


Smoothed Analysis of Information Spreading in Dynamic Networks

Michael Dinitz 

Johns Hopkins University, Baltimore, MD, USA

Jeremy Fineman 

Georgetown University, Washington, DC, USA

Seth Gilbert 

National University of Singapore, Singapore

Calvin Newport 

Georgetown University, Washington, DC, USA

Abstract

The best known solutions for k -message broadcast in dynamic networks of size n require $\Omega(nk)$ rounds. In this paper, we see if these bounds can be improved by smoothed analysis. To do so, we study perhaps the most natural randomized algorithm for disseminating tokens in this setting: at every time step, choose a token to broadcast randomly from the set of tokens you know. We show that with even a small amount of smoothing (i.e., one random edge added per round), this natural strategy solves k -message broadcast in $\tilde{O}(n + k^3)$ rounds, with high probability, beating the best known bounds for $k = o(\sqrt{n})$ and matching the $\Omega(n + k)$ lower bound for static networks for $k = O(n^{1/3})$ (ignoring logarithmic factors). In fact, the main result we show is even stronger and more general: given ℓ -smoothing (i.e., ℓ random edges added per round), this simple strategy terminates in $O(kn^{2/3} \log^{1/3}(n)\ell^{-1/3})$ rounds. We then prove this analysis close to tight with an almost-matching lower bound. To better understand the impact of smoothing on information spreading, we next turn our attention to static networks, proving a tight bound of $\tilde{O}(k\sqrt{n})$ rounds to solve k -message broadcast, which is better than what our strategy can achieve in the dynamic setting. This confirms the intuition that although smoothed analysis reduces the difficulties induced by changing graph structures, it does not eliminate them altogether. Finally, we apply tools developed to support our smoothed analysis to prove an optimal result for k -message broadcast in so-called well-mixed networks in the absence of smoothing. By comparing this result to an existing lower bound for well-mixed networks, we establish a formal separation between oblivious and strongly adaptive adversaries with respect to well-mixed token spreading, partially resolving an open question on the impact of adversary strength on the k -message broadcast problem.

2012 ACM Subject Classification Theory of computation \rightarrow Distributed algorithms

Keywords and phrases Smoothed Analysis, Dynamic networks, Gossip

Digital Object Identifier 10.4230/LIPIcs.DISC.2022.18

Related Version *Full Version:* <https://arxiv.org/abs/2208.05998> [6]

Funding *Michael Dinitz:* Supported in part by NSF award CCF-1909111.

Jeremy Fineman: Supported in part by NSF grants CCF-1918989 and CCF-2106759.

Seth Gilbert: Supported in part by Singapore MOE grant MOE2018-T2-1-160.

1 Introduction

In this paper, we apply smoothed analysis to the study of k -message broadcast in dynamic networks. We prove that even with a small amount of smoothing, a simple distributed random broadcast strategy can significantly outperform the existing worst-case lower bounds. We then prove that in static networks the complexity of this strategy further improves,



© Michael Dinitz, Jeremy Fineman, Seth Gilbert, and Calvin Newport;
licensed under Creative Commons License CC-BY 4.0

36th International Symposium on Distributed Computing (DISC 2022).

Editor: Christian Scheideler; Article No. 18; pp. 18:1–18:22

Leibniz International Proceedings in Informatics



LIPICs Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

establishing that even in the context of smoothing, changing topologies remain more difficult to move information through than their static counterparts. Finally, we apply the tools developed for these analyses to improve the best-known bounds for k -message broadcast, without smoothing, in the *well-mixed* dynamic network setting. This result is significant in part because when combined with an existing lower bound on well-mixed networks [8], it provides a formal separation between strongly adaptive and oblivious adversaries for k -message broadcast.

1.1 Background

In studying distributed network algorithms, it is common to represent the underlying topology with a graph, where nodes correspond to processes and edges to communication links. In the *dynamic network* setting, these graphs can change from round to round as determined by an adversary. An upper bound proved in a dynamic network is considered strong as it can tolerate the many sources of interference, failure or congestion that alter link availability in real world networks (see [11] for a good review).

Kuhn et al. [12] sparked recent interest in the study of the *k -message broadcast problem*, in which nodes in a network of size n must spread k messages (also called *tokens*) to the whole network. In [12], the results assume the Broadcast CONGEST model in which in each round, each node can broadcast a single bounded-size message, containing at most 1 token. A primary result in the paper is a deterministic algorithm that solves k -message broadcast in $O(nk)$ rounds. For larger values of k , this is notably slower than the $O(n+k)$ rounds required to solve this problem in a static network, underscoring the difficulty of dynamic topologies.

Follow-up work by Dutta et al. [8] proved this result close to optimal with a lower bound that establishes $\Omega(nk/\log n + n)$ rounds are necessary to solve k -message broadcast in this setting. This result is strong in that it holds even for randomized algorithms (with a strongly adaptive adversary), and under the *well-mixed* token assumption in which each token has independent constant probability of starting at each node.

1.2 Key question: is $\tilde{\Omega}(nk)$ fundamental?

Given the importance of information dissemination, an $\tilde{\Omega}(nk)$ lower bound on k -message broadcast is unfortunately strong, especially for large networks attempting to disseminate large amounts of information in a setting with limited bandwidth. Following the approach of Dinitz et al. [7], however, we can investigate whether this bound is fundamental.

In more detail, there are two useful possibilities to consider here. First, this $\tilde{\Omega}(nk)$ bound might be *robust* in the sense that something like nk rounds to broadcast k messages is a natural consequence of network topologies that change. This would be reflected, for example, in the existence of large classes of graphs in which this bound is obviously unavoidable. The second possibility is that the bound is instead *fragile* in the sense that it requires carefully-crafted pathological topologies to induce a complexity of this magnitude, and even small changes to these worst-case graphs enable much more efficient solutions. These distinctions are important because if the $\tilde{\Omega}(nk)$ lower bound due to [8] can be shown to be fragile, this provides hope that more efficient information dissemination can be expected in most real world settings. By contrast, if the bound is robust, this indicates that efficient communication should not be expected in practice.

One approach to distinguishing between robustness and fragility is to apply smoothed analysis. In more detail, in the study of sequential algorithms, Spielman and Teng introduced *smoothed analysis* to help explain why the simplex algorithms works well in practice despite

pessimistic worst-case lower bounds [18, 19]. They proved that the introduction of small random perturbations to otherwise worst-case inputs enabled stronger bounds, indicating the existing lower bound was fragile.

Dinitz et al. [7] subsequently adapted smoothed analysis to the study of distributed algorithms in dynamic networks. In this framework, as in the worst-case setting, an adversary generates an arbitrary dynamic graph to describe the changing network. The individual graphs, however, are then each augmented with ℓ additional random edges, for some smoothing parameter ℓ , before the distributed algorithm in question is run.

For $\ell = 0$, this reduces to the standard worst-case setting where existing lower bounds apply. For $\ell = \binom{n}{2}$, this reduces (more or less) to a random graph setting, in which much stronger upper bounds results are typically possible. As argued in [7], if a worst case lower bound is significantly diminished by smoothed analysis for small ℓ values, then this hints that the original bound is fragile.

The processes and problems studied in [7] were flooding, random walks, and token aggregation. (Follow-up work applied smoothed analysis to the study of the minimum spanning tree [3] and leader election [16] in static graphs.) The k -message broadcast problem features arguably the best-known pessimistic lower bound in the dynamic network setting, but its examination using smoothed analysis was left in [7] as an open problem.

1.3 Our Results

We focus in this paper on *random broadcast*, one of the simplest possible algorithms for disseminating tokens: in each round, each node broadcasts a token chosen uniformly at random from its current token set. This simple strategy will enable us to prove a variety of interesting results on k -message broadcast.

Smoothed analysis of random broadcast in dynamic networks

Applying smoothed analysis as our key tool, we prove that even a small amount of smoothing (i.e., one random edge added per round) is sufficient to enable random broadcast to outperform the worst-case lower bound. This implies both that the existing bound is fragile, and that random broadcast is, in some sense, the *right* strategy for spreading tokens through a dynamic network.

We first establish in Section 5 the baseline result that with no smoothing random broadcast solves the problem in $O(nk)$ rounds, with high probability in n . This matches the deterministic bound from [12].¹ We then investigate the impact of smoothing. In Section 6.2, we show that even with a small amount of smoothing (i.e., $\ell = 1$), random broadcast now terminates in $\tilde{O}(n + k^3)$ rounds, with high probability in n , improving on the best-known $O(nk)$ bound for any $k = \tilde{o}(\sqrt{n})$, and matching the static network lower bound of $\Omega(n + k)$ for $k = \tilde{O}(n^{1/3})$.

We emphasize that 1-smoothing adds at most *one* new edge to the network in each round, which enables at most one extra token dissemination. This smoothing therefore changes the overall bandwidth or connectivity of the graph by only a very small amount.² Given that $\Theta(nk)$ rounds might be necessary to solve k -message broadcast, our speed-up in time

¹ Note that [12] gave a deterministic algorithm for the problem, and also explored the problem of termination detection, which further complicates the problem.

² Notice, for example, that to significantly increase the conductance or vertex expansion of the graph, you would need to add many more edges.

complexity in this context does not come simply from adding large amounts of extra capacity to a worst-case network: most of the work of token dissemination must still occur over the adversarially specified edges in the network. The smoothing accomplishes something more subtle: as we elaborate in our below discussion of predecessor paths, these extra edges are not eliminating bottlenecks in the underlying dynamic network, but instead providing just enough random noise to allow us to bypass their corresponding potential for congestion.

In reality, our result for $\ell = 1$ is proved as a corollary of our more general result (proved in Section 6.1) showing that random broadcast solves k -message broadcast in $O\left(\frac{kn^{2/3} \log^{1/3} n}{\ell^{1/3}}\right)$ rounds, which for $\ell = 1$ is upper bounded by the $\tilde{O}(n + k^3)$ bound claimed above for all n and k . Notice that in this general form, for $k = o(n^{1/3})$, random broadcast actually *beats* the $\Omega(n + k)$ lower bound for static networks. This is possible because even a small number of additional random edges enables tokens to not only bypass bottlenecks, but also skip ahead in temporal paths, reducing the effective dynamic diameter of the network. As we increase ℓ , we get further improvements. For $\ell = k^3$, for example, we get a sub-linear result, as the increased smoothing both speeds up the rate at which tokens initially spread, and the rate at which they subsequently jump over smoothed edges to locations near their destinations in time and space (see below for elaboration).

Predecessor paths. A key technique in our analysis, presented in Section 4, is the use of graph structures that we call *predecessor paths*, which capture paths that exist over time. They are represented as a sequence of node/round pairs, $(u_1, r_1), (u_2, r_2), \dots, (u_x, r_x)$, and for each (u_i, r_i) , for $i < x$, it is guaranteed that u_i is connected to u_{i+1} during round r_i . We further customize these paths for a given token t , strengthening the guarantee for each (u_i, r_i) such that not only will u_i be connected to u_{i+1} in round r_i , but it will broadcast token t in this round, if it knows it.

For each given destination u_x and token t , therefore, we can reduce the problem of delivering t to u_x to the problem of seeding token t into the appropriate predecessor path. (Intuitively, we are establishing here a net over time and space that can capture a token and then inexorably guide it to the center of the trap.) At a high-level, we can therefore break our smoothed analysis of random broadcast into three phases. During the first phase, we ignore the smoothed edges, and allow the natural dynamics of information spread in these networks spread out each token to a larger set. During the second phase, we allow the smoothed edges to seed these tokens onto the appropriate predecessor paths. During the final phase, the tokens can then traverse these paths to their final destinations.

The lengths of these phases are inter-dependent. Increasing the initial spreading phase, for example, decreases the second phase as now each smoothed edge has a higher probability of selecting a node with a useful token. Similarly, relying on long predecessor paths also reduces the second phase, as now each smoothed edge has more targets to which to deliver a token. Longer predecessor paths, however, necessitate a longer third phase to given tokens time to traverse to their destinations. Our final result balances these dependencies by optimizing the time complexity when we fix all three phases to be the same length.

Lower bound. In Section 6.3, we complement our upper bound analysis of random broadcast with a nearly-matching lower bound. In more detail we describe and analyze a *dynamic star* topology in which the network graph forms a star in each round, but the identity of the center node rotates over time. In this setting, we prove random broadcast's expected time to solve k -message broadcast is in $\tilde{\Omega}\left(\min\left(\frac{kn^{2/3}}{(\ell(k+\ell))^{1/3}}, \frac{kn}{k+\ell}\right)\right)$. This result is approximately

a factor of $(k + \ell)^{1/3}$ below our upper bound analysis, confirming that significantly more efficient analyses are not possible. Notably, this bound establishes the fundamental nature of the drop from n to $n^{2/3}$ in the presence of even a small amount of smoothing.

Smoothed analysis of random broadcast in static networks

A possible interpretation of our upper bounds is the following: “with a small amount of smoothing, dynamic networks behave like static networks”. In other words, it might be the case that smoothing removes the differences between dynamic and static networks. In Appendix A, we investigate this issue by studying the behavior of random broadcast in the network topologies that do not change from round to round. We prove that in the presence of minimum smoothing (i.e., $\ell = 1$) random broadcast completes in static network a polynomial factor of n faster than what is possible in dynamic networks. Formally, we prove that in any static network with 1-smoothing random broadcast completes in $\tilde{O}(k\sqrt{n})$ rounds, with high probability. (Recall, the relevant upper bound result in dynamic networks for 1-smoothing is $\tilde{O}(kn^{2/3})$ rounds.) We then prove this analysis is tight (within logarithmic) terms with a matching lower bound.

At the core of our analysis is a decomposition of an arbitrary static network into at most $O(\sqrt{n})$ components each with diameter at most $O(\sqrt{n})$. We demonstrate that given a collection of t components that know the token, in a single spreading interval of $\tilde{O}(k\sqrt{n})$ rounds, each of the t components is likely to send a given target token over a smoothed edge to a unique new component, effectively doubling the number that now know it. We leverage this doubling behavior to spread a token to a large fraction of the network in only a logarithmic number of spreading intervals.

Well-mixed networks

In [8], the authors introduced the notion of a *well-mixed* network in the context of studying k -message broadcast. They call a network well-mixed if for every node u and token t , node u starts with token t with some independent constant probability. Surprisingly, they observe that their $\Omega(nk)$ lower bound holds even for well-mixed networks. It turns out that even starting with a very uniform token distribution does not make the problem easy.

Accordingly, they replace the broadcast communication model with the much more powerful Symmetric-Diff CONGEST model in which each node can not only send a different token on each outgoing edge, but also perform a set-difference with each of their neighbors before deciding what tokens to send. Given this extra power, they show that it is possible to solve k -message broadcast in a well-mixed network in $\tilde{O}(n + k)$ rounds, with high probability.

Leveraging our predecessor path constructions introduced for our smoothed analysis results, in Appendix B we prove, perhaps surprisingly, that our simple random broadcast algorithm solves k -message broadcast in $O((k/p) \log n)$ rounds, with high probability, where p is the probability that each nodes starts with each token. For the $p = \Theta(1)$ case considered in [8], we strictly improve on the bound they achieved in their more powerful communication model. Indeed, for constant p , our bound is within a single log factor of matching a trivial $\Omega(k)$ lower bound for *all* algorithms in the broadcast communication model.

The key follow-up question, of course, is why our result does not violate the $\Omega(nk)$ lower bound from [8]. The difference is found in the adversary assumptions. The existing lower bound requires a *strongly adaptive* adversary that knows the nodes’ random choices in advance and can construct the network topology for a given round based on the knowledge of the tokens nodes are broadcasting in that round. We assume, by contrast, an oblivious adversary that designs the network without advance knowledge of these random bits.

Our well-mixed network result, therefore, opens a clear gap between the strongly adaptive and oblivious adversaries in the context of k -message broadcast. This partially resolves the open question presented in [8] as to whether or not the $\Omega(nk)$ lower bound applies to oblivious adversaries as well.

2 Related Work

Many problems have been studied in various dynamic network models; e.g., [14, 10, 8, 4, 1, 5, 17, 9] (see [11] for a good survey). Interest in k -message broadcast in a dynamic network with broadcast communication was sparked by Kuhn et al. [13], who established the original $O(nk)$ upper bound that provides the baseline for the smoothed analysis deployed in this paper. The relevant matching lower bounds for arbitrary and well-mixed token distributions were subsequently proved by Dutta et al. [8].

Dinitz et al. [7] adapted the smoothed analysis technique, originally introduced by Spielman and Teng [18, 19] in the context of sequential algorithms, to dynamic networks. They studied flooding, random walks and aggregation, and identified k -message broadcast as an important open question. Subsequent work applied this smoothed analysis framework to various other graph problems, including minimum spanning tree construction [3] and leader election [16]. Recently, Meir et al. [15] proposed a variation of graph smoothing, suitable for long-lived processes, in which the smoothing parameter ℓ can be fractional. As noted in [7], smoothed analysis is not the only technique deployed in the literature for sidestepping fragile dynamic network lower bounds. Denysyuk et al. [5], for example, circumvent an exponential lower bound for random walks in dynamic graphs due to [2] by requiring the dynamic graph to include a certain number of static graphs from a well-defined set. In the context of the dynamic radio network model, Ghaffari et al. [9] studied the impact of adversary strength, similarly finding a noticeable gap between oblivious and strongly adaptive adversaries in the context of broadcast.

3 Preliminaries

Here we define the dynamic network models we study and the k -message problem we solve. We also formalize ℓ -smoothing and establish some useful notation and probability results leverage throughout the paper to follow.

Model

We study a dynamic network model in which an execution begins with an oblivious adversary that chooses a *dynamic graph*, defined as a sequence $\mathcal{G} = G_1, G_2, \dots$, where each G_i is a connected graph over a common node set V of size $n = |V|$. Time proceeds in synchronous rounds. At the beginning of each round $r \geq 1$, each node $u \in V$ can reliably broadcast a message to its neighbors in G_r . A key difficulty of these models is that u does not know its neighbors in advance.

k -Message Broadcast

The k -message broadcast problem assumes a set T containing $k \geq 1$ unique messages that are also commonly called *tokens* or *rumors*. Each rumor in T starts the execution at one or more nodes. The problem is solved once all nodes have received all k rumors in T . Following the standard convention [12], we assume each node can broadcast at most 1 rumor per round.

ℓ -Smoothing

Fix a dynamic graph $\mathcal{G} = G_1, G_2, \dots$. Fix a smoothing parameter $\ell \geq 1$. Our goal is to define a smoothing process that adds ℓ random edges to each G_i in \mathcal{G} . In an effort to maximize generality, the original definition of ℓ -smoothing from [7] assumed that each G_i was replaced by a graph \hat{G}_i sampled uniformly from the set of graphs that are both “allowable” and within edit distance k of G_i . This was meant to allow both additions and deletions while avoiding illegal topologies (e.g., disconnected graphs).

The results for the Broadcast CONGEST model in [7], however, largely avoided much of this generality, instead applying results (Lemmas 4.1 and 4.2) that establish that this model approximates a simpler model in which edges are randomly added from the set of all edges. For the sake of clarity, in this paper we directly deploy this simpler definition of smoothing.

Formally, after the adversary generates \mathcal{G} , we smooth each G_i as follows: (1) randomly generate ℓ edges (with replacement); (2) for each such edge, if it is not already in the smoothed graph we are generating, add it to the graph.

Notation

In the following, we use \tilde{O} , $\tilde{\Theta}$, and $\tilde{\Omega}$ to suppress logarithmic factors with respect to n . When we specify a result holds *with high probability*, we mean with failure probability upper bounded by n^{-x} for some sufficiently large constant $x \geq 1$. We also use $[x]$, for integer $x \geq 1$, to indicate the set $\{1, 2, \dots, x\}$.

Fix an execution of a k -message broadcast algorithm for some token set T of size k and node set V . For each node $u \in V$ and round $r \geq 1$, let $T_u(r)$ be the set of tokens u started with or received by the beginning of round r . We say u *knows* the tokens in $T_u(r)$ at the beginning of round r . Finally, for a given token $t \in T$, let $n_t(r) = |\{u \mid t \in T_u(r)\}|$ be the number of nodes that know token t at the beginning of r .

Useful Probability Results

Many of our high probability results that follow leverage the following useful form of a Chernoff Bound:

► **Theorem 1.** *Let X_1, \dots, X_j be a series of independent random variables such that $X_i \in [0, 1]$ where $X = \sum_{i=1}^j X_i$ has expectation $E[X] = \mu$. For $\varepsilon \in [0, 1]$, $\Pr[X \leq (1 - \varepsilon) \cdot \mu] \leq \exp(-(1/2) \cdot \varepsilon^2 \mu)$.*

In several places in our analysis, we tame correlated random variables by applying the following stochastic dominance result, which generalizes the above concentration bound. It says that if the probability that the i 'th variable is 1 is at least p no matter how the first $i - 1$ variables are realized, then we can assume that we have independent Bernoulli variables with parameter p .

► **Lemma 2.** *Let X_1, \dots, X_j be j random variables (not necessarily independent), each of which is distributed over $\{0, 1\}$. Suppose there is some $p \in [0, 1]$ such that for all $i \in [j]$ and for all $x_1, x_2, \dots, x_{i-1} \in \{0, 1\}$,*

$$\Pr[X_i = 1 \mid X_k = x_k \forall 1 \leq k < i] \geq p.$$

Then

$$\Pr\left[\sum_{i=1}^j X_i \leq (1 - \varepsilon)pj\right] \leq \exp(-(1/2) \cdot \varepsilon^2 pj)$$

Proof. Consider the following process for sampling random variables $\hat{X}_1, \dots, \hat{X}_j$. For $i = 1$ to j , do the following. Let $x_1, \dots, x_{i-1} \in \{0, 1\}$ be the values of $\hat{X}_1, \dots, \hat{X}_{i-1}$ respectively, and let $q = \Pr[X_i = 1 \mid X_k = x_k \forall 1 \leq k < i]$. Note that by the assumption of the lemma, $q \geq p$. Sample an independent Bernoulli random variable Y_i which is 1 with probability p and is 0 otherwise. If $Y_i = 1$ then set $\hat{X}_i = 1$. If $Y_i = 0$, then set $\hat{X}_i = 1$ with probability $(q - p)/(1 - p)$ and otherwise set $\hat{X}_i = 0$.

It is easy to see that $\hat{X}_1, \dots, \hat{X}_j$ and X_1, \dots, X_j have identical joint distributions. More formally, let $x_i \in \{0, 1\}$ for all $i \in [j]$. Then

$$\begin{aligned} \Pr[\hat{X}_i = x_i \forall i \in [j]] &= \prod_{i=1}^j \Pr[\hat{X}_i = x_i \mid \hat{X}_k = x_k \forall k < i] \\ &= \prod_{i=1}^j \begin{cases} p + (1-p) \frac{\Pr[X_i=1 \mid X_k=x_k \forall 1 \leq k < i] - p}{1-p} & \text{if } x_i = 1 \\ (1-p) \frac{1 - \Pr[X_i=1 \mid X_k=x_k \forall 1 \leq k < i]}{1-p} & \text{if } x_i = 0 \end{cases} \\ &= \prod_{i=1}^j \Pr[X_i = x_i \mid X_k = x_k \forall 1 \leq k < i] \\ &= \Pr[X_i = x_i \forall i \in [j]] \end{aligned}$$

By the definition of \hat{X}_i , we also have the property that $Y_i \leq \hat{X}_i$ for all i . Hence Theorem 1 implies that

$$\begin{aligned} \Pr \left[\sum_{i=1}^j X_i \leq (1 - \varepsilon)pj \right] &= \Pr \left[\sum_{i=1}^j \hat{X}_i \leq (1 - \varepsilon)pj \right] \leq \Pr \left[\sum_{i=1}^j Y_i \leq (1 - \varepsilon)pj \right] \\ &\leq \exp(-(1/2) \cdot \varepsilon^2 pj) \end{aligned}$$

as claimed. \blacktriangleleft

4 Random Broadcast Predecessor Paths

Several of the results that follow are built on a structure that we call *predecessor paths*, which are defined with respect to both a given dynamic graph \mathcal{G} and the collection of random bits that determine the choices during a given execution of the random broadcast algorithm. We define and analyze these structures in a general way here. We will later deploy these results to prove specific bounds on random broadcast.

To formally define a predecessor path, we first introduce the notion of a bit assignment \mathcal{B} to be a function $\mathcal{B} : V \times \mathbb{Z}_{>0} \rightarrow \{0, 1\}^*$, where $\mathcal{B}(u, r)$ are the random bits node u uses to make its choice of which token to broadcast in round r of running random broadcast. Notice, the combination of a dynamic graph \mathcal{G} and bit assignment \mathcal{B} , does not by itself fully specify an execution of random broadcast, as knowledge of the initial token assignment is also required. This information, however, is sufficient for our formal definition.³

► Definition 3. Fix a dynamic graph \mathcal{G} defined over node set V , token set T , target node $u \in V$, target token $t \in T$, round pair r, r' , with $1 \leq r < r'$, and bit assignment \mathcal{B} . A predecessor path $P_{u,t}(r, r')$ for these parameters is a node/round sequence $(u_1, r_1), (u_2, r_2), \dots, (u_h, r_h)$, where $r \leq r_1 < r_2 < \dots < r_h \leq r'$, and $u_i \neq u_j$ for $i, j \in [h]$, $i \neq j$, that satisfies the following with respect to the execution of random broadcast in \mathcal{G} according to bit assignment \mathcal{B} :

³ A brief aside is that although we call these objects *paths*, the definition actually captures a more general in-tree structure in the time expansion graph.

1. For each $i \in [h - 1]$: if $t \in T_{u_i}(r_i)$, then u_i will broadcast t in round r and it will be received by at least one node u_j , for $j > i$.
2. If $t \in T_{u_h}(r_h)$, then u_h will broadcast t during round r_h and it will be received by node u .

A natural corollary of this definition is that if *any* node u_i in a predecessor path $P_{u,t}(r, r')$ learns t by r_i , then u will learn t by r' .

Our goal is to describe and analyze a procedure for generating a predecessor path for a given set of parameters. We will then analyze the expected length of the paths created. Roughly speaking, when studying random broadcast, the longer a predecessor path the better, as it gives more opportunities for a given token to arrive at a node that can then send it on its way to the desired destination.

Preliminaries

As discussed, we will be analyzing the simple algorithm in which every node broadcasts a token that it knows uniformly at random. To ease the analysis, though, we assume without loss of generality that (1) the tokens are labelled from 1 to k and that nodes know k ; and (2) that a node randomly selects a token to send in a given round by randomly permuting the values from 1 to k and then broadcasting the first token from this sequence that it possesses. Note that this gives the exact same process as the randomized algorithm we care about, which is why this is without loss of generality. It allows us, however, to fix the random process for how a token is chosen even when the available set of tokens to send is unknown.

For a given node set V , token set T , bit assignment \mathcal{B} , and round $r \geq 1$, let the *primary token* for u in r , indicated $\delta_u(r)$, be the first token id in u 's random permutation for this round as determined by \mathcal{B} . In the practical setting where u permutes all values from 1 to \hat{k} , then the primary token is the first value in this permutation that correspond to an actual token in T . The important property of a primary token is if $\delta_u(r) = t$, then we know that if $t \in T_u(r)$, node u will broadcast t in this round.

Given a dynamic graph $\mathcal{G} = G_1, G_2, \dots$, a non-empty node subset $S \subset V$, and a round $r \geq 1$, we further define the *predecessor cut* $c(S, r)$ to be the set of nodes in $V \setminus S$ that neighbor nodes in S in G_r . Notice, because each graph is connected and S is a proper subset of V , these cut partitions are always non-empty.

Predecessor Path Construction

We now describe how to construct a predecessor path for a given set of parameters. This construction process works backwards in time from the end of the desired interval to the beginning. While we describe this process algorithmically, we emphasize that this algorithmic construction is used only in the *analysis* of our algorithms.

In the following, we assume a fixed dynamic network $\mathcal{G} = G_1, G_2, \dots$, defined over some node set V of size at least 2, a token set T of size $k \geq 1$, and a fixed bit assignment \mathcal{B} for the nodes in V to run random broadcast. We then parameterize the construction process with a node $u \in V$, token $t \in T$, and round range $1 \leq r < r'$. It returns a predecessor path $P_{u,t}(r, r')$ for these parameters.

We now analyze the paths constructed by this procedure, showing establishing the relationship between interval length ($r' - r$) and path length.

► **Theorem 4.** Fix a dynamic graph \mathcal{G} defined over node set V of size $n > 1$, token set T of size $k \geq 1$, random bit assignment \mathcal{B} for V , and error exponent integer $x > 0$. For every $u \in V$, $t \in T$, and rounds r, r' where $r' - r = z \geq 8xk \ln n$:

1. The sequence $P_{u,t}(r, r')$ produced by Path-Construction is a predecessor path.
2. With probability at least $1 - n^{-x}$: $|P_{u,t}(r, r')| > \frac{z}{2k}$.

■ **Algorithm 1** Path-Construction(u, t, r, r').

```

 $P_{u,t}(r, r') \leftarrow \epsilon;$ 
 $i \leftarrow r';$ 
 $S \leftarrow \{u\};$ 
while  $i > r$  do
     $S_{i-1} \leftarrow c(S, (i-1));$ 
     $S_{i-1}^{(t)} \leftarrow \{v \mid v \in S_{i-1} \wedge \delta_v(i-1) = t\};$ 
    if  $|S_{i-1}^{(t)}| > 0$  then
        fix any  $v$  in  $S_{i-1}^{(t)}$ ;
         $S \leftarrow S \cup \{v\};$ 
        append  $(v, i-1)$  to the front of  $P_{u,t}(r, r')$ ;
    end
     $i \leftarrow i-1;$ 
end
return  $P_{u,t}(r, r')$ 

```

Proof. Fix values for the parameters specified and constrained in the theorem statement. Consider the sequence $P_{u,t}(r, r')$ produced by the Path-Construction algorithm for these parameters. By the definition of this algorithm, $P_{u,t}(r, r')$ is a valid predecessor path. We turn our attention to bounding its size.

At each iteration of the main loop in the procedure, if $|S| < n$, then S_{i-1} is non-empty, as $V \setminus S$ is non-empty and G_{i-1} is connected. Fix some $u \in S_{i-1}$. This node is included into $S_{i-1}^{(t)}$ only if $\delta_v(i-1) = t$, which occurs with probability exactly $1/k$. Therefore, the probability that our path expands in this iteration in the case that $|S| < n$ is at least $1/k$ (it could be larger if there are multiple nodes in S_{i-1}).

For each round $i \in [r, r']$ in the interval, let X_i be the random indicator variable that evaluates to 1 if one of the following conditions holds: (1) $|S| = n$; or (2) $|S| < n$ and $P_{u,t}(r, r')$ grows during the iteration corresponding to round i . Let $Y = \sum_{i \in [r, r']} X_i$. Clearly, Y is an upper bound on the size of $P_{u,t}(r, r')$ returned by the path construction procedure. As we argued that $\Pr(X_i = 1) \geq 1/k$ for each i , we can apply Lemma 2 for these z random variables, $p = 1/k$, and $\varepsilon = 1/2$, to derive $\Pr[Y \leq (1/2) \frac{z}{k}] \leq \exp(-\frac{z}{8k})$. Given our assumption that $z \geq 8xk \ln n$, this error bound is upper bounded by $\exp(-x \ln n) = n^{-x}$, as needed. ◀

5 Random Broadcast in Worst-Case Networks

We begin by showing that in the absence of any additional assumptions, random broadcast matches the best known bound of $O(nk)$ rounds for solving k -message broadcast. The sections that follow will then improve on this baseline. The intuition for this result is straightforward: if token t is not fully spread by round r there is at least one edge over which it could spread with probability at least $1/k$. A union bound and stochastic dominance argument are deployed in the following proof to dispatch dependency and varying token set size issues, respectively.

► **Theorem 5.** Fix a dynamic network \mathcal{G} of size n . Fix a rumor set T of size $k \leq n$. With high probability: random broadcast solves k -message broadcast in $O(nk)$ rounds in \mathcal{G} .

Proof. Fix a token t and round r . If $n_t(r) < n$ then there is at least one edge in the network crossing the cut between nodes that do and nodes that do not know t . Token t is selected for broadcast across this edge with probability at least $1/k$. Let X_r be the random variable that evaluates to 1 under the following two conditions: (1) $n_t(r) = n$; or (2) $n_t(r) < n$ and at least one new node learns t .

Clearly, regardless of the execution, for every r , $\Pr(X_r = 1) \geq 1/k$. Let $Y = \sum_{r=1}^{\alpha nk} X_r$, for a constant $\alpha \geq 1$ that we will fix later. We can apply Lemma 2 for $p = 1/k$, $j = \alpha nk$, and $\varepsilon = 1/2$ to derive that $\Pr[Y \leq (1/2) \cdot (1/k) \cdot (\alpha nk)] \leq \exp(-(1/8) \cdot \alpha n)$.

Notice, due to the large size of the expectation, for $\alpha \geq 8$, this bound gives us a failure probability exponentially small in n . Clearly then, there is a sufficiently large constant α such that this failure probability is less than or equal n^{-2} , allowing us to apply a union bound over all $k \leq n$ tokens to establish that with high probability: *every* token spreads to all nodes in αnk rounds. ◀

6 Random Broadcast in Smoothed Networks

We now consider the random broadcast algorithm when the initial token distributions and network topologies are arbitrary. In this worst-case setting, as mentioned, the best known k -message broadcast solutions require nk rounds, a bound which is known to be tight within log-factors under certain adversary assumptions. In the previous section, we showed that random broadcast matched this bound as well. The goal of this section is to analyze random broadcast under smoothed analysis.

We show here that if we run the simple random broadcast algorithm on an ℓ -smoothed dynamic network (with $\ell > 0$) then with high probability it solves k -message broadcast after only $O\left(\frac{kn^{2/3} \log^{1/3} n}{\ell^{1/3}}\right)$ rounds. Note that even in the case of little smoothing, e.g., $\ell = 1$, this bound represents nearly an $n^{1/3}$ -factor improvement to the round complexity captured by the worst-case bound. To simplify comparison to existing k -message broadcast results, as well as the offline result proved later in this paper, we also prove that for $\ell = 1$, this complexity is upper bounded by $O(n + k^3 \log n)$.

Finally, we note that following the approach of [7], we study only integer ℓ values. As recently demonstrated in [15], fractional smoothing parameters, $0 < \ell < 1$, can also be studied to capture behavior in long-lived networks with slow changes. Our results naturally extend to this case, though we omit this analysis for the sake of clarity.

6.1 Random Broadcast with ℓ -Smoothing

In this section, we look at the case where there are ℓ smoothed edges added per round. Even when $\ell = 1$, we get significant improvements, despite the fact that one edge only enables at most one additional token be transferred per round; thus any non-trivial advantage conveyed by smoothing does not come from directly increasing the capacity of the underlying network. Our goal is to prove the following:

► **Theorem 6.** *Fix any dynamic network \mathcal{G} of size n . Fix any rumor set size $k \leq n$. With high probability, random broadcast solves k -message broadcast in $O\left(\frac{kn^{2/3} \log^{1/3} n}{\ell^{1/3}}\right)$ rounds in \mathcal{G} with ℓ -smoothing.*

Our proof proceeds in phases. Note that the algorithm is the same throughout, but our analysis focuses on different quality guarantees in each phase. For the lemmas that follow, assume we have fixed a dynamic graph \mathcal{G} , size n , and rumor set T of size k , as specified in the theorem.

6.1.1 Phase #1: Spread

The initial distribution of tokens is arbitrary, and tokens may be located at very few nodes initially. The goal of this first phase is to argue that after enough time tokens are spread out sufficiently across the network, specifically spreading to a δ fraction of nodes. The parameter δ is a function of n and k that we shall set later to balance the length of all phases. We show here that $\Theta(k\delta n)$ rounds suffice to spread all tokens to δn nodes. Notice that if $\delta = 1$, this follows from Theorem 5.

► **Lemma 7.** *For any constant $x \geq 1$ and fraction δ with $(1/n) \ln n \leq \delta \leq 1$, there exists a constant $c_1 \geq 1$ such that with probability at least $1 - n^{-x}$: for all $t \in T$, $n_t(R) \geq \delta n$, where $R = c_1 k \delta n$ is the duration of the phase.*

Proof. Fix a given token $t \in T$ and round $r \leq R$. Let X_r be the indicator random variable that evaluates to 1 under two conditions: (1) $n_t(r) \geq \delta n$ (recall that $n_t(r)$ is the number of nodes that know token t at the beginning of round r); or (2) a node learns t for the first time during round r . If the first condition does not hold then not all nodes know t at round r , and hence there is at least one edge connecting a node u that knows t to a node v that does not because the network is assumed to be connected. By the definition of random broadcast, u selects t to broadcast with probability $1/|T_u(r)| \geq 1/k$. It follows that regardless of the execution through the first $r - 1$ rounds: $\Pr(X_r = 1) \geq 1/k$.

Let $Y = \sum_{r=1}^R X_r$. We can apply our stochastic dominance result (Lemma 2) to $p = 1/k$, $j = R$, and $\varepsilon = 1/2$ to derive the following:

$$\begin{aligned} \Pr[Y \leq (1/2) \cdot (1/k) \cdot R] &\leq \exp\left(-\frac{1}{8} \cdot \frac{R}{k}\right) = \exp\left(-\frac{1}{8} \cdot c_1 \delta n\right) \\ &\leq \exp\left(-\frac{1}{8} \cdot c_1 \ln(n)\right) = n^{-c_1/8}. \end{aligned}$$

It is not hard to see that as long as $c_1 \geq 2$, then $Y \geq (1/2) \cdot (1/k) \cdot R$ implies that $n_t(R) \geq \delta n$. This is because otherwise, it must be the case that every X_r which is equal to 1 is because a new node learned token t at round r (not because $n_t(r) \geq \delta n$). But this implies that $n_t(R) \geq Y \geq (c_1/2)\delta n \geq \delta n$.

So if we set $c_1 = 8(x + 1)$, we get that the probability that $n_t(R) < \delta n$ is at most $n^{-(x+1)}$. A union bound over all $k \leq n$ tokens provides that $n_t(R) \geq \delta n$ for every $t \in T$ with probability at least $1 - n^{-x}$, proving the lemma. ◀

6.1.2 Phase #2: Seed

Now that we have spread each token to $\Omega(\delta n)$ nodes, we next rely on the edges added by smoothing to help sparsely seed these tokens to new random nodes in the network. In the third and final phase we will show that this seeding is likely to have foiled the adversary's attempts to keep certain nodes isolated from certain tokens, and will have instead planted seeds sufficiently close on a temporal path to arrive their destinations.

In more detail, our goal is to show that, for some parameter γ , any sufficiently large set S of size at least $\ln n/\gamma$ will have at least one node in the set receive a given token with high probability during the seed phase. This phase will last for $\Theta((\gamma/\delta)kn)$ rounds following the conclusion of the spread phase.

For example, consider $\delta = 1/k$ and $\gamma = 1/k^2$. Then the length of both the spread phase and the seed phase are $\Theta(n)$ rounds. During the seed phase with these parameters, each node has probability of at least $1/k^2$ of receiving a specific token under consideration. Thus

any set of size $k^2 \ln n$ will receive the token with high probability. Note that these are not the best choices of δ and γ . (With these choices, the total number of rounds including the sink phase is $O(n + k^3 \log n)$ by Lemma 9.)

Our analysis of the seed phase focuses almost entirely on the smoothed edges added to the network topology graph in each round.

► **Lemma 8.** *Consider any constant $x \geq 1$ and positive fractions δ and γ . Fix any token $t \in T$, node set $S \subseteq V$ with $|S| \geq (1/\gamma) \ln n$, and round $r_0 \geq 1$ such that $n_t(r_0) \geq \delta n$. With probability at least $1 - n^{-x}$: $|S \cap \{u \mid t \in T_u(r_0 + 2xR)\}| > 0$, where $R = (\gamma/\delta)kn/\ell > k \ln n$ and $2xR$ is the length of the phase.*

Proof. By assumption there are at least δn nodes that know t by the start of this phase (round r_0). If any node in S already knows t at the beginning of this phase, then we are already done. So moving forward, assume no node in S knows t at the start of this phase. Consider the set of potential edges that are useful, denoted E_{useful} , that would connect a node from the initial set that knows t to a node in S . It follows $|E_{\text{useful}}| \geq (\delta n)|S| \geq (\delta n)(\ln n/\gamma) = kn^2 \ln(n)/(R\ell)$. We now calculate, for a given round of phase 2, the probability that a smoothed edge selected is from E_{useful} . To do so we leverage the specific definition of smoothing established in Section 3 that treats this as a purely additive process: select a random edge from all possible edges; if it is not already present in the graph, add it to the graph; otherwise leave the graph the same. Therefore, this probability is:

$$\frac{|E_{\text{useful}}|}{\binom{n}{2}} > \frac{|E_{\text{useful}}|}{n^2} \geq (k \ln n)/(R\ell).$$

As there are ℓ smoothed edges in a round, the probability that none of them are useful is $(1 - k \ln n/(R\ell))^\ell \leq e^{-k \ln n/R} \leq 1 - k \ln n/(2R)$.

We call a round in phase 2 *good* if an edge from E_{useful} is selected *and* the endpoint that knows t selects t to broadcast. Because this latter selection event happens with probability at least $1/k$, we can lower bound the probability of a round being good as $p_{\text{good}} \geq \ln n/(2R)$. If a round is not good we call it *bad*. Assume phase 2 runs for $2xR$ rounds for constant $x > 0$. Then the probability that every round is bad is bounded by:

$$(1 - p_{\text{good}})^{2xR} \leq \left(1 - \frac{\ln n}{2R}\right)^{2xR} < \exp(-x \ln n) = n^{-x},$$

as required by the lemma statement. ◀

6.1.3 Phase #3: Sink

In the second phase, we leveraged smoothed edges to sparsely seed each token throughout the network in a manner that is independent of the adversary's construction of the dynamic graph. In this final phase, we deploy our predecessor path constructions to show it is likely for each token t and destination node u , that t arrived at an appropriate location in both time and topology to subsequently make its way to u like a flow heading toward a sink (hence the phase name). In particular, we simply need the final phase to be long enough to achieve a predecessor path of size $\Omega(\ln n/\gamma)$. Putting this piece together with the previous phases allows us to prove the following lemma, which almost immediately implies Theorem 6.

► **Lemma 9.** *Fix any dynamic network \mathcal{G} of size n . Let δ and γ be any values satisfying $(1/n) \ln n \leq \delta \leq 1$ and $0 < \gamma \leq 1$. Then with high probability, random broadcast completes k message broadcast with ℓ -smoothing in $O(k\delta n + (\gamma/\delta)kn/\ell + k \ln n/\gamma)$ rounds.*

Proof. Our analysis below makes use of a constant $x \geq 1$ that we will fix later. Lemma 7 tells us that there exists a constant $c_1 \geq 1$, such that with probability least $1 - n^{-x}$, for each token $t \in T$, at least δn nodes know t by round $c_1 k \delta n$, for some integer $c_1 > 0$. Let us call this the *spread condition*.

Before we apply the seed phase to each token, we need to identify the sets we are attempting to seed. So we look ahead to the sink phase. Let r_S indicate the first round of the sink phase, which we will calculate later based on the duration of the other two phases. We run this phase for $z = \lceil 8xk \ln n / \gamma \rceil$ rounds. That is, it runs between rounds $r = r_S$ and $r' = r_S + z$. Theorem 4 tells us that with probability at least $1 - n^{-x}$: for every node $u \in V$ and token $t \in T$, the resulting predecessor path $P_{u,t}(r, r')$ is of size greater than $\frac{z}{2k} > \ln n / \gamma$.

This allows to apply our seed phase (Lemma 8) analysis to each such predecessor path. This tells us that so long as the spread condition holds, for each such $S = P_{u,t}(r, r')$, the probability that some node in S receives t during the seed phase is at least $1 - n^{-x}$. The duration of the seed phase is $2x(\gamma/\delta)kn/\ell$ rounds.

Finally, we note that by the definition of a predecessor path, if a node in $P_{u,t}(r, r')$ receives a token t by round r_S , then u receives t by round at most $r_S + z$. We are left to pull together the pieces. To do so, we note that success in solving k -message broadcast by the end of the sink phase requires the following events to occur:

1. The spread condition holds. Call this E_{spread} .
2. For every $u \in V$ and $t \in T$, the predecessor path $P_{u,t}(r, r')$ is sufficient long. Call this $E_{pred}(u, t)$.
3. For each destination node $u \in V$ and token $t \in T$, at least one node in $P_{u,t}(r, r')$ receives t during the seed phase. Call this $E_{seed}(u, t)$.

By our above analysis the probability E_{spread} fails is at most n^{-x} , and the probability $E_{pred}(u, t)$ fails for *any* u and t , as also upper bounded by n^{-x} . For $E_{seed}(u, t)$, the probability of failure for each specific pair u and t , conditioned on E_{spread} and $E_{pred}(u, t)$, is upper bounded by n^{-x} . Therefore, by a union bound, the probability that E_{seed} fails for any such pair is less than $n^{-(x+2)}$. A final union bound then provides that probability any of these events fail is less than $n^{-x} + n^{-x} + n^{-(x+2)} < 1/n$ for a sufficiently large constant x .

Plugging this constant value of x into our above round complexity bounds, and we get that random broadcast succeeds with high probability in $c_1 \delta kn + 2x(\gamma/\delta)kn/\ell + \lceil 8xk \ln n / \gamma \rceil = O(\delta kn + (\gamma/\delta)kn/\ell + k \ln n / \gamma)$ total rounds, as claimed in the theorem. ◀

Proof (of Theorem 6). Choose $\delta = (\log n / (n\ell))^{1/3}$ and $\gamma = (\ell^{1/3})(\log n / n)^{2/3}$. It follows that $\gamma/\delta = (\ell^{2/3})(\log n / n)^{1/3}$. Then by Lemma 9 we get that broadcast completes in at most

$$\begin{aligned} & O((\log n / (n\ell))^{1/3} kn + (\log n / n)^{1/3} kn / \ell^{1/3} + k \log n (n / \log n)^{2/3} / \ell^{1/3}) \\ & = O((kn^{2/3} \log^{1/3} n) / \ell^{1/3}) \end{aligned}$$

rounds, as claimed. ◀

6.2 1-Smoothing

The following result for 1-smoothing is a simple corollary, which allows us to more easily compare to existing results.

► **Corollary 10.** *Fix any dynamic network \mathcal{G} of size n . Fix any rumor set size $k \leq n$. With high probability, random broadcast solves k -message broadcast in $O(n + k^3 \log n)$ rounds in \mathcal{G} with 1-smoothing.*

Proof. Theorem 6 implies that random broadcast takes at most $O(kn^{2/3} \log^{1/3} n)$ rounds. It is easy to see that $kn^{2/3} \log^{1/3} n$ is at most $n + k^3 \log n$ for all values of k . Alternatively, we can set $\delta = 1/k$ and $\gamma = 1/k^2$, and apply Lemma 9 with $\ell = 1$. ◀

6.3 Lower Bound for Random Broadcast in Smoothed Networks

In this section we prove the following theorem.

► **Theorem 11.** *For all $n, k, \ell \geq 1$, there are dynamic networks on n nodes and a starting token distribution of k tokens such that random broadcast with ℓ -smoothing has expected completion time of at least $\tilde{\Omega}\left(\min\left(\frac{kn^{2/3}}{(\ell(k+\ell))^{1/3}}, \frac{kn}{k+\ell}\right)\right)$.*

In most “reasonable” regimes (k and ℓ not overwhelmingly large) the minimum in the above lower bound will be achieved by $\frac{kn^{2/3}}{(\ell(k+\ell))^{1/3}}$. Note that this almost matches the upper bound of Theorem 6: it is off by just a $(k+\ell)^{1/3}$ factor. In addition, we also note that when ℓ is large the upper bound *cannot* be tight, while our lower bound can be. To see this, consider the case where $\ell = n$ and $k = o(n)$. For these parameters, essentially every node is the endpoint of an edge added by smoothing in every round, and for each token the probability of broadcasting it is at least $1/k$, and hence for every token the number of nodes who know it will double in at most $O(k)$ rounds. Thus random broadcast will complete after only $O(k \log n)$ rounds. For this case, where $\ell = n$ and $k = o(n)$, our lower bound from Theorem 11 correctly reduces to $\tilde{\Omega}(k)$, while the upper bound from Theorem 6 remains at $\tilde{O}(kn^{1/3})$. This hints that the modest gap between our upper and lower bounds might ultimately be resolved to be closer to the latter result.

6.3.1 Proof of Theorem 11

Our lower bound instance will be the *dynamic star*. The vertices are v_0, \dots, v_{n-1} , and at time $i \in \{1, 2, \dots, n\}$ the graph will be a star with v_i at the center. Initially node v_0 knows all of the tokens, while nodes v_1, \dots, v_{n-1} know all of the tokens *except* for token 1. So random broadcast is complete once all nodes know token 1. Note that this graph is not defined for more than n rounds, but the lower bound that we are trying to prove is at most n due to the second term in the min, so we will not need to consider rounds past n .

Let $t = \min\left(\frac{kn^{2/3}}{(\ell(k+\ell))^{1/3}}, \frac{kn}{k+\ell}\right) / (2000 \log n)$ (we have not optimized the constant or log factors). We will argue that with constant probability, random broadcast has not completed by time t .

Let $A = \{v_i : 1 \leq i \leq t\}$. For $v_i \in A$, we say that v_i is a *good* node if, conditioned on v_i knowing token 1 in round i , v_i will broadcast token 1 in round i (the round where v_i is the center). Let B denote the set of good nodes. Let C_i denote the set of nodes who know token 1 at the beginning of round i . Without smoothing we would have that $C_i \subseteq \{v_0, v_1, \dots, v_{i-1}\}$, but with smoothing this is not necessarily true. We begin by analyzing C_i under a condition on the good nodes.

► **Lemma 12.** *Suppose that for every $i \leq t$, either v_i is not a good node or v_i does not know token 1 at the beginning of round i . Then $|C_i| \leq 100i \left(\frac{k+\ell}{k}\right) \log n$ for all $i \leq t$ with high probability.*

Proof. By assumption, for all $i \leq t$ the center node of the star does not broadcast token 1. Hence in round i , the nodes who learn token 1 consist of at most the center node and some other nodes who learn token 1 via smoothed edges. Even if *all* of the ℓ smoothed edges

had an endpoint in C_i , the expected number of them who transmit token 1 is at most ℓ/k . Hence a Chernoff bound implies that $|C_i| \leq |C_{i-1}| + 100 \log n \cdot \left(1 + \frac{\ell}{k}\right)$ with high probability. This high probability allows us to do a union bound over the first t rounds, implying that $|C_i| \leq 100i \left(\frac{k+\ell}{k}\right) \log n$ with high probability. \blacktriangleleft

We say that round i has *productive smoothing* if some node in B learns token 1 via an edge added by smoothing. It is easy to see that if after t rounds there has not been any round with productive smoothing, then the assumption in Lemma 12 holds, and hence $|C_i| \leq 100i \left(\frac{k+\ell}{k}\right) \log n$ for all $i \leq t$ with high probability. Since $t < \frac{nk}{100(k+\ell) \log n}$ this means that not all nodes know token 1 at time t , and so random broadcast has not finished. So we just need to argue that with constant probability, there have been no rounds with productive smoothing before round t .

To see this, first note that by the definition of random broadcast, each node in A is good independently with probability $1/k$. Hence the expected number of good nodes is $|A|/k = t/k$. A standard Chernoff bound then implies that $|B| \leq (10t/k) \log n$ with high probability, so from now on we will condition on this being true.

In order for round i to have productive smoothing, at least one of the ℓ edges added by smoothing must have one endpoint in C_i , one endpoint in B , and the endpoint in C_i must choose to broadcast token 1. The probability of this for a single random edge is $\frac{|C_i|}{n} \cdot \frac{|B|}{n} \cdot \frac{1}{k}$, and hence a union bound over all ℓ random edges added in rounded i implies that the probability of productive smoothing in round i is at most

$$\frac{|C_i|}{n} \cdot \frac{|B|}{n} \cdot \frac{\ell}{k} \leq \frac{|C_t|}{n} \cdot \frac{10t \log n}{kn} \cdot \frac{\ell}{k} = |C_t| \cdot \frac{10t\ell \log n}{(kn)^2}.$$

If there has been no productive smoothing before round i , then Lemma 12 implies that with high probability $|C_i| \leq 100i \left(\frac{k+\ell}{k}\right) \log n$. This is high enough probability for us to take a union bound and still have a high probability bound, so we will assume that *if* there has been no productive smoothing before round i then $|C_i| \leq 100i \left(\frac{k+\ell}{k}\right) \log n$.

Let X_i be an indicator random variable for the event that round i has productive smoothing. Then

$$\begin{aligned} \Pr[\exists i \in [t] : X_i = 1] &\leq \sum_{i=1}^t \Pr \left[X_i = 1 \mid \sum_{j=1}^{i-1} X_j = 0 \right] \\ &\leq \sum_{i=1}^t \left(100i \left(\frac{k+\ell}{k} \right) \log n \right) \left(\frac{10t\ell \log n}{(kn)^2} \right) = \frac{1000t^3 \ell (k+\ell) \log^2 n}{k^3 n^2} \end{aligned}$$

Since $t \leq \frac{kn^{2/3}}{(2000\ell(k+\ell) \log^2 n)^{1/3}}$ then this probability is at most $1/2$, which (as discussed) implies Theorem 11.

References

- 1 John Augustine, Gopal Pandurangan, Peter Robinson, and Eli Upfal. Towards robust and efficient computation in dynamic peer-to-peer networks. In *Proceedings of the ACM-SIAM Symposium on Discrete Algorithms*, 2012.
- 2 Chen Avin, Michal Koucký, and Zvi Lotker. How to explore a fast-changing world (cover time of a simple random walk on evolving graphs). In *Proceedings of the International Colloquium on Automata, Languages and Programming*, 2008.
- 3 Soumyottam Chatterjee, Gopal Pandurangan, and Nguyen Dinh Pham. Distributed mst: a smoothed analysis. In *Proceedings of the 21st International Conference on Distributed Computing and Networking*, pages 1–10, 2020.

- 4 Andrea Clementi, Riccardo Silvestri, and Luca Trevisan. Information spreading in dynamic graphs. In *Proceedings of the ACM Symposium on Principles of Distributed Computing*, 2012.
- 5 Oksana Denysyuk and Luís Rodrigues. Random walks on evolving graphs with recurring topologies. In *Proceedings of the International Symposium on Distributed Computing*, 2014.
- 6 Michael Dinitz, Jeremy Fineman, Seth Gilbert, and Calvin Newport. Smoothed analysis of information spreading in dynamic networks, 2022. doi:10.48550/ARXIV.2208.05998.
- 7 Michael Dinitz, Jeremy T Fineman, Seth Gilbert, and Calvin Newport. Smoothed analysis of dynamic networks. *Distributed Computing*, 31(4):273–287, 2018.
- 8 Chinmoy Dutta, Gopal Pandurangan, Rajmohan Rajaraman, Zhifeng Sun, and Emanuele Viola. On the complexity of information spreading in dynamic networks. In *Proceedings of the Twenty-Fourth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 717–736. SIAM, 2013.
- 9 Mohsen Ghaffari, Nancy Lynch, and Calvin Newport. The cost of radio network broadcast for different models of unreliable links. In *Proceedings of the ACM Symposium on Principles of Distributed Computing*, 2013.
- 10 Bernhard Haeupler and David Karger. Faster information dissemination in dynamic networks via network coding. In *Proceedings of the ACM Symposium on Principles of Distributed Computing*, 2011.
- 11 F. Kuhn and R. Oshman. Dynamic networks: models and algorithms. *ACM SIGACT News*, 42(1):82–96, 2011.
- 12 Fabian Kuhn, Nancy Lynch, and Rotem Oshman. Distributed computation in dynamic networks. In *Proceedings of the forty-second ACM symposium on Theory of computing*, pages 513–522, 2010.
- 13 Fabian Kuhn, Nancy Lynch, and Rotem Oshman. Distributed computation in dynamic networks. In *Proceedings of the ACM Symposium on Theory of Computing*, 2010.
- 14 Fabian Kuhn, Rotem Oshman, and Yoram Moses. Coordinated consensus in dynamic networks. In *Proceedings of the ACM Symposium on Principles of Distributed Computing*, 2011.
- 15 Uri Meir, Ami Paz, and Gregory Schwartzman. Models of smoothing in dynamic networks. In *Proceedings of the 34th International Symposium on Distributed Computing*, pages 36:1–36:16, 2020. doi:10.4230/LIPIcs.DISC.2020.36.
- 16 Anisur Rahaman Molla and Disha Shur. Smoothed analysis of leader election in distributed networks. In *International Symposium on Stabilizing, Safety, and Security of Distributed Systems*, pages 183–198. Springer, 2020.
- 17 Calvin Newport. Lower bounds for structuring unreliable radio networks. In *Proceedings of the International Symposium on Distributed Computing*, 2014.
- 18 Daniel A. Spielman and Shang-Hua Teng. Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time. *J. ACM*, 51(3):385–463, 2004.
- 19 Daniel A. Spielman and Shang-Hua Teng. Smoothed analysis: an attempt to explain the behavior of algorithms in practice. *Commun. ACM*, 52(10):76–84, 2009.

A Random Broadcast in Smoothed Static Networks

As previously argued, a minimum amount of smoothing (i.e., $\ell = 1$) improves the performance of random broadcast in dynamic networks from $O(kn)$ to $\tilde{O}(kn^{2/3})$. To better understand how smoothing supports information spreading, a natural follow-up question is to investigate its impact on random broadcast in *static* networks. If smoothed analysis provides the same bounds for both the dynamic and static settings, this would imply that smoothing essentially bypasses the difficulties induced by changing graph edges. Here we show this is not the case. In more detail, we prove that in static networks, 1-smoothing improves the complexity of random broadcast down to $\tilde{\Theta}(k\sqrt{n})$ rounds, beating what we can guarantee in the dynamic

setting. This establishes a gap between static and dynamic networks with respect to random broadcast, confirming the intuition that network dynamism introduces unique difficulties for information dissemination that smoothing alone cannot fully overcome.

Due to space constraints, we only prove an upper bound of $\tilde{O}(k\sqrt{n})$. The matching lower bound, as well as missing proofs, can be found in the full version [6].

Critical to our analysis is the following graph decomposition result:

► **Lemma 13.** *Fix a connected static graph $G = (V, E)$ of size n . There exists a partition of V into components C_1, C_2, \dots, C_x , such that for each i , $1 \leq i \leq x$: (1) $|C_i| \geq \sqrt{n}$; (2) the subgraph of G induced by C_i is connected and has a diameter at most $6\sqrt{n}$.*

We are now ready to prove our upper bound. The key intuition in the following argument is that we can analyze the spread of a given target token within the context of the components provided by Lemma 13. We will show, roughly speaking, that when the target token is first spreading, if \mathcal{A} is the set of components that know the target, it is likely that within $\tilde{O}(k\sqrt{n})$ rounds, each component in \mathcal{A} will succeed in seeding the target over a smoothed edge to a unique component not in \mathcal{A} , effectively *doubling* the number of components that have learned the token. A logarithmic number of such doublings are sufficient to spread the target to at least half the network, at which point the analysis shifts to the perspective of the remaining components, and argues that within an additional $\tilde{O}(k\sqrt{n})$ rounds, each is likely to connect to an already informed component by a smoothed edge and receive the token. Care is needed in the formal argument to deal with both uneven-sized components and dependencies between the smoothed edge behavior in different components during the same spreading interval.

► **Theorem 14.** *Fix some connected static graph G of size n . Fix any rumor set size $k \geq 1$. With high probability, random broadcast solves k -message broadcast in $O(k\sqrt{n} \log^2 n)$ rounds in G with 1-smoothing.*

Proof. Fix a graph $G = (V, E)$ of size n , and a rumor set of size k , as specified by the theorem statement. Fix an arbitrary *target token* from the rumor set. We argue that this target token will spread to full network with the stated complexity.

To do so, we first apply Lemma 13 to partition G into components C_1, C_2, \dots, C_x with the specified properties. Moving forward, in a given round, we call a component *seeded* if at least one node in the component knows the target token, and call it *completed* if every node knows the target. It is straightforward to establish that once a component is seeded, it will be completed, with high probability, in an additional $O(k\sqrt{n} \log n)$ rounds. One approach to making this argument is to fix a breadth-first search tree in the component rooted at a node that knows the target token. This root will broadcast the target in each round with probability at least $1/k$. Because each broadcast decision is independent, we can apply a Chernoff bound to establish that in $O(k \log n)$ rounds, the root will broadcast the target at least once, with high probability.

We can then repeat this analysis to show that within an additional $O(k \log n)$ rounds, every node at level 1 in the tree will have sent the target token, informing every node at level 2, applying union bounds to ensure that every node at level 1 succeeds with sufficient probability. An additional $O(k \log n)$ rounds moves the token to level 3, and so on. By the guarantees of Lemma 13, the tree has $O(\sqrt{n})$ levels, meaning that $O(k\sqrt{n} \log n)$ rounds is sufficient to spread a target token through the whole component with high probability.

Returning to our main argument, fix a round r . Let \mathcal{A}_r be the subset of components that are seeded at the beginning of round r . Let n_r be the number of nodes in components in \mathcal{A}_r . We first consider the case where $n_r < n/2$. To do so, let \mathcal{B}_r be all the components not in

\mathcal{A}_r . Our goal is to show that in an additional $O(k\sqrt{n}\log n)$ rounds, the target token will be delivered over smoothed edges to a collection of components in \mathcal{B}_r , where they will then spread to a total number of new nodes that is at least a constant fraction of n_r – increasing the number of nodes that know the target token by a constant factor.

To make this argument, we divide this period of $O(k\sqrt{n}\log n)$ rounds starting at round r into three parts. During the first part, we wait for the target token to spread sufficiently to complete every component in \mathcal{A}_r . As argued above, with high probability, this takes $O(k\sqrt{n}\log n)$ rounds.

Next we consider a stretch of an additional $O(k\sqrt{n}\log n)$ rounds that we call the *seeding interval*. We will show that during this interval, smoothed edges between components in \mathcal{A}_r and \mathcal{B}_r , will seed the target token into a collection of \mathcal{B}_r components whose collective size is a constant fraction of n_r . The final $O(k\sqrt{n}\log n)$ rounds will be dedicated to allowing these seeded \mathcal{B}_r components to complete. (Of course, it is possible that in the time we spent completing components in \mathcal{A}_r , the target token might have already made its way to components in \mathcal{B}_r and started spreading, but this only speeds up our efforts.)

We are left then to study closely the behavior of the smoothed edges during the smoothing interval of this period. Though Lemma 13 guarantees that each contains at least \sqrt{n} nodes, it is possible that some might be much larger than this lower limit. It is useful for our purposes to temporarily reorganize these already informed nodes into more uniform-sized groups. With this in mind, let \mathcal{A}'_r be a partition of the nodes in components in \mathcal{A}_r into *groups* that are all of size $\Theta(\sqrt{n})$. For our purposes, it does not matter how this partition is defined. The nodes in each $S \in \mathcal{A}'_r$, for example, do not need to be connected. In the next step of our analysis, we will only concern ourselves with the probability that a node in a given group is selected as an endpoint of a smoothed edge.

Next, fix some arbitrary group $S_1 \in \mathcal{A}'_r$ to consider first. Let \hat{n}_r be the number of nodes in \mathcal{B}_r . Because we are still considering the high-level case where $n_r < n/2$, we know $\hat{n}_r \geq n/2$. We now analyze the rounds of the seeding interval in order, stopping at the first round in which: (1) a smoothed edge connects a node $u \in S_1$ to a node in a component in \mathcal{B}_r ; and (2) u broadcasts the target token. It is straightforward to see that in any given round, a *productive* connection of this type occurs with probability at least $\frac{|S_1|}{n} \cdot \frac{\hat{n}_r}{n} \cdot \frac{1}{k} \geq \frac{1}{2k\sqrt{n}}$.

With high probability, therefore, we will find such a productive connection in a seeding interval of length $O(k\sqrt{n}\log n)$. We now want to consider the other groups in \mathcal{A}'_r , and argue that they too will succeed in forming a productive connection during this *same* seeding interval. This will require care to deal properly with dependencies.

The first thing we do is take the component in \mathcal{B}_r at the other end of the productive connection from S_1 , and add it to a set \mathcal{C} of successfully seeded components. Before proceeding in our analysis of this seeding interval, we make two checks to see if we are already done. Let n' be the number of nodes in components in \mathcal{C} at this point. If $n_r + n' \geq n/2$, then we can simply wait an additional $O(k\sqrt{n}\log n)$ rounds to complete the component in \mathcal{C} , and be done with the high-level case we are considering in which less than half of the nodes know the target token. Similarly, if $n' \geq n_r/2$, then we can finish our analysis of this particular seeding interval as we have accomplished our proximate goal of increasing the number of nodes that know the target token by a constant factor.

If we fail both checks, we must then continue with analyzing our same seeding interval in the hopes of seeding more tokens into $\mathcal{B}_r \setminus \mathcal{C}$. Fix a new group $S_2 \in \mathcal{A}'_r \setminus S_1$. As before, consider rounds in the seeding interval one by one, starting from the first round of the interval, until we arrive at a with a productive connection from S_2 to a not yet seeded component in $\mathcal{B}_r \setminus \mathcal{C}$. In each such round, we argue that the probability of a productive connection is still

in $\Omega(\frac{1}{k\sqrt{n}})$. The main difference as compared to prior groups considered is that the number of nodes that can receive a productive smooth edge decreases as we add more components to \mathcal{C} . By our above check, however, we know that the number of nodes in \mathcal{C} is less than $n_r/2 < n/4$. Because we similarly assume that \mathcal{B}_r has at least $n/2$ nodes, then $\mathcal{B}_r \setminus \mathcal{C}$ must still have at least $n/4$ total nodes. The probability of selecting an endpoint in $\mathcal{B}_r \setminus \mathcal{C}$ will therefore always be at least $1/4$, reducing the original productive connection probability calculated for S_1 by only a constant factor for later groups.

As previewed, however, we must also consider dependencies. When considering a given round r' in the seeding interval when considering group S_2 , there are two relevant possibilities: (1) r' was the productive connection from our analysis of S_1 ; (2) r' was not the productive connection for S_1 . The first case introduces a problematic dependency, as being productive for one group prevents you from being productive for another. To deal with this dependency we simply *ignore* in our analysis any round that was part of a productive connection for a previously studied group. The second case, by contrast, introduces a useful dependency: knowing that r' was not productive for S_1 only *increases* the probability that it is productive for S_2 . A standard negative correlation argument tells us that when lower bounding the probability of a productive connection with respect to S_2 , it is fine to treat each round in the second case as succeeding with an independent probability in $\Omega(\frac{1}{k\sqrt{n}})$.

It follows that with high probability, S_2 will also have a productive connection in our seeding interval. At this point, we can repeat the above analysis. Add the newly seeded component to \mathcal{C} . Check if we are done. If not, select a new set $S_3 \in \mathcal{A}'_r \setminus \{S_1, S_2\}$ and study the same interval again, round by round, ignoring now only the two rounds in which S_1 and S_2 succeeded in forming productive connections, and so on, until we finally match the criteria for stopping our analysis. (Notice that removing productive rounds from consideration in the seeding interval is not a problem as there are at most $O(\sqrt{n})$ such productive rounds possible given that $|\mathcal{A}'_r| = O(\sqrt{n})$, and our interval length can be made sufficiently long to still provide us the needed high probability of success even with up to $O(\sqrt{n})$ rounds omitted from consideration.)

Moving on with our argument, recall that for our fixed round r , there are two different criteria that might terminate the above seeding analysis. The first is that at least half the nodes in the network learn the target token, at which point we are ready to move on to the second half of our overall analysis, which we will discuss shortly. The second termination criteria is that the number of nodes in components in \mathcal{B}_r that learn the target token is at least a constant factor of n_r . In this case, we fix a new round r' after the seeding interval in question is done, and after all components in \mathcal{C} complete. We then reapply our above analysis starting from r' , increasing the number of nodes that know the target node by another constant factor. We can repeat this at most $O(\log n)$ times before at least half the nodes know the target token.

We now consider what happens once we succeed in spreading the target token to at least half the nodes in the network. We can now redeploy pieces of our above spreading argument to show that target token will make it to all remaining nodes in just one more additional spreading interval of length $O(k\sqrt{n} \log n)$ rounds.

We first dispense with the sub-case in which less than \sqrt{n} nodes do not the token; i.e., we are almost done. If this is true, each uninformed node u is within \sqrt{n} of at least one informed node, meaning a straightforward spreading analysis will deliver the token to each such u with high probability within $O(k\sqrt{n} \log n)$ rounds – completing the rumor spreading.

We are left with the sub-case in which somewhere between \sqrt{n} and $n/2$ nodes remain that do not have the target token. Let \mathcal{A} be the set of components that contain at least one node that from this set of nodes that do not know the target. Some of these components are

seeded (i.e., at least one node in them knows the target) and some are not (i.e., no node knows the token). Let \mathcal{A}' be the subset of \mathcal{A} that contains only the non-seeded components. Spend $O(k\sqrt{n} \log n)$ to complete the seeded components from \mathcal{A} . We now turn our attention exclusively to the components from \mathcal{A}' , as these are the only components at this point which can possibly contain *any* nodes that do not know the target token. We know each such component is of size at least \sqrt{n} , and that at least half the total nodes in the network are not in these components. We can therefore apply our above spreading analysis to the components in \mathcal{A}' , which establishes that in a single additional spreading interval, every component in \mathcal{A}' will have a productive connection with a completed component, thus seeding it with the target token. We can then complete these components, and therefore complete k -message broadcast, in an additional $O(k\sqrt{n} \log n)$ rounds.

To conclude the proof, we note that all relevant growth and spreading arguments hold with high probability, and that there are at most $\text{poly}(n)$ such events that must succeed. This allows us to deploy union bounds to prove that the entire problem is successful, with high probability, after $O(\log n)$ intervals of length $O(k\sqrt{n} \log n)$, yielding the claimed overall time complexity of $O(k\sqrt{n} \log^2 n)$ rounds. ◀

B Random Broadcast in Well-Mixed Networks

Dutta et al. [8] introduced the notion a *well-mixed* network in the context of the k -message broadcast problem. A network satisfies this property if for each token t and node u , with some independent constant probability, u starts with t . The main $\tilde{\Omega}(nk)$ lower bound from [8] also holds for well-mixed networks. To circumvent this bound, they assume a stronger communication model that allows interactive communication on each edge, and provide a randomized algorithm in this setting that solves k -message broadcast in $\tilde{O}(n+k)$ rounds.

Here we deploy our predecessor path constructions, originally designed to support our smoothed analysis, to now explore an alternative non-smoothing method to circumvent the $\tilde{\Omega}(nk)$ lower bound: weakening the adversary. The lower bound in [8] assumes a *strongly adaptive* adversary that knows all the nodes' random bits, allowing it to generate a network graph in each round based on the specific tokens nodes will broadcast during that round. Another natural option is the *oblivious* adversary, in which the adversary generates the graph without advance knowledge of the nodes' random bits. Indeed, the question of whether an oblivious adversary enabled better bounds was identified as important future work in [8].

Using predecessor paths, we prove that with an oblivious adversary, our simple random broadcast strategy solves k -message broadcast in well-mixed networks in $\tilde{O}(k)$ rounds. This improves on the $\tilde{\Omega}(n+k)$ bound from [8], as it eliminates the n dependence. It also matches the trivial $\Omega(k)$ lower bound that holds for *all* k -message broadcast algorithms in a well-mixed network.⁴ Indeed, our result is actually more general, showing $\tilde{O}(k/p)$ rounds are needed, when p is the token probability; i.e., the definition of well-mixed in [8] assumes $p = \Theta(1)$.

This result opens a clear separation between strongly adaptive and oblivious adversaries in the context of well-mixed networks, and hints such a separation might exist for arbitrary token distributions as well. It also provide further evidence that the simple random broadcast strategy is a highly effective strategy for information dissemination in these settings.

⁴ Fix a static line. Let u be one of the endpoints. With constant probability u is missing $\Omega(k)$ tokens in its initial set. Because u can receive at most one token per round in a line topology, it requires at least $\Omega(k)$ rounds for it to learn these missing tokens one by one.

Formally, we consider the following generalized version of the property concerning initial token distributions introduced in [8]:

► **Definition 15.** Fix a probability $p > 0$. We say an initial token distribution is p -mixed if for each node $u \in V$ and token $t \in T$: u 's initial token set includes t with independent probability p .⁵

Given this definition, we prove our main upper bound result:

► **Theorem 16.** Fix a probability $p > 0$. Random broadcast solves k -message broadcast in $O((k/p) \log n)$ rounds, with high probability, when run in a network with a p -mixed token distribution.

Proof. Fix any u and t . Our first step is to choose a large enough r such that $|P_{u,t}(1, r')| \geq \alpha(1/p) \ln n$, for a constant $\alpha \geq 1$ we will fix later. We want this to hold with probability at least $1 - n^{-4}$. To do so, we can apply Theorem 4 with $x = 4$, $r = 1$, $r' = 32k\alpha(1/p) \ln n$. (Notice, for our parameters, the corresponding $z = r' - r$ value is lower bounded by $8xk \ln n = 32k \ln n$, as required by the theorem statement.) This tells us that with probability at least $1 - n^{-4}$, $|P_{u,t}(1, r')| > \frac{32k\alpha(1/p) \ln n}{2k} \geq \alpha(1/p) \ln n$, as needed. A union bound over the $nk \leq n^2$ possible combinations of processes and tokens, tells us that the probability that any predecessor path is less than this length is less than n^{-2} .

Fix some such path $P_{u,t}(1, r')$ that is sufficiently long; i.e., $q = |P_{u,t}(1, r')| \geq \alpha(1/p) \ln n$. We apply the definition of p -mixed to argue that it is likely that at least one node in this path starts the execution with token t . By definition, of a predecessor path, each node in $P_{u,t}(1, r')$ is unique. By the definition of p -mixed, each u_i starts the execution with token t with independent probability p . Given this independence, we can use the indicator random variable X_i to describe whether or not u_i starts with token t , and then apply our Chernoff Bound from Theorem 1 to $X = \sum_{i=1}^{r'} X_i$, to bound the probability that X is far from its expectation $\mu = pq$. Formally, we get that $\Pr[Y \leq \mu/2] \leq \exp(-\mu/8)$. Given that $\mu = pq \geq \alpha \ln n$, then for $\alpha \geq 32$, it follows that $\Pr[Y = 0] < n^{-4}$.

A union bound over the $nk < n^2$ node and token pairs tells us that if at all predecessor paths are of length at least q , the the probability *any* path does not contain at least one node starting with the path's target token is less than n^{-2} . A union bound over the failure probabilities for these two events gives us the final result that with high probability, for every $u \in V$ and $t \in T$, at least one node in $P_{u,t}(1, r')$ starts with t . As argued in our previous discussion of predecessor paths, if any node on $P_{u,t}(1, r')$ starts with t then u receives t by round r' . Therefore, with high probability, random broadcast solves k -message broadcast in a p -mixed network in $r' = O((k/p) \log(n))$ rounds, as claimed. ◀

⁵ A slight technicality of this definition is that as stated it allows for certain tokens to not show up at all, making termination impossible. A simple fix is to assume that all k tokens are distributed arbitrarily to at least one node, then the random process is deployed to further introduce more tokens into the system.