability that a node in $L_{\gamma_w}^w$ selects its D neighbors within the set

8

$V \setminus H^u_{\log n}$ is $\binom{n - h^u_{\log n}}{D} / \binom{n-1}{D} = \frac{n - h^u_{\log n}}{n-1} \cdots \frac{n - h^u_{\log n} - D + 1}{n - D} \leq (\frac{n - h^u_{\log n}}{n-1})^D$. Then the probability that $L^w_{\gamma_w + 1}$ contains no nodes in $H^u_{\log n}$ is at most $(\frac{n - h^u_{\log n}}{n-1})^{D \cdot \ell^w_{\gamma_w}} \leq (\frac{7n}{8(n-1)})^{\frac{Dn}{16}} = \exp(-\Theta(n))$ This implies that $H^w_{\log n+1} \cap H^u_{\log n} \neq \emptyset$ with failure probability at most $O(\frac{1}{n^2}) + \exp(-\Theta(n)) = O(\frac{1}{n^2})$. Applying a union bound over all nodes in $V \setminus H^u_{\log n}$ yields the result. $\blacksquare$

The connectivity properties of H in Lemmas 4.2 and 4.3 imply that the graph has $O(\log n)$ diameter.

Proof of Theorem 4.2: Lemma 4.2 and 4.3 states that with probability $1 - O(\frac{1}{n})$, for any two distinct nodes, there exists another node such that can be reached by those two nodes via directed paths of lengths at most $\log n + 1$, which means that the underlying undirected graph of the subgraph H has diameter at most $2(\log n + 1) = O(\log n)$.

Then by Theorem 4.1, the finishing time of streaming B pieces in a directed random graph $G_{n,p}$ with n nodes and $p \geq \frac{6 \log n}{n}$ using COP with $D = 8$ is at most $O(B + \log n)$ with probability $1 - O(\frac{1}{n}) - O(\frac{1}{n}) = 1 - O(\frac{1}{n})$. $\blacksquare$

4.2 COP on Preferential Attachment Graphs

Next we consider the finishing time of multiple-piece streaming in *Preferential Attachment graphs* (PA-graphs). The notion of preferential attachment graphs was first introduced by Barabási and Albert [2], and they have been used to model social and peer-to-peer networks. In this work, we follow the formal definition by Bollobás et al. [3]. Let G^n_m be a PA-graph. We denote by $deg_G(v)$ the degree of a vertex v in G^n_m.

Definition 4.3 (Preferential Attachment Graphs) *Let $m \geq 2$ be a fixed constant parameter. The random graph G^n_m is an undirected graph on the vertex set $V := \{1, \ldots, n\}$ inductively defined as follows.*

1. *G^1_m consists of a single vertex with m self-loops.*

2. *For all $n > 1$, G^n_m is built from G^{n-1}_m by adding the new node n together with m edges $e^1_n = \{n, v_1\}, \ldots, e^m_n = \{n, v_m\}$, inserted one after the other in this order. Let $G^n_{m,i-1}$ denote the graph right before the edge e^i_n is added. Let $M_i = \sum_{v \in V} deg_{G^n_{m,i-1}}(v)$ be the sum of the degrees of all the nodes in $G^n_{m,i-1}$. The endpoint v_i is selected randomly such that $v_i = u$ with probability $\frac{deg_{G^n_{m,i-1}}(u)}{M_i + 1}$, except for n that is selected with probability $\frac{deg_{G^n_{m,i-1}}(n) + 1}{M_i + 1}$.*

It can be easily seen that for $m = 1$ the graph is disconnected w.h.p.; so we focus on the case $m \geq 3$. For each node i and $k \in [m]$, let $\tau_{i,k}$ be the node chosen by i in the k-th rounds after i is added to G^n_m. We call the first node added to G^n_m node v_1. Bollobás et al. [3] introduced an equivalent way of generating G^n_m that is more accessible. In that model, each node i is assigned a random variable $\omega_i \in (0,1)$ following some distribution and edges are added to node i with probability proportional to ω_i. The formal definition of the new model is given in [3]. Under this equivalent model, a node i is called *useful* if $\omega_i \geq \frac{\log^2 n}{n}$ and *non-useful* otherwise.

We apply COP with $D = 3$ to stream B pieces in PA-graph and show that the finishing time in a PA-graph is at most $O(B + \log n)$. Note that the PA-graph is a multi-graph and the outgoing neighbors of a node is a multiset in the subgraph constructed by COP.

Theorem 4.3 (Finishing Time of COP In PA-graphs) *For all $B \geq 1$, the finishing time for streaming B pieces starting from any source node in a PA-graph G^n_m using the constant out-degree protocol with $D = 3$ is at most $O(B + \log n)$ with probability $1 - o(1)$.*

Communication Initiation without Repeat. In each round of the downloading process, each node initiates communication uniformly at random with an outgoing neighbor that is not initiated with in the last $\min\{M, deg_G(u) - 1\}$ rounds; for a multi-set of outgoing neighbors, multiplicities are respected. For $D = 3$ and $M = 2$, this is equivalent to round-robin neighbor selection.

Lemma 4.4 *[8, Lemma 5.8] There exists a property $\mathcal{P}$ of G_m^n that holds with probability $1 - o(1)$ such that conditioning on $\mathcal{P}$, with probability at least $1 - n^{-\Omega(n)}$, using only PUSH on G_m^n (avoiding repeated neighbor initiation from the previous $M \geq 2$ rounds), a piece initiated at a node u reaches a useful node in $O(\log \log n)$ rounds.*

Lemma 4.5 *[8, Lemma 5.11] If each node v in G_m^n receives an independent random variable x_v such that $x_v = 1$ with constant probability p and $x_v = 0$ otherwise, then with probability $1 - o(1)$, for each useful node u, there exists a path of length $O(\frac{\log n}{\log \log n})$ between u and node v_1 such that every second node v in the path has degree m, its two neighbors in the path are $\tau_{v,1}$ and $\tau_{v,2}$ and has $x_v = 1$.*

Note that all x_v's are independent and the randomness of x_v comes from the behavior of node v. In the rest of the section, we assume the PA-graph G_m^n under consideration has the properties stated above.

4.3 Upper Bound for the Stopping Time

First, it is obvious that each node in G_m^n has $deg_G(u) \geq m \geq 3$ and the out-degree of each node in H is $\min\{D, deg_G(u)\} = 3$, for the case $D = 3$. A node v in H is called a *fast* node if $deg_G(v) = m$ and it chooses both $\tau_{v,1}$ and $\tau_{v,2}$ as its outgoing neighbors in H.

In Lemma 4.6, we show that for each non-useful node v, there is at least one useful node u that can be reached from v by a directed path of length $O(\log \log n)$. In Lemma 4.7, we show that all useful nodes are connected to node v_1 by paths of length $O(\frac{\log n}{\log \log n})$ in H and every second node in the paths is fast. These two lemmas imply that the diameter of H is at most $O(\frac{\log n}{\log \log n})$. Given a node u and an integer k, we denote by H_k^u the set of nodes that can be reached by a directed path of length at most k in the subgraph H.

Lemma 4.6 (Useful Node in $O(\log \log n)$-Neighborhood) *With probability $1 - o(1)$, for each non-useful node u, at least one useful node will be included in H_K^u for some $K = O(\log \log n)$ following the constant out-degree protocol.*

Proof: We shall use Lemma 4.4 [8, Lemma 5.8] to prove this result. We condition on the same property $\mathcal{P}$ on G_m^n as in Lemma 4.4. In the proof of [8, Lemma 5.8], the authors consider a restricted version of PUSH strategy on G_m^n with one piece starting at a fixed node $u \in V$. We show that there is a correspondence between their restricted piece streaming process and our outgoing neighbor selection process. Note that in the proof of [8, Lemma 5.8], the authors assume that each node performs only PUSH and each node is active for only 3 rounds after it start downloading. With the memory property, this means the node contacts three random distinct neighbors in those 3 rounds (if the node has degree at least 3); this corresponds to the selection of $D = 3$ outgoing neighbors uniformly at random in COP.

Hence, the single rumor spreading process in [8, Lemma 5.8] can be coupled with our outgoing neighbor selection phase such that if a node v becomes downloading after $3K$ rounds in [8, Lemma 5.8], then there is a path from u to v of length K such that the next node in the path is an outgoing neighbor of the previous node selected by COP. With this coupling argument, we know

10

that conditioning on property $\mathcal{P}$, with probability $1 - n^{-\Omega(n)}$, a useful node will be included in H_K^u for some $K = O(\log \log n)$. By union bound on all nodes in V, the conditional probability that this holds for all H_K^u's is at least $1 - n^{-\Omega(n)} \cdot n$. Since property $\mathcal{P}$ holds with probability at least $1 - o(1)$, our result follows. ∎

Lemma 4.7 (Distances Between Useful Nodes) *With probability $1 - o(1)$, for each useful node u, there exists a path of length $O(\frac{\log n}{\log \log n})$ between u and 1 such that every second node in the path is fast.*

Proof: We show that such paths exist by using Lemma 4.5 with appropriately defined independent random variables x_v's. For each node v, we define x_v to an independent $\{0, 1\}$-random variable that takes 1 with probability $p := \frac{6}{m(m-1)}$. Since the outgoing neighbor selection process is independent of the randomness of the PA-graph, we can couple the randomness of the x_v's with the neighbor selection process as follows.

For a node v with degree not equal to m, we ignore x_v and perform outgoing neighbor selection as before. Suppose node v has degree m; observe that $p = (m - 2)/\binom{m}{3}$ is exactly the probability of the event that both $\tau_{v,1}$ and $\tau_{v,2}$ are chosen as outgoing neighbors of v. Hence, we couple x_v with the outgoing neighbor selection for node v such that $x_v = 1$ *iff* both $\tau_{v,1}$ and $\tau_{v,2}$ are chosen. By Lemma 4.5, with probability $1 - o(1)$, each useful node in G_m^n is connected to v_1 via a path such that every second node v in the path has degree m, its two neighbors in the path are $\tau_{v,1}$ and $\tau_{v,2}$ and has $x_v = 1$. By our coupling argument, this means node v is fast and we have proved the result. ∎

Proof of Theorem 4.3: By Lemma 4.6 and 4.7, we know that with probability $1 - o(1)$, each non-useful node can reach a useful node by a path of length at most $O(\log \log n)$ and all useful nodes are connected to node v_1 by paths of length at most $O(\frac{\log n}{\log \log n})$ in H, which implies that the underlying undirected graph of H has diameter at most $O(\log \log n) + O(\frac{\log n}{\log \log n}) = O(\frac{\log n}{\log \log n})$. By Theorem 4.1, the finishing time of streaming B pieces in a PA-graph G_m^n using COP with $D = 3$ is at most $O(B + \log n)$ with probability $1 - o(1) - O(\frac{1}{n}) = 1 - o(1)$. ∎

4.4 Lower Bound for the Stopping Time

We have shown that streaming B pieces in a PA-graph G_m^n using COP with $D = 3$ requires at most $O(B + \log n)$ rounds. However, the PA-graph and the subgraph H of the PA-graph have diameter $O(\frac{\log n}{\log \log n})$, which is $o(\log n)$. Note that by Fact 4.2, each node v will become finished after $O(B + \frac{\log n}{\log \log n})$ with probability $1 - \exp(-\frac{\log n}{\log \log n})$. However, since the failure probability is $\omega(\frac{1}{n})$ and there are n nodes, we cannot use union bound to show that w.h.p., all nodes will be finished after $O(B + \frac{\log n}{\log \log n})$ rounds. Indeed, we show that the finishing time of streaming B pieces in a PA-graph G_m^n using COP with $D = 3$ needs $\Omega(B + \log n)$ rounds, which implies that the upper bound we have proved in the last section is tight and perfect pipelining is achieved. More specifically, we prove the following result.

Theorem 4.4 (Lower Bound For COP in PA-Graphs) *With probability $1 - o(1)$, COP with any constant parameter D needs $\Omega(B + \log n)$ rounds to get all nodes in a PA-graph G_m^n finished.*

Proof: It is obvious that if the protocol needs $\Omega(\log n)$ rounds to spread 1 piece to all nodes in G_m^n, then $\Omega(B + \log n)$ rounds are required for the same protocol to stream B pieces to all nodes in G_m^n. Hence, we only consider the case when $B = 1$ in the following proof.

By the proof of [8, Theorem 3.3], we know that with probability $1 - o(1)$, there are $\Omega(n)$ edges in G_m^n whose incident nodes are of degree at most $m + c$, where c is a constant. We call those edges *link*s and let $L \subseteq E$ be the collection of links, where $|L| = \Omega(n)$. Note that for any two edges e_1 and e_2 in L, e_1 and e_2 may have common endpoints or the endpoints of e_1 and e_2 may have common neighbors. To avoid this and make the proof simpler, we remove one of those two edges if such a situation happens. Let $\tilde{L}$ be the collection of links after the removals. Since each endpoint of a link is of degree at most $m + c$, we conclude that $|\tilde{L}| \geq \frac{1}{(m+c-1)^2+1}|L| = \Omega(n)$ (to keep one edge, at most $(m + c - 1)^2$ edges will be deleted). Note that any two links in $\tilde{L}$ have no common endpoint and the endpoints of any two edges have no common neighbor.

Fix one link in $\tilde{L}$, denoted by (u, v), u will choose v as an outgoing neighbor with probability at least $\frac{D}{m+c}$ and v will choose u as an outgoing neighbor also with probability at least $\frac{D}{m+c}$, independently in COP. Hence for each of those links, with constant probability, the two endpoints of the link will choose each other as outgoing neighbor in the protocol, independently for all links in $\tilde{L}$. If that happens, we call the link *preserved* in H. By Chernoff bound, we know that with probability $1 - \exp(-\Omega(n))$, there are $\Omega(n)$ preserved links.

Fix one pair of nodes (u, v) that are connected by a preserved link, in each round, if the two nodes are waiting, then with probability at least $(\frac{1}{D})^2(1 - \frac{1}{D})^{2(D-1)} = \delta$, they will remain waiting. Note that δ is a constant. The probability that (u, v) remains waiting for $\frac{\log n}{2 \ln \delta^{-1}}$ is therefore at least $n^{-\frac{1}{2}}$. Since each of those pairs are disjoint and the choices of communications are made independently, the probability that none of the $\Omega(n)$ pairs remains waiting after $\frac{\log n}{2 \ln \delta^{-1}}$ is at most $(1 - n^{-\frac{1}{2}})^{\Omega(n)} \leq \exp(-n^{-\frac{1}{2}}\Omega(n)) \leq \exp(-\Omega(n^{\frac{1}{2}}))$. Hence with probability $1 - o(1) - \exp(-\Omega(n)) - \exp(-\Omega(n^{\frac{1}{2}})) = 1 - o(1)$, COP needs $\Omega(\log n)$ rounds to spread one piece to all nodes in G_m^n and $\Omega(B + \log n)$ rounds to stream B pieces to all nodes in G_m^n using COP with any constant parameter D. ∎

4.5 Improving the Upper Bound

It is established by Doerr et al. [8] that if each node in G_m^n avoid repeated neighbor initiation from the previous $M \geq 2$ rounds, the upper bound of the finishing time of spreading one piece in G_m^n can be improved to $O(\frac{\log n}{\log \log n})$, which is tight up to a constant factor. Note that with $M = 2$ and $D = 3$, the pieces streaming phase of COP is the same as round-robin and by Fact 4.1, the next theorem follows immediately.

Theorem 4.5 *For all $B \geq 1$, the finishing time for streaming B pieces starting from any source node in a PA-graph G_m^n using COP with round-robin pieces streaming phase is at most $O(B + \frac{\log n}{\log \log n})$ with probability $1 - o(1)$.*

5 Lower Bound

In Section 4 we show that perfect pipelining can be achieved using protocols based on the basic PUSH-PULL protocol when streaming multiple pieces in directed random graphs and preferential attachment graphs. In this section we give a class of D-regular graphs, where $\omega(1) \leq D \leq o(n)$, with diameter d and show that perfect pipelining is impossible in those regular graphs for any protocols in which communication partners are chosen uniformly at random in each round. We show that streaming B pieces from one source node in those regular graphs requires at least $\Omega(D(B + d))$ rounds, no matter how the pieces are chosen to be transmitted. We denote those protocols by PUSH-PULL based protocols. Note that by the upper bound proved by [1], our lower bound is

actually tight.

The Necklace Graph. A D-regular necklace graph $N(V,E)$ with n nodes consists of $\ell := \frac{n}{D+1}$ units $U_1, U_2, \ldots, U_\ell$. Each unit U_i for $i \in [\ell]$ contains $D+1$ nodes, numbered by $1, 2, \ldots, D+1$. Within a unit U_i, there is an edge between each pair of nodes except the pair $\{1, D+1\}$. That is, the subgraph of $N(V,E)$ induced by nodes in U_i is $K_{D+1} \backslash 1, D+1$. For each $i \in \{1, 2, \ldots, \ell-1\}$, the two units U_i and U_{i+1} are connected via an edge between node $D+1$ of U_i and node 1 of U_{i+1}. Units U_ℓ and U_1 are connected via an edge between node $D+1$ of $U_{\ell+1}$ and node 1 of U_1.

Theorem 5.1 (Lower Bound for PUSH-PULL) *Let D be an integer such that $\omega(1) \leq D \leq o(n)$. Let $N(V,E)$ be a D-regular necklace graph with $|V| = n$ and diameter d. Then for all $B \geq 1$, the finishing time for streaming B pieces starting from one source node in $N(V,E)$ using any PUSH-PULL based protocol is at least $\Omega(D(B+d))$ with probability at least $1 - o(1)$.*

Since there are $\frac{n}{D+1}$ units in the necklace graph and each unit has diameter 2, the diameter of the graph is $d = \Theta(\frac{n}{D})$. Hence given any $\omega(1) \leq D \leq o(n)$, for example, $D = \sqrt{n}$, the finishing time for streaming B pieces requires at least $\Omega(D(B+d)) = \Omega(B\sqrt{n} + n)$ and perfect pipelining is impossible in this case.

Consider a D-regular necklace graph $N(V,E)$ with $|V| = n$ and the source node u_0. Recall that at the beginning node u_0 has all B pieces and all other nodes are waiting. Also recall that $N(V,E)$ consists of $\ell = \frac{n}{D+1}$ units $U_1, \ldots, U_\ell$. Without loss of generality, suppose $u_0 \in U_1$. Let $v \in V$ be a node such that the distance, i.e. the length of a shortest (undirected) path, between u_0 and v is d. Let k be such that $v \in U_k$, then $k = \Theta(d)$.

Let T_1 be the number of rounds required until the first time node v receives at least one piece, and T_2 the number of rounds required after the T_1 rounds until the first time node v becomes finished. Then the finishing time T satisfies $T \geq T_1 + T_2$. To give lower bounds on T_1 and T_2, we make an assumption that can only speed up the actual pieces streaming process.

Speed-up Assumption. Assume that once a node in unit U_i receives a particular piece for the first time, all nodes in U_i have the same piece immediately. Under this assumption, we can simply ignore intra-unit communication since piece transmissions occur only in inter-unit communication. We can view each unit U_i as a super node. In each round, the unit U_i *independently* initiates an instance of communication to U_{i-1} with probability $\frac{1}{D}$ and an instance of communication to U_{i+1} with probability $\frac{1}{D}$.

Recall that node v is in U_k such that the distance between u_0 and v is d. Let T'_1 be the number of rounds required until the first time unit U_k receives at least one piece, and T'_2 the number of rounds required after the T'_1 rounds until the first time unit U_k receives all B pieces under the speed-up assumption. It follows that $T_1 \geq T'_1$ and $T_2 \geq T'_2$.

Lemma 5.1 (Expected Number of Tosses for m Heads) *Suppose there is a biased coin where the probability of heads is p. Let X_i be the number of tosses made to see the i-th head after seeing the $(i-1)$-th head. Let the total number of tosses made to see m heads be $Z := \sum_{i=1}^{m} X_i$. Then, for any $0 < \epsilon < 1$, we have $\Pr(Z \leq (1-\epsilon)\mathbf{E}[Z]) \leq \exp\left(-\frac{\epsilon^2 m}{8}\right)$.*

Proof: Note that the random variables X_i's follow the geometric distribution with success probability p. Then $\mathbf{E}[Z] = \frac{m}{p}$. Let $t < 0$. Then by Markov's inequality and independence of X_i's, we have that

$$\Pr(Z \leq (1-\epsilon)\mathbf{E}[Z]) \leq (\mathbf{E}[\exp(tX_1)])^m \cdot \exp\left(-\frac{t(1-\epsilon)m}{p}\right).$$

Next we give an upper bound on $\mathbf{E}[\exp(tX)]$. Since $t < 0$, we have

$$\mathbf{E}\big[e^{tX}\big] = \sum_{i=1}^{\infty}(1-p)^{i-1}pe^{ti} = pe^{t}\sum_{i=1}^{\infty}[(1-p)e^{t}]^{i-1} = \frac{p}{p-(1-e^{-t})}$$

$$\leq \frac{p}{p-t} = \frac{1}{1-\frac{t}{p}} \leq 1 + \frac{t}{p} + \frac{2t^2}{p^2} \leq \exp\left(\frac{t}{p} + \frac{2t^2}{p^2}\right).$$

Hence we have

$$\Pr(Z \leq (1-\epsilon)\mathbf{E}[Z]) \leq \exp\left(m\left(\frac{2t^2}{p^2} + \frac{\epsilon t}{p}\right)\right).$$

Setting $t = -\frac{\epsilon p}{4}$ yields $\Pr(Z \leq (1-\epsilon)\mathbf{E}[Z]) \leq \exp\left(-\frac{\epsilon^2 m}{8}\right)$. ∎

Proof of Theorem 5.1: From the above analysis it suffices to give a lower bound on $T_1' + T_2'$ under the speed-up assumption. Note that the necklace graph $N(V, E)$ is a simple cycle with $\ell = \frac{n}{D+1}$ super nodes $U_1, \ldots, U_\ell$.

First we give a lower bound on T_1'. Note that there are two paths between U_1 and U_k, the lengths of which are $\Theta(d)$, where d is the diameter of $N(V, E)$. Without loss of generality suppose unit U_k receives the first piece from U_1 via the path $P = \{U_1, U_2, \ldots, U_k\}$. Note that for U_k to become downloading, at least $k - 1$ times of transmissions (PUSH or PULL) between an downloading unit and an waiting unit are required. Also observe that for $i = \{1, 2, \ldots, k-1\}$, if U_i is downloading and U_{i+1} is waiting, the probability that U_{i+1} becomes downloading in a round is at most $\frac{2}{D}$. Let X_i be a random variable following the geometric distribution with success probability $\frac{2}{D}$. Define $S := \sum_{i=1}^{k-1} X_i$. Then T_1' stochastically dominates S. Since $\mathbf{E}[S] = \frac{D(k-1)}{2}$, by Lemma 5.1, it follows that $\Pr(S \leq \frac{D(k-1)}{4}) \leq \exp(-\frac{k-1}{32}) = o(1)$. Therefore we have $\Pr\left(T_1' > \frac{D(k-1)}{4}\right) \geq \Pr\left(S > \frac{D(k-1)}{4}\right) \geq 1 - o(1)$. Since $k = \Theta(d)$, we have $T_1' \geq \Omega(Dd)$ with probability $1 - o(1)$. If B is a constant, then we already have $T \geq T_1' \geq \Omega(D(B+d))$ with probability $1 - o(1)$.

Suppose $B \geq \omega(1)$. We give a lower bound on T_2'. Recall that T_1' is the first time when U_k becomes downloading. After round T_1', at least $B - 1$ times of communication must be performed between U_{k-1} or U_{k+1} and U_k for U_k to become finished. Suppose, in the best case, that both U_{k-1} and U_{k+1} are finished at time T_1'. Observe that in a round U_k independently pulls a piece from U_{k-1} (or U_{k+1}) with probability at most $\frac{1}{D}$ and is pushed a piece by U_{k-1} (or U_{k+1}) with probability at most $\frac{1}{D}$. For $j \in \{1, 2, \ldots, \frac{B-1}{4}\}$, let Y_j be a random variable following the geometric distribution with success probability $\frac{1}{D}$. Define $S' := \sum_{j=1}^{\frac{B-1}{4}} Y_j$. Then T_2' stochastically dominates S'. Since $\mathbf{E}[S'] = \frac{D(B-1)}{4}$, by Lemma 5.1 it follows that $\Pr(S' \leq \frac{D(B-1)}{8}) \leq \exp(-\frac{B-1}{32}) = o(1)$. Therefore we have $\Pr\left(T_2' > \frac{D(B-1)}{8}\right) \geq \Pr\left(S' > \frac{D(B-1)}{8}\right) \geq 1 - o(1)$. Since $M' = \Theta(DB)$, it follows that $T_2' \geq \Omega(DB)$ with probability at least $1 - o(1)$. Therefore, we have $T \geq T_1' + T_2' \geq \Omega(D(B+d))$ with failure probability at most $o(1)$. ∎

Theorem 5.1 indicates that perfect pipelining is impossible when streaming multiple pieces in necklace graphs using any PUSH-PULL based protocol.

References

[1] David Arthur and Rina Panigrahy. Analyzing bittorrent and related peer-to-peer networks. SODA '06, pages 961–969, New York, NY, USA, 2006. ACM.

[2] A. L. Barabasi and R. Albert. Emergence of scaling in random networks. *Science*, 286:509–512, 1999.

[3] Béla Bollobás and Oliver Riordan. The diameter of a scale-free random graph. *Combinatorica*, 24(1):5–34, January 2004.

[4] Keren Censor-Hillel, Bernhard Haeupler, Jonathan Kelner, and Petar Maymounkov. Global computation in a poorly connected world: fast rumor spreading with no dependence on conductance. STOC '12, New York, NY, USA, 2012. ACM.

[5] Ho-Leung Chan, Tak-Wah Lam, and Prudence W. Wong. Efficiency of data distribution in bittorrent-like systems. AAIM '07, Berlin, 2007. Springer-Verlag.

[6] Supratim Deb, Muriel Médard, and Clifford Choute. Algebraic gossip: a network coding approach to optimal multiple rumor mongering. *IEEE/ACM Trans. Netw.*, 14(SI):2486–2507, June 2006.

[7] Alan Demers, Dan Greene, Carl Hauser, Wes Irish, John Larson, Scott Shenker, Howard Sturgis, Dan Swinehart, and Doug Terry. Epidemic algorithms for replicated database maintenance. PODC '87, New York, NY, USA, 1987. ACM.

[8] Benjamin Doerr, Mahmoud Fouz, and Tobias Friedrich. Social networks spread rumors in sublogarithmic time. STOC '11, New York, NY, USA, 2011. ACM.

[9] U. Feige, David Peleg, P. Raghavan, and Eli Upfal. Randomized broadcast in networks. SIGAL '90, pages 128–137, New York, NY, USA, 1990. Springer-Verlag New York, Inc.

[10] Nikolaos Fountoulakis, Anna Huber, and Konstantinos Panagiotou. Reliable broadcasting in random networks and the effect of density. INFOCOM'10, pages 2552–2560, Piscataway, NJ, USA, 2010. IEEE Press.

[11] George Giakkoupis. Tight bounds for rumor spreading in graphs of a given conductance. volume 9 of *STACS '11*, pages 57–68, Dagstuhl, Germany, 2011. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[12] Bernhard Haeupler. Analyzing network coding gossip made easy. STOC '11, pages 293–302, New York, NY, USA, 2011. ACM.

[13] R. Karp, C. Schindelhauer, S. Shenker, and B. Vocking. Randomized rumor spreading. FOCS '00, pages 565–, Washington, DC, USA, 2000. IEEE Computer Society.

[14] David Kempe, Jon Kleinberg, and Alan Demers. Spatial gossip and resource location protocols. STOC '01, pages 163–172, New York, NY, USA, 2001. ACM.

[15] David Kempe and Jon M. Kleinberg. Protocols and impossibility results for gossip-based communication mechanisms. FOCS '02, pages 471–480, Washington, DC, USA, 2002. IEEE Computer Society.

[16] Boris Pittel. On spreading a rumor. *SIAM J. Appl. Math.*, 47(1), March 1987.