Rumor Spreading: Robustness and Limiting Distributions

Dissertation
an der Fakultät für Mathematik, Informatik und Statistik
der Ludwig-Maximilians-Universität München

vorgelegt von
Simon Reisser

1. Dezember 2021

Erstgutachter: Prof. Dr. Konstantinos Panagiotou
Zweitgutachter: Prof. Dr. Benjamin Doerr
Drittgutachter: Prof. Dr. Thomas Sauerwald
Tag der mündlichen Prüfung: 24.06.2022

Eidesstattliche Versicherung

(Siehe Promotionsordnung vom 12.07.11, § 8, Abs. 2 Pkt. 5.)

Ich versichere hiermit an Eidesstatt, dass die vorgelegte Dissertation von mir selbständig, ohne unerlaubte Beihilfe angefertigt ist.

München, 11. August 2022

Simon Reisser

Contents

List of Figures

1 Abstract

In this thesis, we study mathematical aspects of information dissemination. The four collected works investigate randomized rumor spreading with regard to its robustness and asymptotic runtime as well as adversarial effects on opinion forming.

In the first contribution, *Robustness of Randomized Rumor Spreading*, we investigate the popular randomized rumor spreading algorithms *push*, *pull* and *push&pull*. These are used to spread information quickly through large networks, typically modelled by graphs. Starting with one informed vertex and depending on the used algorithm the information is spread in a round based manner. Using *push*, every informed vertex chooses a random neighbour and passes the information forward. With *pull*, each vertex yet uninformed connects to a randomly chosen neighbor and receives the information, if the vertex it connected to is informed. *push&pull* is a combination of *push* and *pull*. Every vertex chooses a random neighbour, if one of them is informed then the other will be informed as well. Their advantages over deterministic algorithms are, that they are easy to implement, fast and very robust against failures. However, there is only sporadic information available to substantiate the claimed robustness. The aim of this work is to close this gap. To that end, three orthogonal properties and their effects on the speed of the dissemination are studied. First, we show that the density of the graph does not play an important role. For fast dissemination it is not relevant how many edges there are, but how evenly they are distributed in the graph. Thus, a network could have many faulty connections, but as long as the remaining ones are spread evenly the speed of the dissemination is not significantly impacted. This begs the question how evenly the remaining edges need to be spread to guarantee a fast dissemination. Surprisingly, the answer to this question is not the same for all three rumor spreading algorithms. *pull* and *push&pull* are very robust. Starting from a graph with evenly distributed edges and thus fast dissemination one may introduce irregularities by deleting up to one half of all edges at each node and the dissemination remains fast. However, for *push* the dissemination already slows down significantly if only few irregularities are introduced. Lastly, we additionally consider random message transmission failures. From previous works, we know that on "nice" graphs all three algorithms only slow down proportionally to the failure probability. However, when considering the effect of density and irregularities together with transmission failures, the picture changes once more. *pull* alone retains its fast dissemination. With a suitable choice of parameters, *push&pull* similar to *push* can be slowed down significantly. Thus, we can not unconditionally confirm the claimed robustness for all three rumor spreading algorithms, only *pull* proved to be robust against all introduced challenges, *push* and *push&pull*, however, did not.

In the second contribution, *Asymptotics for Push on the Complete Graph*, we move from the general approach of quantifying the robustness of all three randomized rumor spreading algorithms on a broad range of networks to very precisely describing the runtime of *push* on complete graphs only. Thereby, the runtime is defined as the time until the information is disseminated to all vertices in the graph. In this work, we completely describe the limiting distribution of the runtime of *push* on the complete graph in terms of a Gumbel distributed random variable. We made a surprising observation, the asymptotic distribution does not converge everywhere, only on suitable subsequences. These subsequences $(n_i)_{i\in\mathbb{N}}$ are characterised by having converging rational part of $(\log_2 n_i)_{i\in\mathbb{N}}$ as well as $(\ln n_i)_{i\in\mathbb{N}}$. This results in the phenomena, that the expected runtime is not constant either but infimum and supremum over all n differ by $\approx 10^{-4}$.

After successfully solving *push* on the complete graph, a natural question is to ask whether the same can be achieved for other rumor spreading algorithms. The third contribution, *Asymptotics for Pull on the Complete Graph*, answers this question for *pull*, describing the asymptotic distribution of the runtime of *pull* on the complete graph in terms of a martingale limit. Again we observed that the limiting distribution only exists on suitable subsequences $(n_i)_{i\in\mathbb{N}}$, those with convergent rational part of $(\log_2 n_i + \log_2 \ln n_i)_{i\in\mathbb{N}}$. We study the expected runtime numerically, finding strong evidence that it is not constant either.

The last contribution, *The Effect of Iterativity on Adversarial Opinion Forming*, deviates from the previously considered model and introduces a second competing piece of information. We interpret them as opinions and assume one to be the truth and the other one to be a falsehood. The opinions are spread through the network by a simple majority rule, i.e. uninformed vertices take the majority opinion of their informed neighbours. Known properties that guarantee robustness are the degree being sufficiently bounded or the edges being evenly distributed. The question considered in this contribution is whether an alternative iterative dissemination process influences robustness. Alon et al. conjecture that iterativity is always beneficial for the adversary. We refute that conjecture by giving a graph where iterativity benefits robustness.

2 Zusammenfassung

In dieser Arbeit beschäftigen wir uns mit mathematischen Aspekten der Informationsverbreitung in Netzwerken. Die vier gesammelten Beiträge untersuchen randomisierte Gerüchteverbreitungsalgorithmen hinsichtlich ihrer Robustheit und asymptotischen Laufzeit, sowie gegnerische Auswirkungen auf die Meinungsbildung.

Der erste Beitrag, *Robustness of Randomized Rumor Spreading*, befasst sich mit den populären randomisierten Gerüchteverbreitungsalgorithmen *push*, *pull* und *push&pull*. Diese werden dazu verwendet, um Informationen schnell durch große, als Graphen modellierte Netzwerke zu verteilen. Beginnend mit einem informierten Knoten und in Runden verfahrend, werden die Informationen abhängig vom verwendeten Algorithmus verteilt. Wird *push* benutzt, so wählt jeder informierte Knoten einen zufälligen Nachbarn und gibt die Information weiter. Mit *pull* wählen uninformierte Knoten zufällige Nachbarn und werden informiert, falls der gewählte Nachbar informiert ist. *push&pull* ist eine Kombination aus *push* und *pull*. Jeder Knoten wählt einen zufälligen Nachbarn aus, ist einer der beiden informiert, so wird auch der andere informiert. Mit einer einfachen Implementierung, hohen Geschwindigkeit und einer starken Robustheit heben sich die randomisierten Gerüchteverbreitungsalgorithmen positiv von deterministischen Algorithmen ab. Bisher liegen jedoch nur sporadische Informationen vor, um die beobachtete Robustheit auch rigoros zu belegen. Ziel dieser Arbeit ist es, diese Lücke zu schließen. Dafür betrachten wir drei verschiedene, strukturelle Eigenschaften der Graphen, um deren Auswirkungen auf die Geschwindigkeit der Verbreitung zu studieren. Als erstes Ergebnis zeigen wir, dass die Dichte des Netzwerks keinen nennenswerten Einfluss hat. Für eine schnelle Verbreitung der Informationen ist nicht die Anzahl der Kanten relevant, sondern deren gleichmäßige Verteilung. Ein Netzwerk könnte folglich viele fehlerhafte Verbindungen haben, aber solange die verbleibenden Verbindungen gleichmäßig verteilt sind, wird die Verbreitung nicht wesentlich verlangsamt. Dies regt die Untersuchung an, wie gleichmäßig die verbleibenden Kanten sein müssen, um eine schnelle Verbreitung zu gewährleisten. Wider Erwarten konnten wir Unterschiede in Abhängigkeit des gewählten Gerüchteverbreitungsalgorithmus aufzeigen. *pull* und *push&pull* sind sehr widerstandsfähig. Denn ausgehend von einem „schönen“ Graph mit gleichmäßig verteilten Kanten können durch Löschen von Kanten Unregelmäßigkeiten eingebracht werden durch die sich die Geschwindigkeit der Gerüchteverbreitung nicht nenneswert verändert. Im Gegensatz dazu verlangsamt sich die Verbreitung mit *push* bereits erheblich, wenn nur wenige Unregelmäßigkeiten auftreten. Abschließend befassen wir uns ergänzend mit zufällig auftretenden Übertragungsfehlern. Aus früheren Arbeiten wissen wir, dass sich bei „schönen“ Graphen alle drei Algorithmen nur proportional zur Ausfallswahrscheinlichkeit verlangsamen. Betrachten wir hingegen die Auswirkungen der Dichte und der Unregelmäßigkeiten mit Übertragungsfehlern zusammen, entsteht eine neue Sachlage. Dabei behält nur *pull* seine schnelle Verbreitung bei, *push&pull* kann bei einer entsprechenden Wahl der Parameter ähnlich wie *push* verlangsamt werden. Somit ist eine Bestätigung der behaupteten Robustheit der drei Gerüchteverbreitungsalgorithmen nicht bedingungslos möglich. Lediglich *pull* erwies sich als widerstandsfähig gegenüber allen betrachteten Problemen, *push* und *push&pull* jedoch nicht.

Im zweiten Beitrag, *Asymptotics for Push on the Complete Graph*, gehen wir vom allgemeinen Ansatz der Beschreibung der Robustheit aller drei randomisierten Gerüchteverbreitungsalgorithmen auf einem breiten Spektrum von Netzwerken zu einer sehr präzise Beschreibung der Laufzeit von *push* auf vollständigen Graphen über. Dabei definiert sich die Laufzeit als die Zeit, in der die Information an alle Knoten im Graph verteilt wird. In dieser Arbeit beschreiben wir die Grenzverteilung der Laufzeit von *push* auf dem vollständigen Graph. Dabei haben wir eine überraschende Beobachtung gemacht, denn die asymptotische Verteilung konvergiert nicht überall, sondern nur auf geeigneten Teilfolgen. Diese Folgen $(n_i)_{i\in\mathbb{N}}$ zeichnen sich dadurch aus, dass sowohl die Nachkommastellen von $(\log_2 n_i)_{i\in\mathbb{N}}$ als auch die von $(\ln n_i)_{i\in\mathbb{N}}$ konvergieren. Dies resultiert in dem Phänomen, dass die erwartete Laufzeit nicht konstant ist, vielmehr unterscheiden sich Supremum und Infimum über alle n um ungefähr 10^{-4}.

Nach dieser erkenntnisreichen Arbeit stellt sich die natürliche Frage, ob dasselbe für die anderen Gerüchteverbreitungsalgorithmen gilt. Die daran anschließende Arbeit *Asymptotics for Pull on the Complete Graph* bejaht die aufgeworfene Frage für *pull*, indem die asymptotische Verteilung der

Laufzeit von *pull* auf vollständigen Graph mit Hilfe eines Martingalgrenzwertes beschrieben wird. Ferner wird beobachtet, dass die Grenzverteilung nur auf geeigneten Teilfolgen existiert, nämlich solche Folgen $(n_i)_{i\in\mathbb{N}}$ bei denen die Nachkommastellen von $(\log_2 n_i + \log_2 \ln n_i)$ konvergieren. Die erwartete Laufzeit wird mit Hilfe dieser Beschreibungen empirisch untersucht, wobei es eine starke Evidenz gibt, dass auch diese nicht konstant ist.

Der letzte Beitrag, *The Effect of Iterativity on Adversarial Opinion Forming*, weicht vom bisher betrachteten Modell ab und führt eine zweite, konkurrierende Information ein. Diese interpretieren wir als Meinungen und nehmen eine davon als wahr an. Die Meinungen werden durch eine einfache Mehrheitsregel im Netzwerk verbreitet, d. h. uninformierte Knoten nehmen die Mehrheitsmeinung ihrer informierten Nachbarn an. Dabei sehen wir ein Netzwerk als robust an, wenn selbst ein Kontrahent die anfangs informierten Knoten nur so wählen kann, dass am Ende der Verbreitung stets die Mehrheit der Knoten von der Wahrheit überzeugt ist. Bekannte Beispiele robuster Netzwerke sind solche mit hinreichend beschränkten Knotengraden oder mit ausreichend gleichmäßig verteilten Kanten. In unserem Beitrag betrachten wir die Frage, inwiefern Robustheit durch einen alternativen, iterativen Verbreitungsprozess beeinflusst wird. Alon et al. vermuten eine negative Auswirkung von Iteration auf Robustheit. Wir widerlegen diese Vermutung durch Konstruktion eines Graphen, auf welchem ein iterativer Prozess die Verbreitung der Wahrheit begünstigt.

3 Acknowledgements

My utmost gratitude goes to Konstantinos Panagiotou for his guidance as my supervisor. Without his commitment, constant invitation for discussion and fruitful insights this work would not have happened.

I want to furthermore thank Rami Daknama for the good teamwork and the many fruitful discussions. He is not just a co-author and colleague but a friend as well. I want to express my appreciation to my colleagues for making my time at the university enjoyable. The many lunches and coffee breaks were always looked forward to. Thank you, Anna, Felix, Kilian, Leon, Leonid, Matija, Nannan, Thomas and Vincent.

I want to thank my parents for unconditionally supporting me and making all of this possible in the first place. Finally I want to thank Julia for having my back at all times and being an awesome person.

4 Introduction

Randomized rumor spreading is used to disseminate information trough a network. It is simple, fast and robust and therefore has a plethora of applications in replicated databases, wireless senor networks and even blockchains. This work, which is structured as follows, looks at the attributed properties 'fast' and 'robust' in more detail. First we review the related literature, giving an overview of the latest results, some applications and core ideas that arise. Afterwards we interpret our new results and highlight where we extend the state of the art. Finally all four collected works are appended in full with details of the author's contribution to each article.

4.1 Related Literature

Let G be an undirected graph on a vertex set V and edge set E. Furthermore, set one vertex $v \in V$ to be informed. Consider the following round-based protocols with the goal of spreading the information from v to all other vertices in $V \setminus \{v\}$. Until all vertices are informed, every round the following steps are performed:

- *push*: In every round, every informed vertex v independently and uniformly at random chooses a neighbour u ($u, v \in V$ are neighbours if $(u, v) \in E$). If u is not informed, we say uninformed, then it becomes informed otherwise nothing happens. This was first introduced in [39].
- *pull*: In every round, every uninformed vertex v independently and uniformly at random chooses a neighbour u, if u is informed, than v becomes informed otherwise nothing happens. This was first introduced in [19].
- *push&pull*: Both informed or uninformed vertices v choose a neighbour u independently and uniformly at random. If either v or u is informed then both of them are, otherwise nothing happens. This was first introduced in [47]. In the following we abbreviate *push&pull* by *pp*.

These three protocols are known as *randomized rumor spreading* and the main property under investigation is the *runtime*, i.e. the random variable $X^p(G, v)$ that counts the rounds required for the protocol $p \in \{push, pull, pp\}$ to spread the information introduced at v to all vertices of G. If the initial vertex does not matter we may omit it, in particular we denote $X^p(G, v)$ by X_n^p for all $v \in V$ and $p \in \{push, pull, pp\}$ if $G = K_n$ is the complete graph on n vertices.

Early works. The study of randomized rumor spreading has first been introduced by Alan Frieze and Geoffrey Grimmett in 1985 [39]. They introduced the *telephone call problem* that corresponds to X_n^{push} in our notation and they found the first asymptotic expression of X_n^{push} ,

$$X_n^{push} = \log_2 n + \ln n + o(\ln n) \quad \text{in probability as} \quad n \to \infty$$

where ln denotes the natural logarithm. Furthermore, they found a large deviation bound, for all $\varepsilon > 0$ and $\gamma > 0$,

$$P\Big(X_n^{push} > (1+\varepsilon)\big(\log_2 n + (1+\gamma)\ln n\big)\Big) = o(n^{-\gamma}).$$

Following the work of Frieze and Grimmett the asymptotic expression of X_n^{push} has been more thoroughly studied by Boris Pittel in 1987 [56]. He narrowed the error in the asymptotic expression down to constant order,

$$X_n^{push} = \log_2 n + \ln n + O(1) \quad \text{in probability as} \quad n \to \infty.$$

In addition Boris Pittel made an important observation. He defined I_t to be the set of informed vertices before round t of *push* and recognized that $|I_t|$ can be described by a deterministic sequence for most t, that is

$$|I_{t+1}| \approx n - \big(n - |I_t|\big)e^{|I_t|/n}. \tag{1}$$

This approach will play a major part in later works and in this thesis as well.

In the same year, Demers et al. [19] introduced randomized rumor spreading into practical applications. They faced the problem of distributed databases and how to propagate updates that originate at one specific site to all participants in the network. Typical approaches were either very specific to the underlying network and did not scale well to lager networks or were not reliable to successfully transmit the update if errors occurred. To overcome these problems they proposed two probabilistic algorithms that are based on randomized rumor spreading.

- *Anti-entropy*: All nodes regularly choose a random neighbour, compare their databases and resolve differences. This is extremely reliable but comparisons can not be made with high frequencies as comparing databases is costly, thus, it might be slower than deterministic algorithms.

- *Rumor mongering*: All nodes that contain the "latest" update periodically choose a random neighbour and pass the update along. After some time they stop and regard the update as "old". This allows for much higher frequencies as *Anti-Entropy* as less data is exchanged in each call, but there is a risk of not passing the update to all nodes. This can be implemented as either *push* or *pull* depending on the frequency of expected updates.

Applications based on *Anti-entropy* and *Rumor mongering* turned out to be highly effective and greatly reduced the workload while being both robust and scalabe. These so-called gossip protocols and variations of them have garnered much attention, they are applied in wireless sensor networks [48], multicast [10], blockchains [55] like Bitcoin [51] and Ripple [4] as well as many more. Moreover, Demers et al. introduced *pull* as an alternative to *push*.

The successful application of randomized rumor spreading prompted more theoretical work. In 1990 Feige et al. [32] studied $X^{push}(G)$ for several types of graphs G on n vertices. They first showed the general bound

$$\ln n \leq X^{push}(G) \leq 12n \ln n \quad \text{for all graphs } G.$$

Additionally, they showed that these bounds are asymptotically tight, in the sense that both $X^{push}(G) = O(\ln n)$ and $X^{push}(G') = O(n \ln n)$ are achieved by some graphs G, G'. Additionally they described $X^{push}(G)$ depending only on two graph parameters: the maximal degree $\Delta(G)$ and the diameter $\operatorname{diam}(G)$

$$X^{push}(G) = O\big(\Delta(G)(\operatorname{diam}(G) + \ln n)\big) \quad \text{with high probability.}$$

With high probability means probability tending to 1 as n tends to infinity, we occasionally abbreviate it as whp. Furthermore, they showed the first theoretical robustness result. Let G be a graph derived from the complete graph K_n by deleting up to $n/3$ edges. Then

$$X^{push}(G) = O(\ln n) \quad \text{with high probability.}$$

Lastly they showed that if G is a hypercube or an Erdős-Rényi random graph with edge probability at least $(1 + \varepsilon)(\ln n)/n$, then

$$X^{push}(G) = \Theta(\ln n).$$

This nicely demonstrates the different research directions taken when studying randomized rumor spreading. These are either finding general bounds that only depend on some graph parameters or finding sharper bounds for specific graph classes. The remainder of this subsection is structured accordingly. First we explore results for graph classes like complete graphs or random graphs, then we look at the more general results using notions of conductance and expansion. Finally we will state some results concerning robustness and conclude with popular variations of randomized rumor spreading.

Complete graphs. The first specific graph class that was investigated and the class that is best understood are complete graphs. Pittel [56] showed that asymptotically $X_n^{push} = \log_2 n + \ln n + O(1)$. This bound was improved by Benjamin Doerr and Marvin Künnemann in 2014 [26], they proved that

$$\lfloor \log_2 n \rfloor + \ln n - 1.116 \leq \mathbb{E}[X_n^{push}] \leq \lceil \log_2 n \rceil + \ln n + 2.765 + o(1) \tag{2}$$

and, furthermore, for large deviations

$$P\big(X_n^{push} > \lceil \log_2 n \rceil + \ln n + 2.188 + r\big) \leq 2e^{-r}.$$

The new insight that allowed them to achieve this precise bound is that *push* on complete graphs can be closely described by $\log_2 n$ plus $1/n$ times the number of draws it takes to collect n distinct coupons. This is the so-called coupon collector's problem (CCP). The CCP is very well understood, for example it is known that in expectation it takes $n \cdot H_n$ draws, where H_n is the nth harmonic number, and the deviation from it can be described by Gumbel distributed random variables, see [30].

In 2000, Karp et al. [47] first introduced *push&pull* as a faster alternative to *push* or *pull* and showed that

$$X_n^{pp} = \log_3 n + O(\ln \ln n) \quad \text{with high probability.}$$

The most precise results for *pull* and *push&pull* on complete graphs to date are given by Benjamin Doerr and Anatolii Kostrygin in 2017 [25]. They developed a general framework to study a very broad range of randomized rumor spreading models that allows to derive sharp bounds for them. In particular, applied to complete graphs, they found that

$$\mathbb{E}[X_n^{pull}] = \log_2 n + \log_2 \ln n + O(1), \quad \text{and} \quad \mathbb{E}[X_n^{pp}] = \log_3 n + \log_2 \ln n + O(1)$$

together with large deviation bounds

$$P\big(\big|X_n^p - \mathbb{E}[X_n^p]\big| \geq r\big) \leq Ae^{-\alpha r} \quad \text{for all } r > 0 \text{, some } A, \alpha > 0 \text{ and all } p \in \{pull, push, pp\}.$$

We see that both *pull* and *push&pull* are quite a bit faster than *push*. This is due to the last phase of the protocol. Assume that there are $\varepsilon \cdot n$ vertices left to inform. Then *push* will inform about e^{-1} of them in the next round, on the other hand *pull* and *push&pull* will inform all but an ε fraction which translates into the double logarithmic term. Furthermore, Doerr and Kostrygin derived some results concerning robustness, this will be discussed in the later part about robustness.

Random graphs. The next result on Erdős-Rényi random graphs is by Fountoulakis et al. in 2010 [35]. They studied *push* on Erdős-Rényi random graphs $G_{n,p}$ on n vertices with edge probability p asymptotically larger than the connectivity threshold $\ln n/n$. Improving on the results in [32], they found that despite these graphs being much sparser than complete graphs they admit the same asymptotic runtime as was found in [39], i.e.

$$X^{push}(G_{n,p}) = \log_2 n + \ln n + o(\ln n) \quad \text{whp. for all } p = \omega(\ln n/n).$$

To consider random graphs that are even sparser one has to consider a different random graph model, as Erdős-Rényi random graphs are typically disconnected for all $p \leq \ln n/n$. Thus, Fountoulakis and Panagiotou in 2010 [36] considered *push* on random $d-$regular graphs $G_{n,d}$ for $d \geq 3$ which are known to be connected with high probability. They showed that

$$X^{push}(G_{n.d}) = \left(\frac{1}{\ln(2(1-1/d))} - \frac{1}{d\ln(1-1/d)}\right)\ln n + o(\log n) \qquad \text{for all } d \geq 3 \text{ whp.}$$

This gives a good understanding of the influence of the degree on the runtime of *push*.

Building on these the two previous results Panagiotou et al. in 2015 [52] computed $X^{push}(G_{n,p})$ for the wider range of $p = c\ln n/n$. For constant c these graphs are no longer as regular and the runtime becomes significantly slower,

$$X^{push}(G_{n,p}) = \log_2 n + c\ln\left(\frac{c}{c-1}\right)\ln n + o(\ln n) \quad \text{for all } p = c\ln n/n \text{ whp.}$$

Furthermore, they studied *push* on expander graphs, these are (almost) regular graphs that have a large spectral gap. Expander graphs are popular in computer science as well as mathematics, e.g. the survey [45]. In particular, they have a pseudorandom property in that the number of edges between two disjoint vertex sets U, V is approximated by $d|U||V|/n$, the expected number of edges in a random graph with edge probability d/n. We denote these expander graphs as $(n, \delta, \Delta, \lambda)$-graphs where n is the number of vertices, δ/Δ are minimal/maximal degree and λ is the second largest eigenvalue of the adjacency matrix. Improving a result from [36] they showed that for all $(n, \delta, \Delta, \lambda)$-graphs G with $\lambda = o(\Delta)$ and $\Delta/\delta = 1 + o(1)$

$$X^{push}(G) = \log_2 n + \ln n + o(\ln n) \quad \text{with high probability.} \tag{3}$$

This supports that density is not an important factor for *push* to be fast, it is more important that the edges are spread evenly and the graph is well connected. However, it is still necessary that the degree is not too small as we saw a larger runtime of constant degree random-regular graphs in [52].

Conductance and Expansion. In addition to these results on specific graph classes, there are results using general graph parameters to describe the runtime. Three parameters proved to be suitable; graph conductance, edge/vertex expansion and mixing time. A series of papers by Chierichetti et al. in 2010 [15, 14] and George Giakkoupis in 2011 [40] investigated the connection between conductance and randomized rumor spreading. Defining the conductance of a graph G by

$$\phi = \min_{S\subseteq V,\ \mathrm{vol}(S)\le|E|} \frac{e(S, V\setminus S)}{\mathrm{vol}(S)} \quad \text{where} \quad G = (V,E) \text{ and } \mathrm{vol}(S) = \sum_{v\in S} d(v)$$

they then showed that

$$X^{pp}(G) = O(\phi^{-1}\log n) \quad \text{with high probability.}$$

This result is tight, as there are graphs of conductance ϕ and diameter $\phi^{-1}\ln n$, which is a trivial lower bound for the runtime. For *push* or *pull* this bound does not hold in general. However, if minimal and maximal degree only differ by a constant factor then this bound applies as well.

Prompted by Chierichetti et al. [15], a second series of papers by Thomas Sauerwald and Alexandre Staufler in 2011 [60], Giakkoupis and Sauerwald in 2012 [43] and Giakkoupis in 2014 [41] took a deeper look at the effect of vertex expansion on randomized rumor spreading. Let the boundary of $S \subseteq V$ be $\partial S = \{v \in V \setminus S : v \in N(S)\}$ and define the vertex expansion of G by

$$\alpha = \min_{S\subseteq V,\ |S|\le n/2} \frac{|\partial S|}{|S|}.$$

They showed that

$$X^{pp}(G) = O\left(\frac{\log\Delta}{\alpha}\ln n\right) \quad \text{for all graph } G \text{ with vertex expansion } \alpha \text{ whp.}$$

A third graph parameter that is closely related to conductance and expansion, is the notion of mixing time. The mixing time $T_{\mathrm{mix}}(G)$ of a graph G is the number of steps needed for a random walk starting at some vertex to converge to the stationary distribution. Improving a result form Robert Elsässer and Thomas Sauerwald [28] in 2007, Sauerwald found in 2012 [59], that whp

$$X^{push}(G) = O(T_{\mathrm{mix}}(G) + \ln n).$$

These results using conductance, vertex expansion and mixing time support the impression gained from random or expander graphs that it is not important how many edges there are but how nicely they are distributed.

Robustness. When randomized rumor spreading was first introduced one of its main proposed advantages was its robustness. However, apart from the result in [32] discussed above, there are not that many theoretical results. The most popular notion of robustness that is investigated are random message transmission failures. That is, each message fails independently with some probability $p > 0$. Elsässer and Sauerwald showed in 2009 [29] that these failures can slow down *push* on any graph by a factor of at most $6/p$. More precise results are available for complete graphs and random graphs. In 2017 Doerr and Kostrygin [25] computed the runtime on complete graphs with independent message transmission failures. They found that

$$\begin{aligned}\mathbb{E}\big[X_n^{push}\big] &= \log_{1+p} n + \frac{1}{p}\ln n + O(1)\\ \mathbb{E}\big[X_n^{pull}\big] &= \log_{1+p} n - \frac{1}{\ln(1-p)}\ln n + O(1)\\ \mathbb{E}\big[X_n^{pp}\big] &= \log_{1+2p} n + \frac{1}{p-\ln(1-p)}\ln n + O(1).\end{aligned} \tag{4}$$

We see a dichotomy; *push* slows down proportional to p, but *pull* and *push&pull* are affected more significantly. That is, in the presence of constant message transmission failures *pull* and *push&pull* loose their speed advantage. The same bound that Doerr and Kostrygin showed for complete graphs is shown by Fountoulakis et al. [35] for Erdős-Rényi random graphs with edge probabilities asymptotically larger than the connectivity threshold.

Social graphs. There are several different ways to model social network graphs. One such way are random geometric graphs G_r, that are created by uniformly distributing n vertices in $[0, \sqrt{n}]^2$ and connecting all vertices with euclidean distance less then $r > 0$. In 2010, Bradonjic et al. [13] showed that with high probability *push* informs all vertices in the largest connected component of G_r within $O(\sqrt{n}/r + \ln n)$ rounds. Moreover, Friedrich et al. in 2012 [38] generalized this result to random geometric graphs in d-dimensions to find a runtime of $O(n^{1/d}/r + \ln n)$.

A different graph class, that got a lot of attention in this context are power-law or so-called scale-free graphs. They share a number of characteristics with social networks, like the small-world phenomena or the occurrence of hubs, i.e. vertices with degrees greatly exceeding the average. In 2011 first Chierichetti et al. [16] and then Doerr et al. [20] studied *push&pull* on preferential attachment graphs. They showed that *push&pull* informs all vertices in $O(\log n)$ rounds and if *push&pull* is modified such that no vertex contacts the same neighbour twice in a row than the runtime reduces to $O(\ln n/\ln\ln n)$. However, *push* and *pull* are much slower, they need $\Omega(n^\alpha)$ with positive probability for some $\alpha > 0$.

One year later, Fountoulakis et al. [37] studied *push&pull* on Chung-Lu random graphs with an underlying power-law distribution with exponent $2 < \beta < 3$ and showed that *push&pull* informs all but εn of all vertices in its giant component in $O(\ln\ln n)$ rounds.

On the other hand, as shown by Abbas Mehrabian and Ali Pourmiri in 2016 [50], on random k-trees *push&pull* performs much worse, needing $O\big(\ln^{1+2/k} n \cdot \ln\ln n \cdot f(n)\big)$ rounds to inform all but $o(n)$ vertices for an arbitrary function $f \in \omega(1)$. To inform all vertices there are at least $\Omega\big(n^{1/(k+3)}\big)$ rounds needed.

Asynchronous Rumor Spreading. The most popular version of randomized rumor spreading and a more practical one is asynchronous rumor spreading. The key difference to "synchronous" rumor spreading is, that the protocol is no longer performed in synchronous rounds, but vertices are equipped with independent rate-1 Poisson processes at which ticks push or pull operations are performed. Asynchronous rumor spreading was first introduced by Boyd et al. in 2006 [12] in the context of distributed averaging. We denote the asynchronous runtime by $Y^p(G, v)$ for $p \in \{push, pull, pp\}$, any graph $G = (V, E)$ and initially informed vertex $v \in V$. A series of papers relates asynchronous rumor spreading to synchronous rumor spreading. Acan et al. in 2015 [1] showed that asynchronous rumor spreading is at most a factor of $O(\ln n)$ slower than synchronous rumor spreading and

$$O(\ln^{-1} n) \cdot Y^p(G, v) \leq X^p(G, v) \leq O(n^{2/3}) \cdot Y^p(G, v) \quad \text{for all } G, v, p.$$

The bound on the right hand side was sharpened by Giakkoupis et al. in 2016 [42] and finally by Angel et al. in 2017 [6] to $O(n^{1/3} \ln^{2/3} n)$. This is tight (up to a factor of order $\ln n$) as shown by the string of diamonds in [1].

Furthermore, asynchronous rumor spreading has been studied on several graph classes. In 2017 Konstantinos Panagiotou and Leo Speidel [54] studied asynchronous *push&pull* on random graphs. They showed that both whp and in expectation

$$Y^{push}(G_{n,p}) = \ln n + o(\ln n) \qquad \text{for all } p = \omega(\ln n/n).$$

Furthermore, on these random graphs they studied independent message transmission failures with failure probability $q > 0$ and found that the runtime is slowed down by a factor of $1/q$. Additionally, they considered node failures, where they marked a set of vertices B as faulty, that will not perform any *push* operations nor respond to *pull* requests. They found that the runtime to inform all non-faulty vertices is asymptotically not affected assuming that B has sublinear cardinality and the initially informed vertex is not faulty.

On scale-free networks, asynchronous *push&pull* has been studied as well. Fountoulakis et al. [37] showed in 2012 a constant time to inform almost all vertices on Chung-Lu random graphs and Doerr et al. in 2012 [21] showed a runtime of $O(\sqrt{\log n})$ to inform all but $o(n)$ vertices on preferential attachment graphs.

The most relevant result on asynchronous rumor spreading in the context of this thesis is by Svante Janson in 1999 [46]. He studied minimal paths on complete graphs with random edge weights. For Exp(1) distributed weights and scaled by n, this problem is equivalent to asynchronous *push* (and *pull*). Among other results, Janson derived the asymptotic distribution of Y_n, showing that in distribution

$$Y_n - 2 \ln n \to G_1 + G_2$$

where G_1, G_2 are independent Gumbel-distributed random variables, see Section 7 for a definition. Furthermore, this implies that

$$\mathbb{E}[Y_n] = 2 \ln n + 2\gamma + o(1)$$

where γ denotes the Euler-Mascheroni constant. This is to date the only result that precisely describes the limiting distribution of any randomized rumor spreading protocol.

Other variants. A second popular variant of (synchronous) randomized rumor spreading is *Quasirandom Rumor Spreading* introduced by Doerr et al. in 2008 [22]. This variant reduces randomness in that vertices no longer choose neighbours independently at random, but each vertex is equipped with a cyclic list of its neighbours. Neighbours are then contacted in the order of the list, starting form a random position. Reducing randomness is desirable from a computational perspective, as producing random bits is a costly operation. Doerr et al. ([22] and [23]) showed that quasirandom rumor spreading performs as well as (fully) randomized rumor spreading on complete, random, random regular and Ramanujan graphs. Furthermore, they showed that quasirandom rumor spreading can be faster on very sparse graphs like random graphs with edge probability only slightly higher than the connectivity threshold or hypercubes. Intuitively one might assume that this reduction in randomness comes at the cost of robustness, however, Doerr et al. in 2013 [24] showed that on complete graphs the robustness of (fully random) *push* is the same as that of quasirandom *push*.

On complete graphs quasirandom *push* has been studied more thoroughly by Angelopoulos et al. in 2009 [7] as well as by Nikolaos Fountoulakis and Anna Huber in 2010 [34]. They showed a runtime of $\log_2 n + \ln n + \Theta(\ln \ln n)$ to inform all vertices, which almost matches the bound of Pittel [56] $\log_2 n + \ln n + O(1)$ for (fully random) *push*.

There are two more interesting variants worth mentioning. First Daum et al. in 2018 [18] considered *pull* where nodes are restricted to answering one call only. If the answered call is chosen randomly than they find a runtime of $O(\psi(G) \ln n)$ with $\psi(G) \leq \Delta/\delta$. Going in the other direction Panagiotou et al. in 2015 [53] considered *push* and *push&pull* where nodes are allowed to make multiple calls per round. If the number of calls is power-law distributed with exponent $\beta \in (2, 3)$, then the runtime of *push&pull* is $\Theta(\ln \ln n)$ and if $\beta = 3$ then *push* has a runtime of $O(\ln n / \ln \ln n)$.

Other Information Spreading Protocols. So far, we only considered the setting, when there is one piece of information present in the network. Now we will loosen that assumption and allow for additional (contradicting) pieces of information. We look at this from the context of opinion forming, each piece of information is an opinion and a vertex being informed means that it formed its opinion. Furthermore, we will not start with only one vertex having an opinion, but with a set, the so-called experts or early-adopters, in which each vertex has some opinion. This is motivated by an influential work of Everett Rogers in 2003 [58]. He studied the diffusion of innovations through a network and his key insight is, that there is a group of early-adopters forming a first hand experience about said innovation and all other participant in the network form their opinion according to those of the early-adopters.

There are countless theoretical as well as empirical results on opinion forming, so we will only focus on a small sample that is far from an exhaustive list.

In 1992 and 1998 Bikhchandani et al. [8, 9] introduced the theory of informational cascades to explain fragile "mood swings" of the populace. Each member has an intrinsic preference between some binary choice, but also values being part of the majority. When voting is done sequentially (and publicly), participants have to decide whether to vote their personal preference or to conform to the majority. This gives rise to so called informational cascades, where one opinion rapidly spreads through the network. They study the question of when cascades occur and which opinion dominates.

This so-called binary decision making with externalities was also studied by Watts in [62]. He computed the probability of cascades occurring when the underlying network is given as a random graph.

In 2012, Alon et al. [2] showed that when considering binary decision making with externalities, sequential voting is strictly better than simultaneous voting, in the sense that with higher probability the most preferred alternative will dominate.

Feldman et al. in 2014 [33] also studied binary decision making with externalities, where there is some ground truth that people are more likely to favor. Voting is no longer sequential, but asynchronous and the vote may be updated at any time. They looked at the question of consensus, where the process converges to one opinion only. They showed that on sparse expander graphs consensus on the ground truth is very likely, but they also give graphs where no consensus occurs or it converges to the falsehood.

A different approach is the study of so-called group recommendation. Starting from some opinionated experts, the goal is to find a suggestion for all uninformed participant, based on the opinions of their nearest experts. This was studied by Anderson et al. in 2008 [5], where they took an axiomatic approach to these recommendation systems. Furthermore, they examined how adversarial influences can affect the recommendations. A similar question, from the view-point of bribery, is raised by Umberto Grandi and Paolo Turrini in 2016 [44]. They analysed the effect of bribery and found conditions that make a group recommendation system bribery-proof. See also the survey of Faliszewski et al. in 2016 [31] for more details.

Alon et al. in 2015 [3] combined the adversarial approach of Andersen et al. and Grandi et al. and the study of binary decision making. There is a binary opinion, either the truth ("1") or a falsehood ("0") and each expert is initially labelled "1" or "0". However, which vertices are the experts and what opinions they have is determined by an adversary. Setting the number of experts $\mu n,\ 0 < \mu < 1/2$ and the number of "1" labeled experts $(1/2 + \delta)\mu n,\ 0 < \delta < 1/2$ as fixed, the *strong-adversary* is allowed to choose the experts and their opinions. The *weak-adversary* is allowed to choose the experts, but their opinions are distributed randomly, i.e. with probability $(1/2 + \delta)$ label "1" and "0" otherwise.

Starting with the experts as the only vertices having an option, the opinions are then disseminated through the network by a simple majority rule, where each vertex without opinion takes the opinion of the majority of its opinionated neighbours (ties are broken uniformly at random, this includes vertices without expert neighbours)

The key question Alon et al. answered, is given some graph, whether an adversary can influence the experts in such a way that after the dissemination the majority of all vertices believes the falsehood. Conversely they say that a graph G is *robust* if for any distribution of the experts after the dissemination more than half of all vertices know the truth.

Alon et al. showed that most random-graphs and expander graphs are robust against the *strong-adversary*. Furthermore, if the maximal degree is sufficiently bounded than the graph is robust against the weak adversary.

Additionally, they proposed a modified dissemination process. Instead of all vertices breaking their ties uniformly at random, only vertices with at least one opinionated neighbour do so. All vertices without expert neighbours stay uninformed. Considering all opinionated vertices as experts, this process is then repeated until all possible vertices have formed an opinion. Alon et al. finally considered the question, whether iterativity harms or helps the adversary. For both types of adversary, they provided examples where it helps the adversary. Additionally, they provided an example where iterativity hinders the weak adversary. Finally, they conjectured that there is no graph that is robust against the iterative strong adversary but not robust against the non-iterative strong adversary.

There has been further research in that precise model. In his PhD-thesis Rami Daknama [17] studied the local resilience of being robust against the strong adversary on random graphs. He showed that Erdős-Rényi random graphs with edge probability of $\omega(\ln n/n)$ remain robust even if one is allowed to delete up to an $2(1 - \mu + 2\delta\mu)\delta/(1 + 2\delta)$ fraction of edges at each vertex.

4.2 Contribution

In this subsection we detail the individual contributions of the works included in this thesis.

4.2.1 Robustness of Randomized Rumor Spreading

We saw that there are several different types of robustness results in the literature. Random message transmission failures [29, 35, 25], edge failures [32] and the non-dependence on the density [52]. The goal of this contribution is to study the question of robustness more thoroughly. If we modify a graph, we say it is robust, if *push*/*pull*/*push&pull* has asymptotically the same runtime as on the complete graph, up to terms of order $o(\ln n)$. The first part of this paper extends the validity of (3) shown in [52] to *pull* and *push&pull*, as well as incorporating random message transmission failures with probability $q > 0$. We set $X^p(G, q)$ as the runtime of the algorithm $p \in \{push,\ pull,\ push\&pull\}$ on the graph G with message transmission failure probability q. We find that all three algorithms are robust in terms of non-dependence on the density, as we recover the runtimes on complete graphs (4) found in [25]. That is, one can delete a vast amount of edges as long as the remaining edges are distributed regularly. We quantify this by studying expander sequences G_n, which are sequences of $(n, \delta_n, \Delta_n, \lambda)$-graphs as defined in (3). We find that

$$X^p(G_n, q) = c_p(q) \ln n + o(\ln n), \quad p \in \{push,\ pull,\ push\&pull\} \quad whp,$$

where we set for $q \in (0, 1)$

$$c_{push}(q) = \frac{1}{\ln(1+q)} + \frac{1}{q}, \quad c_{pull}(q) = \frac{1}{\ln(1+q)} - \frac{1}{\ln(1-q)}, \quad c_{pp}(q) = \frac{1}{\ln(1+2q)} + \frac{1}{q - \ln(1-q)}.$$

For $q = 1$ this can be continuously extended to be consistent with the model without random message transmission failures. Thus, density is not important, but good expansion properties are. However, how important are they? Can we also reduce expansion and still retain robustness? To answer this question we introduce the notion of *local resilience*. Local resilience of a graph with respect to a property quantifies the maximal fraction of edges that can be deleted at each vertex such that the property remains. This was first introduced by Sudakov and Vu in [61], where they studied the resilience of random graphs with respect to hamiltonicity, chromatic number and perfect matchings.

Applying this definition to our setting, we find that there is a mixed bag. Let $\varepsilon > 0$ and $G_n(\varepsilon)$ be a graph that is obtained by deleting up to $1/2 - \varepsilon$ of all edges at any node starting with an expander graph G_n. Then on the one hand *pull* and *push&pull* remain robust, that is

$$X^{pull}\big(G_n(\varepsilon)\big) = c_{pull}(1)\ln n + o(\ln n)$$

and even

$$X^{pp}\big(G_n(\varepsilon)\big) \le c_{pp}(1)\ln n + o(\ln n).$$

This might come as a surprise, that *push&pull* is not slowed down, but might even speed up when deleting edges. However, we think this is expected, as for example on the star graph *push&pull* deterministically has a runtime of at most 2. On the other hand *push* is already slowed down when only a small fraction is deleted, i.e

$$\text{for every } \varepsilon > 0 \text{ there is } \eta > 0 \;\text{ such that }\quad X^{push}\big(G_n(\varepsilon)\big) > (1+\eta)c_{push}(1)\ln n.$$

Going one step further, we look at the local resilience when random transmission failures are added. This changes the picture once more. *pull* keeps its local resilience of $1/2$, however, *push&pull* does not,

$$X^{pull}\big(G_n(\varepsilon), q\big) = c_p(q)\ln n + o(\ln n)$$

but

$$X^{pp}\big(G_n(\varepsilon), q\big) = \big(c_{pp}(q) + f(\varepsilon, q)\big)\ln n + o(\ln n).$$

See Figure 1 for a plot of $f(\varepsilon, q)$.

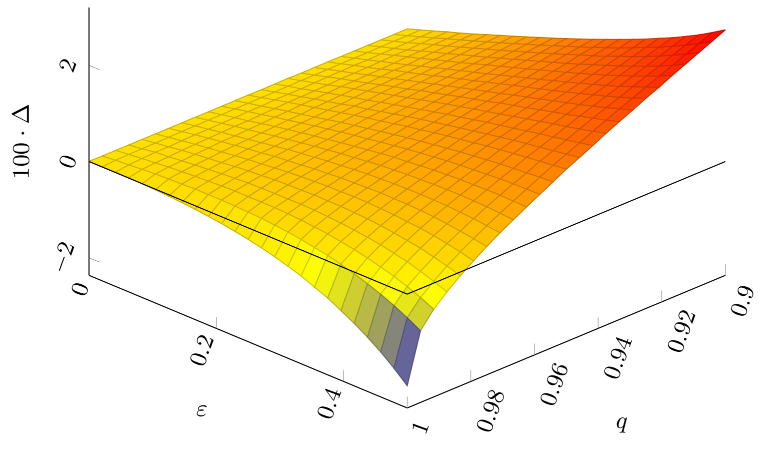

Figure 1: Plotted values of $f(\varepsilon, q)$ in $T_{pp}(G(n,\varepsilon), q) - c_{pp}(q)\log n = f(\varepsilon, q)\log n + o(\log n)$, for $0.9 < q < 1$ and $0 < \varepsilon < 1/2$.

We utilize three very different approaches to study each of these algorithms that differ from typical approaches. For *push* we show that in the beginning it behaves the same way as on a graph without deleted edges by giving a suitable coupling, then we prove that in the end it may be slowed down by giving an example for all $\varepsilon > 0$. For *pull* instead of analysing the growth of informed vertices, we look at the growth of edges between informed and uninformed vertices and then relating these quantities gives the desired result. The study of *push&pull* is the most involved one. We give a suitable partition such that pairwise most parts behave like random regular bipartite graphs. On these pairs, we analyse the growth of informed vertices, which yields a linear recursion. We solve that recursion by computing the largest eigenvalue of the underlying matrix.

Apart from these results, enhancing all existing results on robustness, we introduce two novel methods to the study of randomized rumor spreading. Usually, one analyses the expected growth of the number of informed vertices after performing one round of the algorithm. To translate this to a statement about the actual number of informed vertices one needs to show concentration around the expectation. Usually this is done by bounding the variance, which can be rather involved. Moreover, this approach only works if the expected growth only depends on the number of informed vertices and not on the precise vertices being informed, which is the case for graphs with significant irregularities. To facilitate the general approach and overcome the problem of missing regularity, we employ to very strong results.

First, consider so-called self-bounding functions [49]. A function $f : \mathcal{X}^m \to \mathbb{R}$ is self-bounding if there are functions $f_i : \mathcal{X}^{m-1} \to \mathbb{R}$ such that for all $x \in \mathcal{X}^m$

$$0 \leq f(x) - f_i(x^{(i)}) \leq 1$$

and

$$\sum_{i=1}^{m} \left(f(x) - f_i(x^{(i)}) \right) \leq f(x)$$

where $x^{(i)} = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_m) \in \mathcal{X}^{m-1}$ is obtained by dropping the i-th entry of x. These functions have the striking property that, for independent random variables $X_1, \dots X_m$, the variance is uniformly bounded by its expectation

$$\mathrm{Var}[f(X_1, \dots, X_m)] \leq \mathbb{E}[f(X_1, \dots, X_m)].$$

We show that the number of informed vertices after performing one round of *push*/ *pull*/ *push&pull* can be described as a self-bounding function of independent random variables. This immediately yields that its expectation is bounded by its variance, allowing an easy application of Chebychev's inequality. Furthermore, an even stronger result holds true, self-bounding functions admit exponential concentration bounds [11], but in our cases the simpler ones are sufficient. As the characterization of being self-bounding is independent of the underlying graph, this approach is applicable in a wide range of settings. We are positive that this will facilitate further research in this area.

Secondly, in order to handle the irregularities introduced by deleting edges, we employ one more very powerful tool. This time from a quite unexpected field, namely extremal graph theory. Szemerédis regularity lemma [57] guarantees that for any graph with sufficiently many vertices (here sufficiently many is a *very* large constant) there is a partition, such that pairwise almost all parts behave like bipartite random regular graphs. On these pairs we have all the nice properties we need in order to study our protocols in the round-based manner that we usually do. This results in a linear recursion (with dimension equal to the cardinality of the partition), which we can asymptotically solve by analysing the maximal eigenvalue of the underlying matrix. There is one drawback, however, this is only true for dense graph, where the degree is linear in the number of vertices. However, we are certain that it also holds true in the sparse case, as there is an extension of the regularity lemma to that case.

4.2.2 Asymptotics for Push on the Complete Graph

There have been substantial efforts to characterise the runtime of rumor spreading protocols as precisely as possible. So far, complete characterisation of the runtime (its distribution and related quantities) has only been achieved for asynchronous *push* on complete graphs by studying the equivalent problem of minimal paths with exponentially distributed random edge weights. In 1999 Janson [46] solved this problem, giving both its asymptotic distribution Y_n

$$Y_n - 2\ln n \to G_1 + G_2$$

where G_1, G_2 are Gumbel distributed random variables, as well as derived quantities like the expectation

$$\mathbb{E}[Y_n] = 2\ln n + 2\gamma + o(1).$$

γ denotes the Euler-Mascheroni constant. For synchronous *push*, first Pittel [56] then Doerr and Künnemann [26] characterised the distribution and expectation of its runtime up to constant errors. The goal of this work is to close the remaining gap, that is we completely characterise the asymptotic distribution and derived properties like the limiting distribution and its expectation. Will synchronous *push* behave the same same way as asynchronous *push* or are there fundamental differences between these variants?

We answer that question in a surprising way. Let X_n^{push} be the runtime of *push* on the complete graph. Then there is a continuous and periodic function c and a Gumbel distributed random variable G such that asymptotically and in distribution X_n^{push} is given by

$$\lceil G + \log_2 n + \ln n + \gamma + c(\log_2 n - \lfloor \log_2 n \rfloor) \rceil. \tag{5}$$

This result is interesting in several ways. First of all it implies that the asymptotic distribution X_n only converges on subsequences $(n_i)_{i\in\mathbb{N}}$ where the rational parts of $(\log_2 n_i)_{i\in\mathbb{N}}$ and $(\ln n_i)_{i\in\mathbb{N}}$ converge to constants. Furthermore, c is a continuous, 1-periodic function with amplitude $\approx 10^{-9}$, see Figure 2. If we want to give the limiting distribution, we need to restrict ourselves to the

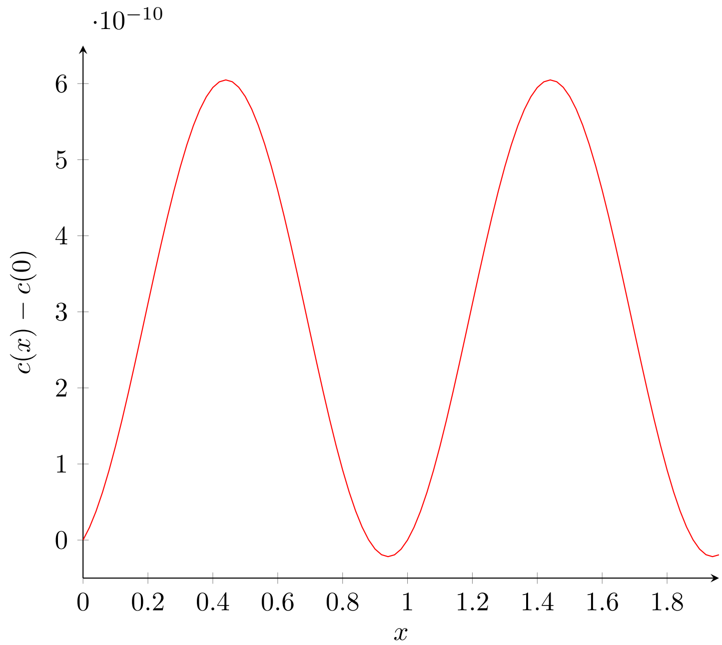

Figure 2: The function $c(x) - c(0)$, $c(0) \approx 0.105$, plotted for values of x between 0 and 2. The periodic nature of the function and its small amplitude are evident.

aforementioned subsequences, i.e. for $x, y \in [0, 1)$ subsequences n_i such that $\log_2 n_i - \lfloor \log_2 n_i \rfloor \to x$ and $\ln n_i - \lfloor \ln n_i \rfloor \to y$. Such subsequences exist and on them the limiting distribution is

$$X_{n_i}^{push} - (\lfloor \log_2 n_i \rfloor + \lfloor \ln n_i \rfloor) \to \mathrm{dGum}(-x - y - c(x)),$$

where we write dGum for a discretised Gumbel distribution. From this we can derive (sharp) bounds on the expectation, improving (2) best possible,

$$\log_2 n + \ln n + 1.18242 \leq \mathbb{E}[X_n^{push}] \leq \log_2 n + \ln n + 1.18263$$

as the non-constant part in the expectation is of order 10^{-4}.

The proof is based on ideas introduced by Pittel [56] as well as Doerr and Künnemann [26]. On the complete graph, *push* behaves differently depending on the number of informed vertices. At first, the number of informed vertices perfectly doubles every round, as the probability of any informed vertex to choose a previousy chosen vertex is small as long as there are $\leq \sqrt{n}$ informed vertices. This is also observed in the well known birthday paradox. Once there are $\approx \sqrt{n}$ informed vertices, we can again use the notion of self-bounding functions, see Section 4.2.1, to find that the number of informed vertices is closely concentrated around its expected value in all following rounds. This allows us to describe the behaviour of *push* by a deterministic recurrence relation, see (1), given by iteratively taking expected values.

Once there is only a small linear fraction of uninformed vertices left, the behaviour changes once more, the deterministic recurrence relation breakes down and a different phenomena takes over. Instead of vertices choosing random neighbours, we can also view it as independent sampling from (almost) all vertices and once a vertex is sampled we set it as informed. This point of view has the advantage, that we know how often we have to sample in order to draw each vertex at least once, as that is described by the Coupon Collector's Problem. This problem is solved and we know all desired quantities, like the expectation or the limiting distribution. So far, this was also utilized by Doerr and Künnemann [26], however, we are not directly interested in the number of samples or "pushes" but in the number of rounds. As we only have a small number of uninformed vertices left, a sensible approach is to rescale by n and that is exactly what Doerr and Künnemann did in their work. Unfortunately this is not precise enough, but finding the right relation of "pushes" to rounds is the last building block that allowed us to precisely characterise the limiting distribution.

Summarizing, this work concludes the search for the asymptotic distribution of *push* on the complete graph. Furthermore, on all sequences with convergent distribution we additionally give the limiting distribution as well as its moments.

4.2.3 Asymptotics for Pull on the Complete Graph

After characterising the asymptotic distribution of *push* on the complete graph, a natural question is to ask whether the same can be done with *pull*. This has attracted much less attention though, the best result in this direction is by Doerr and Kostrygin in 2017 [25].They showed that for the runtime of *pull* on the complete graph X_n^{pull}

$$X_n^{pull} = \log_2 n + \log_2 \ln n + O(1) \quad \text{with high probability and in expectation.}$$

Compared to *push*, *pull* behaves quite a bit differently. *push* had its randomness at the end of the process; i.e at first the number of informed vertices perfectly doubles every round, then they follow a deterministic recurrence relation and at the end they behave like the Coupon Collector's Problem. Another benefit is, that the Coupon Collector's Problem is very well understood as we know its limiting distribution and related properties. In total, this allowed for a nice description of the asymptotic distribution. In contrast, *pull* has its randomness at the start of the process. Looking at the informed vertices I_t before some round t we have

$$|I_{t+1}| = |I_t| + \mathrm{Bin}\big(n - |I_t|, |I_t|/n\big).$$

That is, every uninformed vertex gets informed independently with probability $|I_t|/n$. As long as $|I_t| = o(n)$, this binomial distribution can be approximated by a Poisson one, i.e.

$$|I_{t+1}| \approx |I_t| + \mathrm{Po}(|I_t|)$$

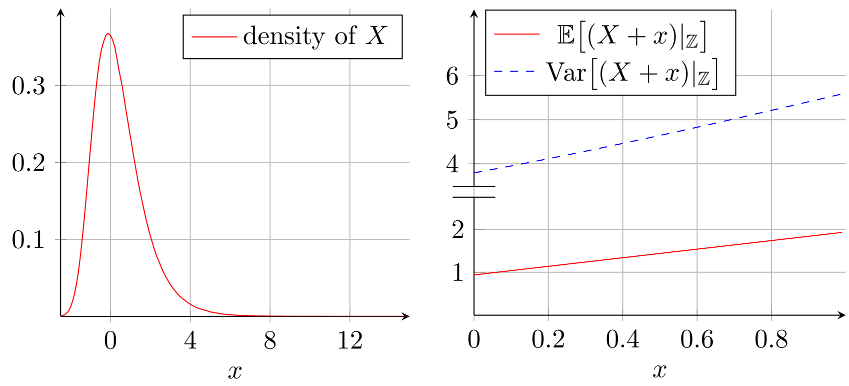

Figure 3: The left plot shows an estimate of the density of the random variable X. The right plot shows, as a function of $x \in [0,1]$, the estimated expectation and variance of the random variable $(X+x)|_{\mathbb{Z}}$.

which behaves like 2^t in expectation. Appropriately normalized (by 2^{-t}), we show that this sequence is a convergent martingale, thus, there is a random variable H such that almost surely (and in $\mathcal{L}^2$)

$$2^{-t}|I_t| \to H \quad \text{for } t \to \infty.$$

Compared to *push* we have $2^t H$ informed vertices for $t \approx (1/3)\log_2 n$ instead of 2^t for some random variable H. Next, *pull* can also be described by a deterministic recurrence relation, until there are $\approx \sqrt{n}$ uninformed vertices remaining. Then there occurs an interesting phenomena, once there are less then $\approx \sqrt{n}$ uninformed vertices for the first time the algorithm will terminate with high probability in the next round.

In this paper we show that there is a continuous random variable $X = -\log_2 H$ such that in distribution X_n^{pull} is given by

$$\lfloor \log_2 n + \log_2 \ln n + X + 1 \rfloor \tag{6}$$

Set $(X+x)|_{\mathbb{Z}}$ as X discretized and translated by x, that is

$$P\big((X+x)|_{\mathbb{Z}} \leq k\big) := P(X \leq k - x), \qquad k \in \mathbb{Z}.$$

Again, (6) only converges on suitable subsequences $(n_i)_{i\in\mathbb{N}}$, where $\log_2 n_i + \log_2 \ln n_i - \lfloor \log_2 n_i + \log_2 \ln n_i \rfloor \to x$ for some $x \in [0,1)$. Thus, on these subsequences we can give the limiting distribution

$$X_{n_i}^{pull} - \lfloor \log_2 n_i + \log_2 \ln n_i \rfloor \to (X+x)|_{\mathbb{Z}} \quad \text{for } i \to \infty \quad \text{in distribution}$$

and all of its moments

$$\mathbb{E}\left[\big(X_{n_i}^{pull} - \lfloor \log_2 n_i + \log_2 \ln n_i \rfloor\big)^k\right] \to \mathbb{E}\left[\big((X+x)|_{\mathbb{Z}}\big)^k\right] \qquad \forall\, k \in \mathbb{N}.$$

To give more detailed statements requires further knowledge about the distribution/ moments of X that we do not have. However, we have some numerically results, see Figure 3. To get these numbers, we have drawn one million instances of $-\log_2 H_{28}$ as a substitute for X and with these approximated its density, as well as first and second moments of its discretized version $(X+x)|_{\mathbb{Z}}$.

There is one unexpected similarity between *push* and *pull*. Consider the following function

$$g = g^{(1)} : [0,1] \to [0,1], \quad x \mapsto xe^{x-1}$$

and its iterations

$$g^{(i)} : [0,1] \to [0,1], \quad g^{(i)} = g \circ g^{(i-1)}, i \geq 2.$$

This function plays a fundamental role in the distributions of the runtime of *push* and *pull* on the complete graph. The function c in (5), shown in Figure 2, is defined as

$$c(x) = -x + \lim_{a\to\infty, a\in\mathbb{N}} \lim_{b\to\infty, b\in\mathbb{N}} -a + b + \ln\left(g^{(b)}(1 - 2^{-a-x})\right),$$

and the random variable X in (6) is given by $X = -\log_2 H$ and H has characteristic function

$$\varphi(x) = \lim_{t\to\infty} g^{(t)}\left(e^{ix2^{-t}}\right).$$

To be able to state our results in a nicer way, expressing $g^{(t)}$ in a closed form would be helpful. However, it appears to have no closed form, at least we were not able to find one. We leave it as an open problem to further study $g^{(t)}, t \in \mathbb{N}$.

4.2.4 The Effect of Iterativity on Adversarial Opinion Forming

We want to model the spread of opinions through a network using the theory of early-adopters, introduced by Rogers in 2003 [58]. His insight was, that new information is introduced into the network by so-called early-adopters, or experts, that form a first hand experience. The other participants in the network, however, do not form a first hand experience themselves, but form their opinion according to those of the experts they know.

Unfortunately, these experts are a prime target for any adversary that would like to influence the opinion of the network. It might only take a few key experts to spread a falsehood, to successfully sway the majority of the network from true to false opinion. Two real world examples are:

The common practice of online vendors to buy positive reviews for their products by either directly giving monetary incentives to reviewers or providing them with free products. In particular, on Amazon in certain product categories, like Bluetooth speakers and headphones, ReviewMeta finds more than 50% of reviews to be fake, see [27].

The second example is a newly opened restaurant, that in its opening phase invites food critics to try and rate the restaurant. However, when those critics dine at the restaurant, the restaurant puts in more effort than it would when catering to a regular customer, e.g. by providing better quality food and service.

In these two examples we saw different sorts of adversaries: The seller that bought his reviews could select the reviewers as well as their opinion, the restaurant owner could choose the critics but could only influence their opinion to some degree.

Alon et al. in 2015 [3] proposed a model that implements the above ideas. Let $0 < \mu, \delta < 1/2$ and $G = (V, E)$ be a graph on n vertices. The set of experts is given by $\mathcal{E} \subseteq V$, $|\mathcal{E}| = \mu n$. Moreover, $\mathcal{E}$ is partitioned into $\mathcal{E}_1 \cup \mathcal{E}_0$, the set of experts knowing the truth that are labelled "1" as well as the set of experts believing a falsehood, labelled "0". How $\mathcal{E}_1$ and $\mathcal{E}_0$ are chosen depends on the type of adversary.

Like in the first example, the *strong-adversary* is allowed to choose $\mathcal{E}$ as well as the partition $\mathcal{E} = \mathcal{E}_1 \cup \mathcal{E}_0$ as long as $|\mathcal{E}_1| = (1/2 + \delta)\mu n$ and consequently $|\mathcal{E}_0| = (1/2 - \delta)\mu n$.

The *weak-adversary* models the second example. The adversary is allowed to choose $\mathcal{E}$, but the partition $\mathcal{E} = \mathcal{E}_1 \cup \mathcal{E}_0$ is chosen randomly. Each vertex is added to $\mathcal{E}_1$ with probability $(1/2 + \delta)$ and to $\mathcal{E}_0$ otherwise.

There are two different ways the experts can disseminate their opinions. In the *non-iterative* setting, every vertex in $V \setminus \mathcal{E}$ takes the majority of its expert neighbours. That is, it is labelled "1" if it has more neighbours in $\mathcal{E}_1$ than in $\mathcal{E}_0$. Vertices with an equal number of neighbours in $\mathcal{E}_1$ and $\mathcal{E}_0$ are labelled "1" with probability $1/2$ and "0" otherwise.

In the *iterative* setting, vertices also take the majority opinion, however, ties are only broken, if there is at least one expert vertex involved. Vertices that don't have any expert neighbours stay uninformed. This process is than iterated by considering all vertices labelled "1" as $\mathcal{E}_1$ and vertices labelled "0" as $\mathcal{E}_0$. The process stops once all possible vertices are labelled.

The main question that Alon et al. posed in this context is whether iterativity helps the adversary. Are graphs less robust, if the dissemination process is iterated compared to a single-round dissemination?

For the weak-adversary Alon et al. answered this question by giving two examples. First, they showed that for suitable parameters μ and δ a path is robust against the non-iterative weak adversary but not robust against the iterative weak adversary. Secondly, they considered the union of a star graph with a suitable expander. They found that in the non-iterative setting, putting one "0" labelled vertex in the center of the star graph can swing the majority, as most vertices in the expander will be decided uniformly at random, however, in the iterative setting, the expander will be deterministically decided, resulting in a clear majority for "1"s.

The path that we just saw, is also an example where iterativity helps the strong adversary. However, Alon et al. did not provide a second example for the strong-adversary, they rather conjectured that there is no such example, that is they conjectured that iterativity always helps the strong adversary.

In this work, we refute that conjecture. For $\mu = \delta = 1/5$ the graph G in Figure 4 is robust against the iterative strong adversary but not against the non-iterative strong adversary.

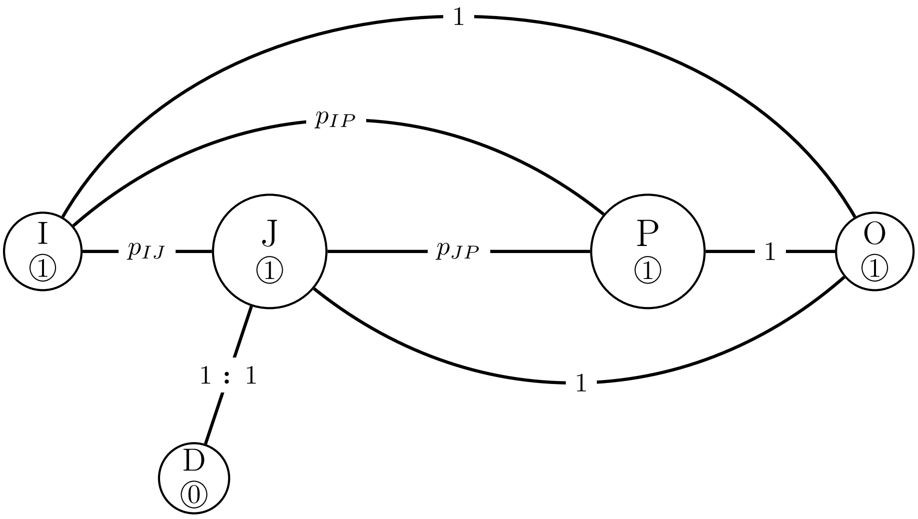

Figure 4: This figure shows the graph G. We set $|I| = (719/2000)n$, $|P| = (879/2000)n$, $|D| = 10^{-4}n$, $|I| = (7/50)n$ and $|O| = (3/50)n$. Furthermore, let $p_{IJ} = p_{JP} = 3/7 + 10^{-4}$ and $p_{IP} = 3/7 - 10^{-6}$. The numbers on the edges and in the nodes give the probability that each edge is present between the components and in the components respectively. For example any edge with one node in I and one in J exists independently with probability p_{IJ}. Every node in D has exactly one distinct neighbour in J and no other neighbours, i.e. every node in D has degree 1 and no two nodes in D have a common neighbour.

References

[1] H. Acan, A. Collevecchio, A. Mehrabian, and N. Wormald. On the push&pull protocol for rumor spreading. *SIAM Journal on Discrete Mathematics*, 31(2):647–668, 2017.

[2] N. Alon, M. Babaioff, R. Karidi, R. Lavi, and M. Tennenholtz. Sequential voting with externalities: herding in social networks. In *Proceedings of the 13th ACM Conference on Electronic Commerce (EC). ACM*, 2012.

[3] N. Alon, M. Feldman, O. Lev, and M. Tennenholtz. How robust is the wisdom of the crowds? In *Twenty-Fourth International Joint Conference on Artificial Intelligence*, pages 2055–2061, 2015.

[4] I. Amores-Sesar, C. Cachin, and J. Mićić. Security Analysis of Ripple Consensus. In *24th International Conference on Principles of Distributed Systems (OPODIS 2020)*, volume 184, pages 10:1–10:16, 2021.

[5] R. Andersen, C. Borgs, J. Chayes, U. Feige, A. Flaxman, A. Kalai, V. Mirrokni, and M. Tennenholtz. Trust-based recommendation systems: an axiomatic approach. In *Proceedings of the 17th international conference on World Wide Web*, pages 199–208, 2008.

[6] O. Angel, A. Mehrabian, and Y. Peres. The string of diamonds is tight for rumor spreading. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2017)*, 2017.

[7] S. Angelopoulos, B. Doerr, A. Huber, and K. Panagiotou. Tight bounds for quasirandom rumor spreading. *the electronic journal of combinatorics*, 16(1):R102, 2009.

[8] S. Bikhchandani, D. Hirshleifer, and I. Welch. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of political Economy*, 100(5):992–1026, 1992.

[9] S. Bikhchandani, D. Hirshleifer, and I. Welch. Learning from the behavior of others: Conformity, fads, and informational cascades. *Journal of economic perspectives*, 12(3):151–170, 1998.

[10] K. P. Birman, M. Hayden, O. Ozkasap, Z. Xiao, M. Budiu, and Y. Minsky. Bimodal multicast. *ACM Transactions on Computer Systems (TOCS)*, 17(2):41–88, 1999.

[11] S. Boucheron, G. Lugosi, and P. Massart. On concentration of self-bounding functions. *Electronic Journal of Probability*, 14(none):1884 – 1899, 2009.

[12] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. Randomized gossip algorithms. *IEEE transactions on information theory*, 52(6):2508–2530, 2006.

[13] M. Bradonjic, R. Elsasser, T. Friedrich, T. Sauerwald, and A. Stauffer. Efficient broadcast on random geometric graphs. In *Proceedings of the Twenty-First Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1412–1421, 2010.

[14] F. Chierichetti, S. Lattanzi, and A. Panconesi. Almost tight bounds for rumour spreading with conductance. In *Proceedings of the forty-second ACM symposium on Theory of computing*, pages 399–408. ACM, 2010.

[15] F. Chierichetti, S. Lattanzi, and A. Panconesi. Rumour spreading and graph conductance. In *Proceedings of the twenty-first annual ACM-SIAM symposium on Discrete Algorithms*, pages 1657–1663. SIAM, 2010.

[16] F. Chierichetti, S. Lattanzi, and A. Panconesi. Rumor spreading in social networks. *Theoretical Computer Science*, 412(24):2602–2610, 2011.

[17] R. Daknama. Theoretical runtime bounds for information spreading and a new vehicle routing algorithm, Oktober 2018.

[18] S. Daum, F. Kuhn, and Y. Maus. Rumor spreading with bounded in-degree. *Theoretical Computer Science*, 810:43–57, 2018.

[19] A. Demers, D. Greene, C. Hauser, W. Irish, J. Larson, S. Shenker, H. Sturgis, D. Swinehart, and D. Terry. Epidemic algorithms for replicated database maintenance. In *Proceedings of the sixth annual ACM Symposium on Principles of distributed computing*, pages 1–12. ACM, 1987.

[20] B. Doerr, M. Fouz, and T. Friedrich. Social networks spread rumors in sublogarithmic time. In *Proceedings of the forty-third annual ACM symposium on Theory of computing*, pages 21–30, 2011.

[21] B. Doerr, M. Fouz, and T. Friedrich. Asynchronous rumor spreading in preferential attachment graphs. In *Scandinavian Workshop on Algorithm Theory*, pages 307–315. Springer, 2012.

[22] B. Doerr, T. Friedrich, and T. Sauerwald. Quasirandom rumor spreading. In *Proceedings of the nineteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 773–781, 2008.

[23] B. Doerr, T. Friedrich, and T. Sauerwald. Quasirandom rumor spreading: Expanders, push vs. pull, and robustness. In *International Colloquium on Automata, Languages, and Programming*, pages 366–377. Springer, 2009.

[24] B. Doerr, A. Huber, and A. Levavi. Strong robustness of randomized rumor spreading protocols. *Discrete Applied Mathematics*, 161(6):778–793, 2013.

[25] B. Doerr and A. Kostrygin. Randomized Rumor Spreading Revisited. In *44th International Colloquium on Automata, Languages, and Programming, ICALP 2017, July 10-14, 2017, Warsaw, Poland*, pages 138:1–138:14, 2017.

[26] B. Doerr and M. Künnemann. Tight Analysis of Randomized Rumor Spreading in Complete Graphs. In *Proceedings of the Meeting on Analytic Algorithmics and Combinatorics*, pages 82–91, 2014.

[27] E. Dwoskin and C. Timberg. How merchants use Facebook to flood Amazon with fake reviews, 2018. `https://www.washingtonpost.com/business/economy/how-merchants-secretly-use-facebook-to-flood-amazon-with-fake-reviews/2018/04/23/5dad1e30-4392-11e8-8569-26fda6b404c7_story.html?`, visited 24.09.20.

[28] R. Elsässer and T. Sauerwald. Broadcasting vs. mixing and information dissemination on cayley graphs. In *Annual Symposium on Theoretical Aspects of Computer Science*, pages 163–174. Springer, 2007.

[29] R. Elsässer and T. Sauerwald. On the runtime and robustness of randomized broadcasting. *Theoretical Computer Science*, 410(36):3414–3427, 2009.

[30] P. Erdös and A. Rényi. On a classical problem of probability theory. *Magyar Tud. Akad. Mat. Kutató Int.Közl. 6*, page 215–220, 1961.

[31] P. Faliszewski, J. Rothe, and H. Moulin. *Control and Bribery in Voting*, page 146–168. Cambridge University Press, 2016.

[32] U. Feige, D. Peleg, P. Raghavan, and E. Upfal. Randomized broadcast in networks. *Random Structures & Algorithms*, 1(4):447–460, 1990.

[33] M. Feldman, N. Immorlica, B. Lucier, and S. M. Weinberg. Reaching consensus via non-bayesian asynchronous learning in social networks. *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 192–208, 2014.

[34] N. Fountoulakis and A. Huber. Quasirandom rumor spreading on the complete graph is as fast as randomized rumor spreading. *SIAM Journal on Discrete Mathematics*, 23(4):1964–1991, 2010.

[35] N. Fountoulakis, A. Huber, and K. Panagiotou. Reliable broadcasting in random networks and the effect of density. In *INFOCOM, 2010 Proceedings IEEE*, pages 1–9. IEEE, 2010.

[36] N. Fountoulakis and K. Panagiotou. Rumor spreading on random regular graphs and expanders. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 560–573. Springer, 2010.

[37] N. Fountoulakis, K. Panagiotou, and T. Sauerwald. Ultra-fast rumor spreading in social networks. In *Proceedings of the twenty-third annual ACM-SIAM symposium on Discrete Algorithms*, pages 1642–1660. SIAM, 2012.

[38] T. Friedrich, T. Sauerwald, and A. Stauffer. Diameter and broadcast time of random geometric graphs in arbitrary dimensions. *Algorithmica*, 67(1):65–88, 2012.

[39] A. M. Frieze and G. R. Grimmett. The shortest-path problem for graphs with random arc-lengths. *Discrete Applied Mathematics*, 10(1):57–77, 1985.

[40] G. Giakkoupis. Tight bounds for rumor spreading in graphs of a given conductance. In *28th International Symposium on Theoretical Aspects of Computer Science, STACS 2011, March 10-12, 2011, Dortmund, Germany*, pages 57–68, 2011.

[41] G. Giakkoupis. Tight Bounds for Rumor Spreading with Vertex Expansion. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2014, Portland, Oregon, USA, January 5-7, 2014*, pages 801–815, 2014.

[42] G. Giakkoupis, Y. Nazari, and P. Woelfel. How asynchrony affects rumor spreading time. In *Proceedings of the 2016 ACM Symposium on Principles of Distributed Computing*, pages 185–194, 2016.

[43] G. Giakkoupis and T. Sauerwald. Rumor spreading and vertex expansion. In *Proceedings of the twenty-third annual ACM-SIAM symposium on Discrete Algorithms*, pages 1623–1641. SIAM, 2012.

[44] U. Grandi and P. Turrini. A network-based rating system and its resistance to bribery. In *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*, pages 301–307, 2016.

[45] S. Hoory, N. Linial, and A. Wigderson. Expander graphs and their applications. *Bulletin of the American Mathematical Society*, 43(4):439–561, 2006.

[46] S. Janson. One, two and three times log n/n for paths in a complete graph with random weights. *Combinatorics, Probability and Computing*, 8(4):347–361, 1999.

[47] R. M. Karp, C. Schindelhauer, S. Shenker, and B. Vöcking. Randomized Rumor Spreading. In *41st Annual Symposium on Foundations of Computer Science, FOCS 2000, 12-14 November 2000, Redondo Beach, California, USA*, pages 565–574, 2000.

[48] D. Kempe, A. Dobra, and J. Gehrke. Gossip-based computation of aggregate information. In *Annual Symposium on Foundations of Computer Science - Proceedings*, pages 482– 491, 11 2003.

[49] G. Lugosi. Concentration-of-measure inequalities, 2004.

[50] A. Mehrabian and A. Pourmiri. Randomized rumor spreading in poorly connected small-world networks. *Random Structures & Algorithms*, 49(1):185–208, 2016.

[51] S. Nakamoto. Bitcoin: A peer-to-peer electronic cash system. Technical report, Manubot, 2019.

[52] K. Panagiotou, X. Pérez-Giménez, T. Sauerwald, and H. Sun. Randomized Rumour Spreading: The Effect of the Network Topology. *Combinatorics, Probability & Computing*, 24(2):457–479, 2015.

[53] K. Panagiotou, A. Pourmiri, and T. Sauerwald. Faster rumor spreading with multiple calls. *The Electronic Journal of Combinatorics*, 22(1):P1–23, 2015.

[54] K. Panagiotou and L. Speidel. Asynchronous Rumor Spreading on Random Graphs. *Algorithmica*, 78(3):968–989, 2017.

[55] C. Patsonakis and M. Roussopoulos. Revisiting Asynchronous Rumor Spreading in the Blockchain Era. In *2019 IEEE 25th International Conference on Parallel and Distributed Systems (ICPADS)*, pages 284–293, Dec 2019.

[56] B. Pittel. On Spreading a Rumor. *SIAM J. Appl. Math.*, 47(1):213–223, Mar. 1987.

[57] V. Rödl and M. Schacht. Regularity lemmas for graphs. In *Fete of combinatorics and computer science*, pages 287–325. Springer, 2010.

[58] E. M. Rogers. *Diffusion of innovations (5. ed.)*. Free Press, 2003.

[59] T. Sauerwald. On mixing and edge expansion properties in randomized broadcasting. *Algorithmica*, 56(1):51–88, 2010.

[60] T. Sauerwald and A. Stauffer. Rumor spreading and vertex expansion on regular graphs. In *Proceedings of the twenty-second annual ACM-SIAM symposium on Discrete Algorithms*, pages 462–475. SIAM, 2011.

[61] B. Sudakov and V. Vu. Local resilience of graphs. *Random Structures & Algorithms*, 33(4):409–433, 2008.

[62] D. Watts. A simple model of global cascades on random networks. *Proceedings of the National Academy of Sciences of the United States of America*, 99(9):5766–5771, 2002.

5 Robustness of Randomized Rumor Spreading

This chapter is the published version of

> Daknama, R., Panagiotou, K., & Reisser, S. (2021). Robustness of randomized rumour spreading. *Combinatorics, Probability and Computing*, 30(1), 37-78.

The published version is online at `https://doi.org/10.1017/S0963548320000310`.

My own contribution. This paper is joint work with fellow PhD student Rami Daknama and my supervisor Konstantinos Panagiotou. We developed all results in joint discussion. Rami Daknama wrote down Theorem 1.6 a) as well as Theorem 1.7 a), which take up about 1/8th of the paper. His part is also included in his thesis [17]. All other parts of this paper are written by me, in particular Theorem 1.4, Theorem 1.5, Theorem 1.6 b), Theorem 1.7 b) and their proofs as well as the final editing. All of it includes continual improvements by Konstantinos Panagiotou.

Combinatorics, Probability and Computing (2021), **30**, pp. 37–78
doi:10.1017/S0963548320000310

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Robustness of randomized rumour spreading[1]

Rami Daknama, Konstantinos Panagiotou* and Simon Reisser

Institute for Mathematics, Ludwig-Maximilians-Universität München, 80333 Munich, Germany
*Corresponding author. Email: kpanagio@math.lmu.de

(Received 27 September 2019; revised 8 June 2020; accepted 15 May 2020; first published online 12 August 2020)

Abstract

In this work we consider three well-studied broadcast protocols: *push*, *pull* and *push&pull*. A key property of all these models, which is also an important reason for their popularity, is that they are presumed to be very robust, since they are simple, randomized and, crucially, do not utilize explicitly the global structure of the underlying graph. While sporadic results exist, there has been no systematic theoretical treatment quantifying the robustness of these models. Here we investigate this question with respect to two orthogonal aspects: (adversarial) modifications of the underlying graph and message transmission failures.

We explore in particular the following notion of *local resilience*: beginning with a graph, we investigate up to which fraction of the edges an adversary may delete at each vertex, so that the protocols need significantly more rounds to broadcast the information. Our main findings establish a separation among the three models. On one hand, *pull* is robust with respect to all parameters that we consider. On the other hand, *push* may slow down significantly, even if the adversary may modify the degrees of the vertices by an arbitrarily small positive fraction only. Finally, *push&pull* is robust when no message transmission failures are considered, otherwise it may be slowed down.

On the technical side, we develop two novel methods for the analysis of randomized rumour-spreading protocols. First, we exploit the notion of self-bounding functions to facilitate significantly the round-based analysis: we show that for any graph the variance of the growth of informed vertices is bounded by its expectation, so that concentration results follow immediately. Second, in order to control adversarial modifications of the graph we make use of a powerful tool from extremal graph theory, namely Szemerédi's Regularity Lemma.

2020 MSC Codes: Primary 05C85; Secondary 68R10

1. Introduction

Randomized broadcast protocols are highly relevant for data distribution in large networks of various kinds, including technological, social and biological networks. Among many others there are three basic models in the literature, introduced in [9], [19] and [27], namely *push*, *pull* and *push&pull* (or *pp* for short). Consider a connected graph in which some vertex holds a piece of information; we call this vertex (initially) informed. All three models have the common characteristic that they proceed in rounds. In the *push* model, in every round every informed vertex chooses a neighbour independently and uniformly at random and informs it; this of course only

[1] An extended abstract of this paper was published in the *Proceedings of the European Symposium on Algorithms 2019 (ESA '19)*.

has an effect if the target vertex was previously uninformed. Conversely, in the *pull* model every round every *un*informed vertex chooses a neighbour independently and uniformly at random and asks for the information. If the asked vertex has the information, then the asking vertex becomes informed as well. The third model, *push&pull*, combines both worlds: in each round, each vertex chooses a neighbour independently and uniformly at random, and if one of two vertices is informed, then afterwards both become so. We additionally assume that each message transmission succeeds independently with probability $q \in (0, 1]$. For these algorithms, the main parameter that we consider is the random variable that counts how many rounds are needed until all vertices are informed, and we call these quantities the *runtimes* of the respective algorithms.

In the remainder we will denote the runtime of *push* by $T_{push}(G, v, q)$, where G is the underlying graph, initially the vertex v is informed, and we have a transmission success probability of $q \in (0, 1]$. Analogously we denote the runtimes of *pull* and *push&pull* by $T_{pull}(G, v, q)$ and $T_{pp}(G, v, q)$ respectively. If the choice of v does not matter we will omit it in our notation. The most basic case is when G is the complete graph K_n with n vertices. Then (see *e.g.* Doerr and Kostrygin [11]) it is known that, for $\mathcal{P} \in \{push, pull, pp\}$ and $q \in (0, 1]$ in expectation and with probability tending to 1 as $n \to \infty$,

$$T_{\mathcal{P}}(K_n, q) = c_{\mathcal{P}}(q) \log n + o(\log n),$$

where, for $q \in (0, 1)$,

$$c_{push}(q) := \frac{1}{\log(1+q)} + \frac{1}{q},$$
$$c_{pull}(q) := \frac{1}{\log(1+q)} - \frac{1}{\log(1-q)},$$
$$c_{pp}(q) := \frac{1}{\log(1+2q)} + \frac{1}{q - \log(1-q)},$$

and where we set $c_{\mathcal{P}}(1) := \lim_{q \to 1} c_{\mathcal{P}}(q)$. If q is clear from the context, we write $c_{\mathcal{P}}$ instead of $c_{\mathcal{P}}(q)$. In fact, the results in [11] and also [12] are much more precise, but the stated forms will be sufficient for what follows.

Contribution and related work. In this article our focus is on quantifying the *robustness* of all three models. Indeed, robustness is a key property that is often attributed to them, since they are simple, randomized and, crucially, do not exploit explicitly the structure of the underlying graph (apart from considering the neighbourhoods of the vertices, of course). Clearly the runtime can vary tremendously between different graphs with the same number of vertices. Hence it is essential to understand the impact of structural graph characteristics on the runtime of rumour-spreading algorithms.

One result in this spirit for the *push* model was shown in [28]. Roughly speaking, in that paper it is shown that even on graphs with low density, if the edges are distributed rather uniformly, then *push* is as fast as on the complete graph. This can be interpreted as a robustness result: starting with a complete graph, one can delete a vast amount of edges, and as long as this is done rather uniformly, the runtime of *push* is affected insignificantly. To state the result more precisely, we need the following notion.

Definition 1.1 ($(n, \delta, \Delta, \lambda)$-graph)**.** Let G be a connected graph with n vertices that has minimum degree δ and maximum degree Δ. Let $\mu_1 \geqslant \mu_2 \geqslant \cdots \geqslant \mu_n$ be the eigenvalues of the adjacency matrix of G, and set

$$\lambda = \max_{2 \leqslant i \leqslant n} |\mu_i| = \max\{|\mu_2|, |\mu_n|\}.$$

We will call G an $(n, \delta, \Delta, \lambda)$-graph.

In this paper we are interested in the case where G gets large, that is, when $n \to \infty$. Hence all asymptotic notation in this paper is with respect to n; in particular, 'with high probability', or w.h.p. for short, means with probability $1 - o(1)$ when $n \to \infty$.

Definition 1.2 (expander sequence)**.** Let $\mathcal{G} = (G_n)_{n\in\mathbb{N}}$ be a sequence of graphs, where G_n is a $(n, \delta_n, \Delta_n, \lambda_n)$-graph for each $n \in \mathbb{N}$. We say that $\mathcal{G}$ is an *expander sequence* if $\Delta_n/\delta_n = 1 + o(1)$ and $\lambda_n = o(\Delta_n)$.

Note that if we consider any sequence $\mathcal{G} = (G_n)_{n\in\mathbb{N}}$ of graphs this always implicitly defines δ_n, Δ_n and λ_n as in Definition 1.2. Expander graphs have found numerous applications in computer science and mathematics; see *e.g.* the survey [25]. If $\mathcal{G}$ is an expander sequence, then intuitively this means that for n large enough, the edges of G_n are rather uniformly distributed. For a more formal statement see Lemma 2.7. Moreover, note that our definition of expander sequences excludes the case when Δ_n is bounded. This is actually a necessary condition for our robustness results to hold; see [13]. With all these definitions at hand we can state the result from [28] that quantifies the robustness of *push* with respect to the network topology, that is, the runtime is asymptotically the same as on the complete graph K_n.

Theorem 1.1. *Let $\mathcal{G} = (G_n)_{n\in\mathbb{N}}$ be an expander sequence. Then w.h.p.*

$$T_{push}(G_n) = c_{push}(1) \log n + o(\log n).$$

Apart from expander sequences, results in the form of Theorem 1.1 (where the asymptotic runtimes of one or more of these algorithms are determined) were also shown for sufficiently dense Erdős–Rényi random graphs [16], random regular graphs [15] as well as hypercubes [28]. Moreover, the order of the runtime on various models that describe social networks was investigated. The Chung–Lu model was studied in [17], preferential attachment graphs were explored in [10], and geometric graphs were examined in [18]. A somewhat different approach is to derive general runtime bounds that hold for all graphs and depend only on some graph parameter, *e.g.* conductance [6, 20], vertex expansion [21] or diameter [5, 14, 23]. Furthermore, several variants of *push*, *pull* and *push&pull* were studied. These include vertices being restricted to answer only one *pull* request per round [7], vertices being allowed to contact multiple neighbours per round [11, 28], vertices not calling the same neighbour twice [10] and asynchronous versions [1, 2, 4, 29]. Finally, besides [11], robustness of these rumour-spreading algorithms with respect to message transmission failures was also studied by Elsässer and Sauerwald in [13]. It was shown for any graph that if a message fails with probability $1 - p$, then the runtime of *push* increases at most by a factor of $6/p$.

In this work our focus is on three subjects concerning the robustness of rumour spreading. Our first (and not unexpected) result extends Theorem 1.1 to the runtimes of *pull* and *push&pull*. In particular, we show that none of the three protocols slows down or speeds up on graphs with good expansion properties compared to its runtime on the complete graph. This motivates us to investigate how severely a graph with good expansion properties has to be modified to increase the respective runtimes.

In our second contribution, which is also the main result and which differs from what was treated in previous works, we propose and study a novel approach to quantifying robustness. In particular, we investigate the impact of adversarial edge deletions, where we use the well-known concept of *local resilience*; see *e.g.* [8, 31]. To be specific, we explore up to which fraction of edges an adversary needs to be allowed to delete at each vertex to slow down the process by a significant amount of time, *i.e.* by $\Omega(\log n)$ rounds. Here we discover a surprising dichotomy in the following sense. On the one hand, we show that neither *pull* and *push&pull* can be slowed down by such adversarial edge deletions – in essentially all but trivial cases, where the fraction is so large that the graph may become (almost) disconnected. On the other hand, we demonstrate that even a small number of edge deletions is sufficient to slow down *push* by $\Omega(\log n)$ rounds. In other words, we find that in contrast to *pull* and *push&pull*, the *push* protocol is not resilient to adversarial deletions and lacks (in this specific sense) the robustness of the other two protocols.

As our third subject, we generalize the previous results by additionally considering message transmission failures that occur independently with probability $1 - q \in [0, 1)$. On the positive side, we show that for arbitrary $q \in (0, 1]$, all three algorithms inform *almost* all vertices at least as fast as in an expander sequence in spite of adversarial edge deletions. However, if we want to inform all vertices, only *pull* is not slowed down by adversarial edge deletions for all values of q; *push* can be slowed down as before, and *push&pull* is a mixed bag, in that for $q = 1$ it cannot be slowed down whereas for $q < 1$ it can. Furthermore, in general it is also possible to speed up *push&pull* by deleting edges, which is however not surprising as the star-graph deterministically finishes in at most two rounds.

Summarizing, this work enhances previous (robustness) results, particularly the ones concerning precise asymptotic runtimes and random transmission failures. Crucially, we introduce and study the concept of local resilience as a method to investigate robustness. However, apart from that, in this paper we develop two new general methods for the analysis of rumour-spreading algorithms.

- The most common approach in the current literature for the study of the runtime is to determine the expected number of newly informed vertices in one or more rounds and to show concentration, for example by bounding the variance. Achieving this, however, is often quite complex and makes laborious and lengthy technical arguments necessary. Here we use the theory of *self-bounding* functions (see Section 2), which allows us to cleanly upper-bound the variance by the *expected value*. The argument works for all three investigated algorithms and the bound is valid for all graphs. We are certain that this method will also facilitate future work on the analysis of rumour-spreading algorithms.
- Studying the robustness of the protocols is a challenging task, as the adversary (as described previously) has various options to modify the graph, for example by introducing a high variance in the degrees of the vertices; this turns out to be particularly problematic in the case of *push&pull*. Here we demonstrate that such types of irregularities can be handled universally by applying a powerful tool from a completely different area, namely extremal graph theory. In particular, we use Szemerédi's Regularity Lemma (see *e.g.* [30]), which allows us to partition the vertex set of a graph such that nearly all pairs of sets in the partition behave nearly like perfect regular bipartite graphs. This allows us to apply our methods on these regular pairs; eventually we obtain a linear recursion that can be solved by analysing the maximal eigenvalue of the underlying matrix.

1.1 Results

Our first result addresses the question about how fast rumours spread on expander graphs; in order to obtain a concise statement, the occurrence of independent message transmission failures is also considered.

Theorem 1.2. *Let $\mathcal{G} = (G_n)_{n\in\mathbb{N}}$ be an expander sequence and let $q \in (0, 1]$. Then w.h.p.*

(a) $T_{push}(G_n, q) = c_{push}(q) \log n + o(\log(n))$,
(b) $T_{pull}(G_n, q) = c_{pull}(q) \log n + o(\log(n))$,
(c) $T_{pp}(G_n, q) = c_{pp}(q) \log n + o(\log(n))$.

The first statement is an extension of Theorem 1.1 and its proof is a straightforward adaptation of the proof in [28]. We omit it. The contribution here is the proof of (b) and (c). Next we consider the case with edge deletions in addition to the message transmission failures.

Theorem 1.3. *Let $0 < \varepsilon < 1/2, q \in (0, 1]$ and $\mathcal{G} = (G_n)_{n\in\mathbb{N}}$ be an expander sequence. Let $\tilde{\mathcal{G}} = (\tilde{G}_n)_{n\in\mathbb{N}}$ be such that each $\tilde{G}_n$ is obtained by deleting edges of G_n such that each vertex keeps at least a $(1/2 + \varepsilon)$ fraction of its edges. Then w.h.p.*

(a) $T_{pull}(\tilde{G}_n, q) = c_{pull}(q) \log n + o(\log n)$,

(b) $T_{pp}(\tilde{G}_n, 1) \leqslant c_{pp}(1) \log n + o(\log n)$, *when additionally assuming that* $\delta(G_n) \geqslant \alpha n$ *for some constant* $0 < \alpha \leqslant 1$.

This result demonstrates unconditionally the robustness of *pull*, and conditionally on $q = 1$ the robustness of *push&pull* on dense graphs, in the case of edge deletions, that is, the runtime is asymptotically the same as in the complete graph. Moreover, we even show that *push&pull* may profit from edge deletions in contrast to being slowed down; see Subsection 3.6 for an example. The proof of this result, especially the statement about *push&pull*, is rather involved, since the original graph may become quite irregular after the edge deletions. Here we use, among many other ingredients, the aforementioned decomposition of the graph given by Szemerédi's Regularity Lemma.

Note that Theorem 1.3 does not consider *push* and *push&pull* (when $q \neq 1$) at all. Indeed, our next result states that in these cases the behaviour is rather different and that the algorithms may be slowed down.

Theorem 1.4. *Let $\varepsilon > 0$ and $q \in (0, 1]$. Then there is an expander sequence $\mathcal{G} = (G_n)_{n\in\mathbb{N}}$ and a sequence of graphs $\tilde{\mathcal{G}} = (\tilde{G}_n)_{n\in\mathbb{N}}$ with the following properties. Each $\tilde{G}_n$ is obtained by deleting edges of G_n such that each vertex keeps at least a $(1 - \varepsilon)$ fraction of its edges. Moreover, w.h.p.*

(a) $T_{push}(\tilde{G}_n, q) \geqslant c_{push}(q) \log n + \varepsilon/(2q) \log n + o(\log n)$,

(b) $T_{pp}(\tilde{G}_n, q) \geqslant c_{pp}(q) \log n + (\varepsilon/(8q) - \varepsilon q^3/5) \log n + o(\log n)$.

Nevertheless, not all hope is lost. On the positive side, the next result states that *push* and *push&pull* are able to inform *almost* all vertices as fast as on the complete graph in spite of adversarial edge deletions. In this sense, we obtain an almost-robustness result for these cases.

Theorem 1.5. *Let $0 < \varepsilon < 1/2, q \in (0, 1]$ and $\mathcal{G} = (G_n)_{n\in\mathbb{N}}$ be an expander sequence. Let $\tilde{\mathcal{G}} = (\tilde{G}_n)_{n\in\mathbb{N}}$ be such that each $\tilde{G}_n$ is obtained by deleting edges of G_n such that each vertex keeps at least a $(1/2 + \varepsilon)$ fraction of its edges. For $\mathcal{P} \in \{push, pp\}$, let $\tilde{T}_{\mathcal{P}}$ denote the number of rounds needed to inform at least $n - n/\log n$ vertices. Then w.h.p.*

(a) $\tilde{T}_{push}(\tilde{G}_n, q) = \log_{1+q}(n) + o(\log n)$,

(b) $\tilde{T}_{pp}(\tilde{G}_n, q) \leqslant \log_{1+2q}(n) + o(\log n)$, *when additionally assuming that* $\delta(G_n) \geqslant \alpha n$ *for some* $0 < \alpha \leqslant 1$.

We conjecture that there is also a version of Theorem 1.5(b) that is true for *push&pull* on sparse graphs; to be precise we conjecture that in the setting of Theorem 1.5(b), $\tilde{T}_{pp}(\tilde{G}_n) \leqslant \log_{1+2q}(n) + o(\log n)$, without further restrictions on G_n, that is, *push&pull* cannot be slowed down, informing *almost* all vertices.

As a final remark, note that Theorems 1.3 and 1.5 are tight in the sense that if an adversary may delete up to half of the edges at each vertex, then there are expander graphs that become disconnected such that their components have linear size. On those graphs a linear fraction of the vertices will remain uninformed forever.

Outline. The rest of this paper is structured as follows. In Section 2 we collect and prove several important facts; this part of the paper also contains our technical contribution concerning the analysis through self-bounding functions. In Section 3.1 we show that *pull* is as fast on expanders with (or without) deleted edges as it is on the complete graph. Section 3.2 treats *push&pull* on expanders without deleted edges. In the remaining subsections we focus on the cases that may be slowed down by edge deletions. In Section 3.3 we show that adversarial edge deletions cannot slow down the time until *push* has informed almost all vertices, by giving a coupling to the case without edge deletions. Conversely, in Section 3.4 we show that the time until *push* has informed all vertices can be slowed down by edge deletions, even if only a few edges are deleted. Then in Section 3.5 we show that *push&pull* informs almost all vertices of dense graphs fast in spite of adversarial edge deletions. We utilize a version of Szemerédi's Regularity Lemma to get a well-behaved partition of the vertex set that is suitable for performing a round-based analysis. However, if $q < 1$, adversarial edge deletions can slow down or speed up the time until *push&pull* has informed all vertices for nearly all values of q; we show this in Section 3.6.

Further notation. Let $G = (V, E)$ denote a graph with vertex set V and edge set $E \subseteq \binom{V}{2}$. We will denote the set of neighbours of any vertex $v \in V$ by $N_G(v)$ or by $N(v)$, and we will denote its degree by $d_G(v) := |N_G(v)|$ or by $d(v)$; δ_G or δ and Δ_G or Δ denote the minimum and maximum degree of G. Similarly the neighbourhood of any set of vertices $S \subseteq V$ is defined by $N_G(S) := \cup_{v \in S} N_G(v)$. Furthermore, let $U, W \subseteq V$ with $U \cap W = \emptyset$ be two disjoint vertex sets; then $E(U, W) = E_G(U, W)$ denotes the set of edges with one vertex in U and one vertex in W and let $e(U, W) = e_G(U, W) := |E_G(U, W)|$. With $E_G(U)$ we denote the set of edges with both vertices in U; $e_G(U) := |E_G(U)|$. For any round $t \in \mathbb{N}$ and $\mathcal{P} \in \{push, pull, pp\}$, we let $I_t^{(\mathcal{P})}(G)$ denote the set of vertices of G informed by *push*, *pull* and *push&pull* respectively at the beginning of round t and $|I_1^{(\mathcal{P})}| = 1$; if the underlying graph is clear from the context we will omit it; if we consider a sequence of graphs $\mathcal{G} = (G_n)_{n \in \mathbb{N}}$ and a sequence of times $t = (t(n))_{n \in \mathbb{N}}$, then

$$I_t^{(\mathcal{P})}(\mathcal{G}) = (I_{t(n)}^{(\mathcal{P})}(G_n))_{n \in \mathbb{N}}$$

is also a sequence. Similarly, $U_t^{(\mathcal{P})} := V \backslash I_t^{(\mathcal{P})}$ denotes the set of uninformed vertices. By log we refer to the natural logarithm. For any event A we will write $\mathbb{E}_t[A]$ instead of $\mathbb{E}[A|I_t]$ for the conditional expectation and $\mathbb{P}_t[A]$ instead of $\mathbb{P}[A|I_t]$ for the conditional probability. Finally we want to clarify our use of Landau symbols. Let $a, b \in \mathbb{R}$ and f be a function. The terms $a \leqslant b + o(f)$ and $a \geqslant b - o(f)$ mean that there exist positive functions $g, h \in o(f)$ such that $a \leqslant b + g$ and $a \geqslant b - h$. Consequently $a = b + o(f)$ means that there exists a positive function $g \in o(f)$ such that $a \in [b - g, b + g]$.

2. Tools and techniques

In this section we collect and prove statements about our protocols and properties of expander sequences. We begin by applying the previously mentioned notion of self-bounding functions to derive universal and simple-to-apply concentration results for our random variables, *i.e.* the number of informed vertices after a particular round. Then we extend the concentration results to more than one round. In the last part we recall the well-known Expander Mixing Lemma and utilize it to derive properties (weak expansion, path enumeration) for the case where we delete edges from our graphs.

Self-bounding functions. Our main technical new result in this section is the following bound on the variance for the number of informed vertices in any given round; it is true for any graph and any set of informed vertices.

Lemma 2.1. *Let G be a graph, $t \in \mathbb{N}$ and $I_t = I_t^{(\mathcal{P})}(G)$ for $\mathcal{P} \in \{push, pull, pp\}$. Then*

$$Var[|I_{t+1}| \, |I_t] \leqslant \mathbb{E}[|I_{t+1}| \, |I_t].$$

Lemma 2.1 follows directly from Lemmas 2.2 and 2.3. Before stating them we introduce the notion of self-bounding functions.

Definition 2.1 (self-bounding function)**.** Let X be a set and $m \in \mathbb{N}$. A non-negative function $f : X^m \to \mathbb{R}$ is self-bounding if there exist functions $f_i : X^{m-1} \to \mathbb{R}$ such that, for all $x_1, \ldots, x_m \in X$ and all $i = 1, \ldots, m$,

$$0 \leqslant f(x_1, \ldots, x_m) - f_i(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_m) \leqslant 1$$

and

$$\sum_{1\leqslant i\leqslant m} (f(x_1, \ldots, x_m) - f_i(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_m)) \leqslant f(x_1, \ldots, x_m).$$

A striking property of self-bounding functions is the following bound on the variance.

Lemma 2.2 ([3])**.** *For a self-bounding function f and independent random variables $X_1, \ldots, X_m$, $m \in \mathbb{N}$,*

$$Var[f(X_1, \ldots, X_m)] \leqslant \mathbb{E}[f(X_1, \ldots, X_m)].$$

Lemma 2.3. *Let G be a graph, $t \in \mathbb{N}$, and let $I_t = I_t^{(\mathcal{P})}(G)$ for $\mathcal{P} \in \{push, pull, pp\}$. Then, conditional on I_t, there exists $m \in \mathbb{N}$, independent random variables $X_1, \ldots, X_m$ and a self-bounding function $f = f^{(\mathcal{P})}$ such that $|I_{t+1}| = f(X_1, \ldots, X_m)$.*

Proof. We will prove in detail the result for *push*, and then we show what needs to be modified in order to obtain the statement in the case of *pull* and *push&pull*. Let $I_t = I_t^{(push)}$, $n \in \mathbb{N}$ be the number of vertices of G, *i.e.* $V = [n]$, and $f\colon [n]^{|I_t|} \to \mathbb{R}$ with

$$(x_1, \ldots, x_{|I_t|}) \mapsto |I_t| + \sum_{1\leqslant k\leqslant |I_t|} \mathbb{1}[x_k \in U_t]\mathbb{1}[\forall\, \ell < k\colon x_k \neq x_\ell].$$

Moreover, let $(X_i)_{1\leqslant i\leqslant |I_t|}$ be independent random variables, where X_i is a uniformly random neighbour of the ith vertex – according to an arbitrary ordering – in I_t. We argue that $f(X_1, \ldots, X_{|I_t|}) = |I_{t+1}|$. Consider $v \in I_t$; then v is counted by the $|I_t|$ term in f. For $v \in I_{t+1} \backslash I_t$, let $v_1, \ldots, v_s \in I_t$, $s \in \mathbb{N}$ be the informed vertices with random neighbour v in round t, *i.e.* $X_{v_1} = \cdots = X_{v_s} = v$ and $X_u \neq v$ for all other $u \in I_t$. Assume further that $v_1 < v_2 < \cdots < v_s$. For $k = v_1$ the term $\mathbb{1}[X_k \in U_t]\mathbb{1}[\forall\, \ell < k\colon x_k \neq x_\ell] = 1$ as $X_{v_1} = v \in U_t$ and for all $i \leqslant v_1$ it holds that $X_i \neq X_{v_i}$. For $k = v_r$, $2 \leqslant r \leqslant s$ the term $\mathbb{1}[\forall\, \ell < k\colon x_k \neq x_\ell] = 0$ as $v_1 < v_r$ and $X_{v_1} = X_{v_r} = v$. Thus every vertex $v \in I_{t+1} \backslash I_t$ is counted exactly once by f. Further, set

$$f_i(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_{|I_t|}) = |I_t| + \sum_{k=1,k\neq i}^{|I_t|} \mathbb{1}[x_k \in U_t]\mathbb{1}[\forall\, j < k, j \neq i\colon x_j \neq x_k], \quad 1 \leqslant i \leqslant |I_t|.$$

The function f_i arises from f by leaving the ith variable out of consideration, that is, the push of the ith vertex has no effect. Then by definition $f - f_i \in \{0, 1\}$ for all $1 \leqslant i \leqslant |I_t|$, and in fact we have

$$f - f_i = \mathbb{1}[x_i \in U_t]\mathbb{1}[\forall\, j \neq i\colon x_i \neq x_j].$$

This quantity is precisely the difference in informed vertices after round t, assuming the ith vertex did not push. Furthermore

$$\sum_{1 \leqslant i \leqslant |I_t|} (f - f_i) \leqslant \sum_{1 \leqslant i \leqslant |I_t|} \mathbb{1}[x_i \in U_t]\mathbb{1}[\forall\, j \neq i\colon x_i \neq x_j] \leqslant f.$$

Thus f has the self-bounding property, which establishes the claim in the case of *push*. The proof for *pull* is completely analogous, where we use

$$f^{(pull)}\colon [n]^{|U_t|} \to \mathbb{R}, \ \ (x_1, \ldots, x_{|U_t|}) \mapsto |I_t| + \sum_{k \in U_t} \mathbb{1}[x_k \in I_t]$$

and, similarly, for *push&pull* we use $f^{(pp)}\colon [n]^n \to \mathbb{R}$ with

$$\begin{aligned}(x_1, \ldots, x_n) \mapsto |I_t| &+ \sum_{1 \leqslant k \leqslant n} \mathbb{1}[k \in I_t]\mathbb{1}[x_k \in U_t]\mathbb{1}[\forall\, j \in \{1, \ldots, k\} \cap I_t\colon x_k \neq x_j] \\ &+ \sum_{1 \leqslant k \leqslant n} \mathbb{1}[k \in U_t]\mathbb{1}[x_k \in I_t]\mathbb{1}[\forall\, w \in I_t\colon x_w \neq k].\end{aligned}$$

Here it is useful to see that the two sums in $f^{(pp)}$ are complementary, that is, only one of the summands for index k can be 1. Thus the functions $f_i^{(pull)}$ and $f_i^{(pp)}$ are obtained analogously to the push case. □

Remark 2.1. Let $G = (V, E)$ be a graph. Lemma 2.3 also applies to subsets of I_{t+1}, that is, for any $U \subset V$ and conditioned on I_t we have that $|I_{t+1} \cap U|$ and $|(I_{t+1} \cap U) \setminus I_t|$ are self-bounding.

The following proposition gives a tool that we will use in order to extend our round-wise analysis to longer phases.

Proposition 2.4. *Let $(\mathcal{A}_i)_{i \in \mathbb{N}_0}$ be a sequence of events, $0 < c < 1$, $\delta > 0$ and $t_1 \geqslant t_0 \geqslant 1$, such that*

$$\mathbb{P}[\mathcal{A}_t \mid \mathcal{A}_{t_0}, \ldots, \mathcal{A}_{t-1}, \mathcal{A}_0] \geqslant 1 - c^{t-t_0}\delta \quad \textit{for all } t_0 \leqslant t \leqslant t_1.$$

Then

$$\mathbb{P}\left[\bigcap_{t=t_0}^{t_1} \mathcal{A}_t \mid \mathcal{A}_0\right] \geqslant 1 - \delta/(1 - c).$$

Proof. Using the definition of conditional probability we obtain, as $c < 1$,

$$\begin{aligned}\mathbb{P}\left[\bigcap_{t=t_0}^{t_1} \mathcal{A}_t \mid \mathcal{A}_0\right] &= \prod_{t=t_0}^{t_1} \mathbb{P}[\mathcal{A}_t \mid \mathcal{A}_{t_0}, \ldots, \mathcal{A}_{t-1}, \mathcal{A}_0] \\ &\geqslant \prod_{t=t_0}^{t_1} (1 - c^{t-t_0}\delta) \\ &\geqslant 1 - \sum_{t=t_0}^{t_1} (c^{t-t_0}\delta)\end{aligned}$$

$$= 1 - \delta \sum_{t=0}^{t_1 - t_0} c^t$$
$$\geqslant 1 - \delta/(1-c). \qquad \square$$

We give two typical applications of the previous lemmas, similar to what we will encounter several times later in the paper. The first lemma addresses the case where we have a lower bound for the expected number of informed vertices after one round.

Lemma 2.5. *Let $\mathcal{P} \in \{push, pull, pp\}$ and $I_t = I_t^{(\mathcal{P})}$. Assume that there is $c > 1$ such that $\mathbb{E}_t[|I_{t+1}|] \geqslant c|I_t|$ for all t as long as $n/f(n) \leqslant |I_t| \leqslant n/g(n)$ for some functions $1 \leqslant f(n) \leqslant g(n) \leqslant n, f = o(n)$. Assuming $|I_{t_0}| \geqslant n/f(n)$, then there is $\tau = \log_c(f(n)/g(n)) + o(\log n)$ such that w.h.p.*

$$|I_{t_0+\tau}| \geqslant n/g(n).$$

Proof. Let $t \geqslant t_0$ and $n/f(n) \leqslant |I_t| \leqslant n/g(n)$. Lemma 2.1 guarantees that $\mathrm{Var}_t[|I_{t+1}|] \leqslant \mathbb{E}_t[|I_{t+1}|]$, and applying Chebyshev's inequality gives

$$\mathbb{P}_t[|\,|I_{t+1}| - \mathbb{E}_t[|I_{t+1}|]| \leqslant \mathbb{E}_t[|I_{t+1}|]^{2/3}] \geqslant 1 - \mathbb{E}_t[|I_{t+1}|]^{-1/3} \geqslant 1 - |I_t|^{-1/3}. \tag{2.1}$$

Consider the events

$$\mathcal{A}_t = |I_t| \geqslant \mathbb{E}_{t-1}[|I_t|] - \mathbb{E}_{t-1}[|I_t|]^{2/3} \quad \text{or} \quad |I_t| \geqslant n/g(n).$$

The intersection of $\mathcal{A}_{t_0+1}, \ldots, \mathcal{A}_t$ implies inductively that either $|I_t| \geqslant n/g(n)$ or

$$\begin{aligned} |I_t| &\geqslant (1 - \mathbb{E}_{t-1}[|I_t|]^{-1/3})\mathbb{E}_{t-1}[|I_t|] \\ &\geqslant (1 - (c|I_{t-1}|)^{-1/3})c|I_{t-1}| \\ &\geqslant ((1 - (c|I_{t_0}|)^{-1/3})c)^{t-t_0}|I_{t_0}|. \end{aligned} \tag{2.2}$$

We obtain with (2.1)

$$\mathbb{P}_{t_0}[\mathcal{A}_{t+1} \mid \mathcal{A}_{t_0+1}, \ldots, \mathcal{A}_t, |I_t| < n/g(n)] \geqslant 1 - ((1 - (c|I_{t_0}|)^{-1/3})c)^{-(t-t_0)/3}|I_{t_0}|^{-1/3},$$

and otherwise

$$\mathbb{P}_{t_0}[\mathcal{A}_{t+1} \mid \mathcal{A}_{t_0+1}, \ldots, \mathcal{A}_t, |I_t| \geqslant n/g(n)] = 1.$$

Choose

$$\tau := t - t_0 = \log_c(f(n)/g(n)) + o(\log n)$$

as small as possible such that the lower bound for $|I_{t+1}|$ in (2.2) is $\geqslant n/g(n)$, that is, the lower bound in (2.2) is $< n/g(n)$ for $t = t_0 + \tau$. Combining the two conditional probabilities we obtain for all $t_0 \leqslant t \leqslant t_0 + \tau$

$$\mathbb{P}_{t_0}[\mathcal{A}_{t+1} \mid \mathcal{A}_{t_0+1}, \ldots, \mathcal{A}_t] \geqslant 1 - ((1-(c|I_{t_0}|)^{-1/3})c)^{-(t-t_0)/3}|I_{t_0}|^{-1/3}.$$

Applying Proposition 2.4 then yields the claim. □

In the second lemma we make the stronger assumption that we can determine asymptotically the expected number of informed vertices after one round. Here we assume that we begin with a 'small' set of informed vertices, say of size $\sqrt{\log n}$, and want to reach a set of size nearly linear in n.

Lemma 2.6. *Assume that there is some $c > 1$ such that $\mathbb{E}_t[|I_{t+1}|] = (1+o(1))c|I_t|$ for all t as long as $\sqrt{\log n} \leqslant |I_t| \leqslant n/\log n$. Assume furthermore that $|I_{t_0}| \geqslant \sqrt{\log n}$. Then there are $\tau_1, \tau_2 = \log_c(n/|I_{t_0}|) + o(\log n)$ such that w.h.p.*

$$|I_{t_0+\tau_1}| \leqslant \frac{n}{\log n} \leqslant |I_{t_0+\tau_2}|.$$

Proof. Lemma 2.5, setting $f = n/\sqrt{\log n}$ and $g = \log n$ directly implies the existence of τ_1. To find τ_2, let $\mathcal{A}_t$ be the event $||I_t| - \mathbb{E}_{t-1}[|I_t|]| \leqslant \mathbb{E}_{t-1}[|I_t|]^{2/3}$. There is $h(n) \in o(1)$ such that, for $c^- := (1-h(n))c$ and $c^+ := (1+h(n))c$, we have that $\mathbb{E}_t[|I_{t+1}|] \leqslant c^+|I_t|$ and $\mathbb{E}_t[|I_{t+1}|] \geqslant c^-|I_t|$. Using this notation, the events $\mathcal{A}_{t_0+1}, \ldots, \mathcal{A}_{t+1}$ together imply inductively that

$$\begin{aligned}|I_{t+1}| &\leqslant (1+\mathbb{E}_t[|I_{t+1}|]^{-1/3})\mathbb{E}_t[|I_{t+1}|]\\ &\leqslant (1+(c^-|I_t|)^{-1/3})c^+|I_t|\\ &\leqslant ((1+(c^-|I_{t_0}|)^{-1/3})c^+)^{t-t_0}|I_{t_0}|\end{aligned}$$

for all t such that the right-hand side is bounded by $n/\log n$. Moreover, for all such t,

$$\begin{aligned}|I_{t+1}| &\geqslant (1-\mathbb{E}_t[|I_{t+1}|]^{-1/3})\mathbb{E}_t[|I_{t+1}|]\\ &\geqslant (1-(c^-|I_t|)^{-1/3})c^-|I_t|\\ &\geqslant ((1-(c^-|I_{t_0}|)^{-1/3})c^-)^{t-t_0}|I_{t_0}|.\end{aligned}$$

Thus, as $\mathcal{A}_t$ only depends on I_t, it follows with (2.1) that

$$\mathbb{P}_{t_0}[\mathcal{A}_{t+1} \mid \mathcal{A}_{t_0+1}, \ldots, \mathcal{A}_t] \geqslant 1 - ((1-(c^-|I_{t_0}|)^{-1/3})c^-)^{-(t-t_0)/3}|I_{t_0}|^{-1/3}.$$

Applying Proposition 2.4 yields the existence of τ_2. □

Expander sequences. In this section we collect some important properties of expander sequences that we are going to use later. We start by stating a version of the well-known Expander Mixing Lemma applied to our setting of expander sequences.

Lemma 2.7 ([28, Corollary 2.4]). *Let $\mathcal{G} = (G_n)_{n\in\mathbb{N}} = ((V_n, E_n))_{n\in\mathbb{N}}$ be an expander sequence. Then, for $S_n \subseteq V_n$ such that $1 \leqslant |S_n| \leqslant n/2$, it is*

$$\left| e(S_n, V_n \backslash S_n) - \frac{\Delta_n |S_n|(n-|S_n|)}{n} \right| = o(\Delta_n)|S_n|.$$

The following result is a consequence of the Expander Mixing Lemma that applies to graphs in which some edges were removed. It seems very simple but it turns out to be surprisingly useful.

Lemma 2.8. *Let $\mathcal{G} = (G_n)_{n\in\mathbb{N}} = ((V_n, E_n))_{n\in\mathbb{N}}$ be an expander sequence. Let $\varepsilon > 0$ and set $\tilde{\mathcal{G}} = (\tilde{G}_n)_{n\in\mathbb{N}}$, where each $\tilde{G}_n$ is obtained from G_n by deleting edges such that each vertex keeps at least* *a* $(1/2+\varepsilon)$ *fraction of its edges. For each* $n \in \mathbb{N}$ *let* $S_n \subseteq V_n$*. Then there is* $n_0 \in \mathbb{N}$ *such that, for all* $n \geqslant n_0$,

$$e_{\tilde{G}_n}(S_n, V_n \backslash S_n) \geqslant \varepsilon e_{G_n}(S_n, V_n \backslash S_n).$$

Proof. Without loss of generality we assume that $|S_n| \leqslant n/2$. Since at most $(1/2-\varepsilon)\Delta_n$ edges are deleted at each vertex, we immediately obtain that

$$e_{\tilde{G}_n}(S_n, V_n \backslash S_n) \geqslant e_{G_n}(S_n, V_n \backslash S_n) - \Delta_n(1/2-\varepsilon)|S_n|.$$

Using Lemma 2.7 and choosing n_0 large enough such that

$$\frac{o(\Delta_n)}{\Delta_n}\frac{n}{n-|S_n|} < \varepsilon \quad \text{for all } n \geqslant n_0,$$

we obtain that

$$\begin{aligned}
&(1-\varepsilon)e_{G_n}(S_n, V_N \backslash S_n) - \Delta_n(1/2-\varepsilon)|S_n| \\
&\geqslant (1-\varepsilon)\frac{\Delta_n|S_n|(n-|S_n|)}{n} - o(\Delta_n)|S_n| - \Delta_n(1/2-\varepsilon)|S_n| \\
&= \frac{\Delta_n|S_n|(n-|S_n|)}{n}\left(1-\varepsilon-\frac{o(\Delta_n)}{\Delta_n}\frac{n}{n-|S_n|}-\frac{n(1/2-\varepsilon)}{n-|S_n|}\right).
\end{aligned}$$

As $n-|S_n| \geqslant n/2$, the last expression is > 0. Hence

$$\begin{aligned}
e_{\tilde{G}_n}(S_n, V_n \backslash S_n) &\geqslant \varepsilon e_G(S_n, V_n \backslash S_n) + (1-\varepsilon)e_G(S_n, V_n \backslash S_n) - \Delta_n(1/2-\varepsilon)|S_n| \\
&\geqslant \varepsilon e_{G_n}(S_n, V_n \backslash S_n).
\end{aligned}$$
□

Next we give a lemma that counts the number of paths between two arbitrary vertices of a dense graph satisfying a weak expander property (as for example guaranteed by Lemma 2.8). This will later be used to give a lower bound on the probability of any vertex being informed after a given constant number of rounds.

Lemma 2.9. *Let* $G=(V,E)$, $|V|=n$*. Assume that there is* $\alpha > 0$ *such that* $d(v) \geqslant \alpha n$ *for all* $v \in V$ *and* $e(W, V\backslash W) \geqslant \alpha|W|\,|V\backslash W|$ *for all* $W \subseteq V$*. Then, for all* $u, w \in V$*, there is* $1 \leqslant d \leqslant 8/\alpha^2+2$ *such that there are at least* $(\alpha^4/64)^{d+1}n^{d-1}$ *paths of length* d *from* u *to* w.

Proof. Assume $\alpha \leqslant 1/2$, as otherwise the claim is trivial (with $d \in \{1,2\}$). We define sequences $(S_i)_{i\in\mathbb{N}}$ and $(H_i)_{i\in\mathbb{N}} \subseteq V$ as follows. Set $S_1 = \{u\} \cup N(u)$, $W = \{w\} \cup N(w)$ and $H_1 = V \backslash (S_1 \cup W)$ and proceed for $i \geqslant 1$ as follows. Let $\tilde{S}_{i+1} \subseteq H_i$ be the set of vertices $v \in H_i$ with $|N(v) \cap S_i| \geqslant \alpha^2 n/8$. Set $S_{i+1} = S_i \cup \tilde{S}_{i+1}$ and $H_{i+1} = H_i \backslash \tilde{S}_{i+1}$. Then we claim that, for all $i \geqslant 1$,

$$e(S_i, W) \geqslant \alpha^3 n^2/2 \quad \text{or} \quad |S_{i+1}| \geqslant |S_i| + \alpha^2 n/8. \tag{2.3}$$

To see this, assume that $e(S_i, W) \leqslant \alpha^3 n^2/2$. Since $|S_i|, |W| \geqslant \alpha n$, the weak expansion property guarantees that

$$e(S_i, H_i) = e(S_i, H_i \cup W) - e(S_i, W) \geqslant \alpha|S_i|\,|H_i \cup W| - \alpha^3 n^2/2 \geqslant \alpha^2(1-\alpha)n^2 - \alpha^3 n^2/2,$$

and using $\alpha \leqslant 1/2$ we obtain that $e(S_i, H_i) \geqslant \alpha^2 n^2/4$. To complete the proof of (2.3) we compute the size of $\tilde{S}_{i+1}$. As $|N(v) \cap S_i| \leqslant \alpha^2 n/8$ for all $v \in H_i \backslash \tilde{S}_{i+1}$ and $|N(v) \cap S_i| \leqslant n$, we get

$$\frac{\alpha^2 n^2}{4} \leqslant e(S_i, H_i) \leqslant |\tilde{S}_{i+1}|n + |H_i|\frac{\alpha^2 n}{8}.$$

Since $|H_i| \leqslant n$ we immediately get that $|\tilde{S}_{i+1}| \geqslant \alpha^2 n/8$, which shows (2.3). We next show that there are (sufficiently) many paths for each vertex in S_i to u. More precisely, let $1 \leqslant j \leqslant 8/\alpha^2$ be such that $e(S_i, W) < \alpha^3 n^2/2$ for all $1 \leqslant i \leqslant j$. For those i we have by (2.3) that $|S_i| \geqslant i \cdot \alpha^2 n/8$. We claim that for all $v \in S_i \setminus \{u\}$ there is $d \leqslant i$ such that v has at least $(\alpha^4/64)^d \cdot n^{d-1}$ paths of length d with endpoint u. We show the claim by induction on i. The base case $v \in S_1 \setminus \{u\}$ is clear, as $1 \geqslant \alpha^4/64$. For the induction step assume that $v \in S_{i+1} \setminus S_i$, $v \neq u$. Then by construction $|N(v) \cap S_i| \geqslant \alpha^2 n/8$. Thus, by induction hypothesis, there is $d \leqslant i$ such that v has at least $\alpha^2 n/(8i)$ neighbours with at least $(\alpha^4/64)^d n^{d-1}$ paths with endpoint u. As $i \leqslant 8/\alpha^2$ this gives that v has at least $\alpha^2 n/(8i) \cdot (\alpha^4/64)^d n^{d-1} \geqslant (\alpha^4/64)^{d+1} n^d$ paths of length $d+1 \leqslant i+1$ with endpoint u, and this accomplishes the induction step. With all these facts at hand we finally show the claim of the lemma. Let $j \leqslant 8/\alpha^2$ be the first index such that $e(S_j, W) \geqslant \alpha^3 n^2/2$, and let $W' \subseteq W$ be such that $|N(v) \cap S_j| \geqslant \alpha^3 n/4$ for all $v \in W'$. Thus

$$\frac{\alpha^3 n^2}{2} \leqslant e(S_j, W) \leqslant |W'|n + |W|\frac{\alpha^3 n}{4},$$

and thus $|W'| \geqslant \alpha^3 n/4$. Then there is $d \leqslant j$ and $W'' \subseteq W'$ such that $|W''| \geqslant |W'|/j$ and every v in W'' has at least $\alpha^3 n/(4j)$ neighbours with at least $(\alpha^4/64)^d n^{d-1}$ paths of length d with endpoint u. Therefore every $v \in W''$ has at least

$$(\alpha^4/64)^d n^{d-1} \cdot \alpha^3 n/(4j) \geqslant (\alpha^4/64)^{d+1} n^d/j$$

paths of length $d+1$ with endpoint u. This in turn gives that there are at least

$$|W'|/j \cdot (\alpha^4/64)^{d+1} n^d/j \geqslant \alpha^3/4 \cdot (\alpha^4/64)^{d+2} n^{d+1}$$

paths of length $d+2$ from w to u, and the proof is completed. □

Next comes a technical lemma that, given a small set, quantifies the number of vertices for which only a small fraction of their neighbourhood intersects that given set.

Lemma 2.10. *Let $\mathcal{G} = (G_n)_{n\in\mathbb{N}} = ((V_n, E_n))_{n\in\mathbb{N}}$ be an expander sequence. Let $\varepsilon > 0$ and let $\tilde{\mathcal{G}} = (\tilde{G}_n)_{n\in\mathbb{N}}$, where each $\tilde{G}_n$ it is obtained from G_n by deleting edges such that each vertex keeps at least a $(1/2+\varepsilon)$ fraction of its edges. Let $A_n \subseteq V_n$ with $|A_n| = o(n)$.*

(a) *There is $B_n \subseteq A_n$ with $|B_n| = (1-o(1))|A_n|$ such that, for all $u \in B_n$,*

$$\frac{|N_{\tilde{G}_n}(u) \cap A_n|}{|N_{\tilde{G}_n}(u)|} = o(1).$$

(b) *There is $B_n \subseteq V_n \setminus A_n$ with $|V_n \setminus (A_n \cup B_n)| = o(|A_n|)$ such that, for all $v \in B_n$,*

$$\frac{|N_{\tilde{G}_n}(v) \cap A_n|}{|N_{\tilde{G}_n}(v)|} = o(1).$$

Proof. Let δ_n, Δ_n denote the minimum and maximum degree of G_n. Lemma 2.7 yields that

$$e_{G_n}(A_n, V_n \setminus A_n) = \frac{\Delta_n |A_n|\,|V_n \setminus A_n|}{n} + o(\Delta_n)|A_n| = (1+o(1))\Delta_n|A_n|.$$

As there are a maximum of $\Delta_n|A_n|$ edges with at least one point in A_n, we get that $e_{G_n}(A_n) = o(\Delta_n)|A_n|$. Since we obtain $\tilde{G}_n$ from G_n by deleting edges,

$$e_{\tilde{G}_n}(A_n) = o(\Delta_n)|A_n|. \tag{2.4}$$

With this fact at hand we show (a). Let $\eta > 0$ and call a vertex $u \in A_n$ bad if $|N_{\tilde{G}_n}(u) \cap A_n| \geqslant \eta|N_{\tilde{G}_n}(u)|$. Since $N_{\tilde{G}_n}(u) \geqslant \delta_n/2$, we obtain for any bad u that $|N_{\tilde{G}_n}(u) \cap A_n| \geqslant \eta\delta_n/2$. As $\delta_n = (1 - o(1))\Delta_n$, we infer from (2.4) that the number of bad vertices is $o(|A_n|)$.

To see (b), again let $\eta > 0$ and this time call a vertex $v \in V_n \setminus A_n$ bad if $|N_{\tilde{G}_n}(v) \cap A_n| \geqslant \eta|N_{\tilde{G}_n}(v)|$. Then, for any such bad v, we know that $|N_{\tilde{G}_n}(v) \cap A_n| \geqslant \eta\delta_n/2$. As before, using (2.4) we readily get that the number of bad v is $o(|A_n|)$. □

We conclude our preliminary section by giving a lemma that crudely bounds the time needed until at least $\omega(1)$ vertices are informed.

Lemma 2.11. *Let $0 < \varepsilon \leqslant 1/2, q \in (0, 1]$ and $\mathcal{G} = (G_n)_{n\in\mathbb{N}}$ be an expander sequence. Let $\tilde{\mathcal{G}} = (\tilde{G}_n)_{n\in\mathbb{N}}$ be such that each $\tilde{G}_n$ is obtained by deleting edges of G_n such that each vertex keeps at least a $(1/2 + \varepsilon)$ fraction of its edges. Let $\mathcal{P} \in \{push, pull, pp\}$ and suppose that $|I_t^{(\mathcal{P})}| < \sqrt{\log n}$. Then there is $\tau = o(\log n)$ such that w.h.p. $|I_{t+\tau}^{(\mathcal{P})}| \geqslant \sqrt{\log n}$.*

Proof. Recall that the probability that $v \in U_t$ gets informed by *pull* is $q|N(v) \cap I_t|/|N(v)|$. Thus

$$\mathbb{P}_t[|I_{t+1}^{(pull)} \backslash I_t| = 0] = \prod_{u\in N(I_t)\cap U_t} \left(1 - \frac{q|N(u) \cap I_t)|}{|N(u)|}\right) \leqslant e^{-qe(U_t,I_t)/\Delta_n}.$$

Similarly we obtain for *push*

$$\mathbb{P}_t[|I_{t+1}^{(push)} \backslash I_t| = 0] = \prod_{v\in I_t} \frac{|N(v) \cap I_t|}{|N(v)|} = \prod_{v\in I_t}\left(1 - \frac{|N(v) \cap U_t|}{|N(v)|}\right) \leqslant e^{-qe(I_t,U_t)/\Delta_n}.$$

The same bound is obviously also true for *push&pull*. Thus, for all $\mathcal{P} \in \{push, pull, pp\}$,

$$\mathbb{P}_t[|I_{t+1}^{(\mathcal{P})} \backslash I_t| \geqslant 1] \geqslant 1 - e^{-qe(U_t,I_t)/\Delta_n}.$$

As Lemmas 2.7 and 2.8 imply that $e(U_t, I_t) \geqslant (1 + o(1))\varepsilon\Delta_n|I_t|$, there is $c \in (0, 1)$ such that $\mathbb{P}[|I_{t+1}^{(\mathcal{P})} \backslash I_t| \geqslant 1] > c$. Define $\tau := \lceil(2/c)\sqrt{\log n}\rceil$ and $X = \mathrm{Bin}(\tau, c)$ with $\mathbb{E}[X] = c\tau$ and $\mathrm{Var}[X] = \tau(1 - c)c$. Then, using Chebyshev,

$$\mathbb{P}_t[|I_{t+\tau}^{(\mathcal{P})}| \leqslant \sqrt{\log n}] \leqslant \mathbb{P}_t[X \leqslant \sqrt{\log n}] \leqslant \mathbb{P}_t[|X - \mathbb{E}[X]| \leqslant \mathbb{E}[X]/2] \leqslant 4\mathrm{Var}[X]/\mathbb{E}[X]^2 = o(1). \qquad \square$$

3. Proofs

3.1 Proof of Theorems 1.2(b) and 1.3(a) – edge deletions do not slow down pull

Let $0 < \varepsilon \leqslant 1/2$. In this section we study the runtime of *pull* in the case in which the input graph is an expander, and where at each vertex at most a $(1/2 - \varepsilon)$ fraction of the edges is deleted. The runtime on expander sequences without edge deletions, *i.e.* the setting in Theorem 1.2(b), is included as the special case where we set $\varepsilon = 1/2$. In contrast to previous proofs, in the analysis of *pull* the 'standard' approach that consists of showing, for example, that $\mathbb{E}_t[|I_{t+1} \setminus I_t|] \approx |I_t|$ fails. The main reason is that the graph between I_t and U_t might be quite irregular, so that, depending on the actual state, $\mathbb{E}_t[|I_{t+1} \setminus I_t|] \approx c|I_t|$ for some $c < 1$. However, we discover a different invariant that is preserved, namely that the number of edges between I_t and U_t behaves in an exponential way. With Lemmas 2.7 and 2.8 we can then relate this to the number of informed vertices.

Lemma 3.1. *Consider the setting of Theorem 1.3(a) and let $I_t = I_t^{(pull)}$.*

(a) *Let $\sqrt{\log n} \leqslant |I_t| \leqslant n/\log n$. Then*

$$|e(U_{t+1}, I_{t+1}) - (1+q)e(U_t, I_t)| \leqslant |I_t|^{-1/3} e(U_t, I_t)$$

with probability at least $1 - O(|I_t|^{-1/3})$.

(b) *Let $|U_t| \leqslant n/\log n$. Then $\mathbb{E}_t[|U_{t+1}|] = (1 - q + o(1))|U_t|$.*

Proof. We start with (a). Let $D_t = e(U_{t+1}, I_{t+1}) - e(U_t, I_t)$ and for $u \in U_t$ let X_u be the random variable that indicates whether u gets informed in round $t+1$. Then

$$\mathbb{E}_t[D_t] = \sum_{u \in U_t} \sum_{v \in N(u) \cap U_t} \mathbb{E}_t[X_u(1 - X_v)] - \sum_{u \in U_t} \mathbb{E}_t[X_u] \cdot |N(u) \cap I_t|$$

$$= \sum_{u \in U_t} q \frac{|N(u) \cap I_t|}{|N(u)|} \left(\left(\sum_{v \in N(u) \cap U_t} 1 - q \frac{|N(v) \cap I_t|}{|N(v)|} \right) - |N(u) \cap I_t| \right).$$

The second sum is at most $|N(u)|$, so obviously $\mathbb{E}_t[D_t] \leqslant qe(U_t, I_t)$. To get a lower bound consider a largest set $\tilde{U} \subseteq U_t$ such that $|N(u) \cap I_t|/|N(u)| = o(1)$ for all $u \in \tilde{U}$. From Lemma 2.10(b) we obtain that $|U_t \setminus \tilde{U}| = o(|I_t|)$, and so

$$\mathbb{E}_t[D_t] \geqslant \sum_{u \in U_t} q|N(u) \cap I_t| \left(\left(\sum_{v \in N(u) \cap \tilde{U}} \frac{1}{|N(u)|} - o\left(\frac{1}{|N(u)|} \right) \right) - \frac{|N(u) \cap I_t|}{|N(u)|} \right).$$

Consider furthermore $\hat{U} \subseteq \tilde{U}$ such that $|N(u) \cap \tilde{U}|/|N(u)| = 1 - o(1)$ and thus also $|N(u) \cap I_t|/|N(u)| = o(1)$ for all $u \in \hat{U}$. Lemma 2.10(b) again yields that we can choose $\hat{U}$ such that $|U_t \setminus \hat{U}| = o(|I_t|)$, and thus

$$\mathbb{E}_t[D_t] \geqslant (1 - o(1)) \sum_{u \in \hat{U}} q|N(u) \cap I_t| \left(\frac{|N(u) \cap \tilde{U}|}{|N(u)|} - \frac{|N(u) \cap I_t|}{|N(u)|} \right) - \sum_{u \in U_t \setminus \hat{U}} |N(u) \cap I_t|$$

$$\geqslant (q - o(1))e(U_t, I_t) - 2e(U_t \setminus \hat{U}, I_t).$$

According to Lemmas 2.7 and 2.8 we have that $e(U_t, I_t) = \Theta(|I_t|\Delta_n)$. But

$$e(U_t \setminus \tilde{U}, I_t) \leqslant |U_t \setminus \tilde{U}|\Delta_n = o(|I_t|\Delta_n).$$

Thus $\mathbb{E}_t[e(U_{t+1}, I_{t+1})] = (1 + q - o(1))e(U_t, I_t)$. In the next step we bound the variance. For each edge e let X_e be the indicator random variable that denotes the events that $e \in E(U_{t+1}, I_{t+1})$. Thus

$$e(U_{t+1}, I_{t+1}) = \sum_{e \in E} X_e = \frac{1}{2} \sum_{u \in V} \sum_{v \in N(u)} X_{\{u,v\}}.$$

Using the fact that X_e and $X_{e'}$ are independent for all $e, e' \in E$ with $e \cap e' = \emptyset$,

$$\mathrm{Var}[e(U_{t+1}, I_{t+1})] = \mathrm{Var}\left[\sum_{e \in E} X_e \right]$$

$$= \sum_{e, e' \in E} \mathbb{E}[X_e X_{e'}] - \mathbb{E}[X_e]\mathbb{E}[X_{e'}]$$

$$\leqslant \sum_{u \in V} \sum_{v, v' \in N(u)} \mathbb{E}[X_{\{u,v\}} X_{\{u,v'\}}]$$

$$\leqslant \Delta_n \sum_{u\in V} \sum_{v\in N(u)} \mathbb{E}[X_{\{u,v\}}]$$
$$= 2\Delta_n \mathbb{E}[e(U_{t+1}, I_{t+1})].$$

Since

$$\mathbb{E}_t[e(U_{t+1}, I_{t+1})] = (1 + q - o(1))e(U_t, I_t) = \Theta(\Delta_n|I_t|),$$

by Lemmas 2.7 and 2.8 and

$$\mathrm{Var}[e(U_{t+1}, I_{t+1})] \leqslant 2\Delta_n \mathbb{E}_t[e(U_{t+1}, I_{t+1})]$$

we immediately obtain for $|I_t| \geqslant \sqrt{\log n}$ with Chebyshev's inequality that

$$\mathbb{P}[|e(U_{t+1}, I_{t+1}) - \mathbb{E}_t[e(U_{t+1}, I_{t+1})]| \geqslant e(U_t, I_t)|I_t|^{-1/3}] \leqslant O(|I_t|^{-1/3}).$$

Next we show (b). We bound the expected number of uninformed vertices after one additional round. Lemma 2.10(a) asserts that there is a set $\tilde{U} \subseteq U_t$ such that $|\tilde{U}| = (1 - o(1))|U_t|$ and $|N(u) \cap I_t|/|N(u)| = 1 - o(1)$ for all $u \in \tilde{U}$. Thus

$$\begin{aligned}\mathbb{E}_t[|U_{t+1}|] &= \sum_{u\in U_t} 1 - q\frac{|N(u)\cap I_t|}{|N(u)|}\\ &\leqslant |U_t| - q\sum_{u\in\tilde{U}} \frac{|N(u)\cap I_t|}{|N(u)|}\\ &= |U_t| - q(1 - o(1))|\tilde{U}|\\ &= (1 - q - o(1))|U_t|.\end{aligned}$$

As $|N(u) \cap I_t| \leqslant |N(u)|$ we also have

$$\mathbb{E}_t[|U_{t+1}|] = \sum_{u\in U_t} 1 - q\frac{|N(u)\cap I_t|}{|N(u)|} \geqslant \sum_{u\in U_t} (1 - q) = (1 - q)|U_t|. \qquad \square$$

Lemmas 3.2 and 2.11 give lower bounds which, together with an upper bound provided by Lemma 3.3, imply Theorems 1.2(b) and 1.3(a).

Lemma 3.2. (upper bound in Theorem 1.3(a))**.** *Consider the setting of Theorem 1.3(a) and let $I_t = I_t^{(pull)}$. Then the following statements hold w.h.p.*

(a) *Let $\sqrt{\log n} \leqslant |I_t| \leqslant n/\log n$. Then there are $\tau_1, \tau_2 = \log_{1+q}(n/|I_t|) + o(\log n)$ such that $|I_{t+\tau_2}| < n/\log n < |I_{t+\tau_1}|$.*
(b) *Let $n/\log n \leqslant |I_t| \leqslant n - n/\log n$. Then there is $\tau = o(\log n)$ such that $|I_{t+\tau}| > n - n/\log n$.*
(c) *Let $|I_t| \geqslant n - n/\log n$.*
  (i) *Case $q = 1$. Then there is $\tau = o(\log n)$ such that $|I_{t+\tau}| = n$.*
  (ii) *Case $q \neq 1$. Then there is $\tau \leqslant -\log n/\log(1-q) + o(\log n)$ such that $|I_{t+\tau}| = n$.*

Proof. We start with (a). Let $|I_t| \in [\sqrt{\log n}, n/\log n]$. First note that any bound on $e(U_t, I_t)$ translates to a bound for $|I_t|$, as with Lemmas 2.7 and 2.8 we obtain

$$(1 - o(1))\varepsilon\Delta_n|I_t| \leqslant e(U_t, I_t) \leqslant \Delta_n|I_t|. \tag{3.1}$$

In particular, up to constant factors, $|I_t|$ is $e(U_t, I_t)/\Delta_n$ and *vice versa*. From Lemma 3.1(a) we obtain that $e(U_{t+1}, I_{t+1}) = (1 + q \pm |I_t|^{-1/3})e(U_t, I_t)$ with probability $1 - O(|I_t|^{-1/3})$. Proceeding as in Lemmas 2.5 and 2.6 and their proofs, where we replace the events

$$|I_t| \geqslant \mathbb{E}_{t-1}[|I_t|] - \mathbb{E}_{t-1}[|I_t|]^{2/3} \quad \text{or} \quad |I_t| \geqslant n/g(n)$$

and

$$||I_t| - \mathbb{E}_{t-1}[|I_t|]| \leqslant \mathbb{E}_{t-1}[|I_t|]^{2/3}$$

with

$$e(U_t, I_t) \geqslant (1 + q - |I_{t-1}|^{-1/3})e(U_{t-1}, I_{t-1}) \quad \text{or} \quad |I_t| \geqslant n/\log n$$

and

$$e(U_{t+1}, I_{t+1}) = (1 + q \pm |I_t|^{-1/3})e(U_t, I_t),$$

we obtain the statement.

We continue with (b). Consider first the case $|I_t| \in [n/\log n, n/2]$. Using Lemmas 2.7 and 2.8, *i.e.* $e(U_t, I_t) \geqslant \varepsilon|U_t|\,|I_t|\Delta_n/n + o(\Delta_n)|I_t|$, together with $|U_t| \geqslant n/2$ implies

$$\begin{aligned}
\mathbb{E}_t[|I_{t+1} \backslash I_t|] &= \sum_{u \in U_t} q \frac{|N(u) \cap I_t|}{|N(u)|} \\
&\geqslant \frac{q \cdot e(U_t, I_t)}{\Delta_n} \\
&\geqslant \frac{q\varepsilon|U_t|\,|I_t|\Delta_n/n + o(\Delta_n)|I_t|}{\Delta_n(1 + o(1))} \\
&\geqslant \left(\frac{q\varepsilon}{2} + o(1)\right)|I_t|.
\end{aligned}$$

Applying Lemma 2.5, where we set $g = 2, f = \log n$ and $c = q\varepsilon/2 + o(1)$, we are finished with this part as well. Now let $|I_t| \in [n/2, n - n/\log n]$. We switch our focus to the set of uninformed vertices. Using again the fact that $e(U_t, I_t) \geqslant \varepsilon|U_t|\,|I_t|\Delta_n/n + o(\Delta_n)|U_t|$, we have

$$\begin{aligned}
\mathbb{E}_t[|U_{t+1}|] &= \sum_{u \in U_t} 1 - q\frac{|N(u) \cap I_t|}{|N(u)|} \\
&= \sum_{u \in U_t} 1 - q\frac{|N(u) \cap I_t|}{\Delta_n(1 + o(1))} \\
&= |U_t| - \frac{q \cdot e(U_t, I_t)}{\Delta(1 + o(1))} \\
&= |U_t| - \frac{q\varepsilon|U_t|\,|I_t|\Delta_n/n + o(\Delta_n)|U_t|}{\Delta_n(1 + o(1))} \\
&\leqslant \left(1 - \frac{q\varepsilon}{2} + o(1)\right)|U_t|.
\end{aligned}$$

Inductively we obtain for any integer $\tau \geqslant 1$ the bound $\mathbb{E}_t[|U_{t+\tau}|] \leqslant (1 - q\varepsilon/2 + o(1))^\tau|U_t|$, and so for some $\tau := 2\log\log n/\log(1/(1 - q\varepsilon/2 + o(1))) = o(\log n)$ we have

$$\mathbb{E}_t[|U_{t+\tau}|] \leqslant |U_t|/\log^2 n = o(n/\log n).$$

Hence, by Markov's inequality, $\mathbb{P}_t[|U_{t+\tau}| \geqslant n/\log n] = o(1)$.

In order to show (c), let $|I_t| \in [n - n/\log n, n]$. As for $q = 1$ the term $1 - q$ in Lemma 3.1(b) vanishes, we distinguish the cases $q = 1$ and $q \neq 1$. We start with $q = 1$. By induction, it follows that for any round $\tau > 0$ and suitable $f = o(1)$,

$$\mathbb{E}_t[|U_{t+\tau}|] \leqslant (f(n))^\tau|U_t|.$$

We choose $\tau = \log_{1/f(n)}(n) = o(\log n)$ as $1/f = \omega(1)$. Hence we obtain $\mathbb{E}_t[|U_{t+\tau}|] \leqslant |U_t|/n \leqslant 1/\log n$. Therefore we have $\mathbb{P}_t[|U_{t+\tau}| \geqslant 1] \leqslant o(1)$ by Markov's inequality. For $q \neq 1$ we have by induction, for any number of rounds $\tau \geqslant 1$,

$$\mathbb{E}_t[|U_{t+\tau}|] \leqslant (1-q+o(1))^{\tau}|U_t|.$$

We choose

$$\tau = \log_{1/(1-q+o(1))}(n) = -\log n/\log(1-q) + o(\log n).$$

Thus, using Markov's inequality, analogously to the case $q = 1$, we obtain the desired upper bound. □

Note that for $q = 1$ this already implies Theorems 1.2(b) and 1.3(a). This leaves the case for $q \neq 1$.

Lemma 3.3. *Let $0 < \varepsilon \leqslant 1/2, q \in (0, 1]$ and $\mathcal{G} = (G_n)_{n\in\mathbb{N}}$ be an expander sequence. Let $\tilde{\mathcal{G}} = (\tilde{G}_n)_{n\in\mathbb{N}}$ be such that each $\tilde{G}_n$ is obtained by deleting edges of G_n, such that each vertex keeps at least a $(1/2+\varepsilon)$ fraction of its edges and abbreviate $I_t = I_t^{(pull)}$. Let $q \in (0, 1)$ and $|I_t| \leqslant n/2$. Then, for $\tau = -\log n/\log(1-q)$ and all $c < 1$, w.h.p. $|I_{t+c\tau}| < n$.*

Proof. We consider a modified process in which vertices have a higher chance of getting informed. In particular, note that the probability that $u \in U_t$ gets informed is at most $q|N(u) \cap I_t|/|N(u)| \leqslant q$ and that all these events are independent; now we assume that each such u gets independently informed with probability exactly q. Then the runtime in this modified model constitutes a lower bound for the runtime in the original model.

Let $c < 1, u \in U_t$ and E_u be the event that u does not get informed in $c\tau$ rounds in this model. Thus

$$\mathbb{P}[E_u] = (1-q)^{c\tau} = (1-q)^{-c\log n/\log(1-q)} = n^{-c} = \omega(1/n),$$

and as the events E_u are independent and $|U_t| = \Theta(n)$,

$$\mathbb{P}\Bigg[\bigwedge_{u\in U_t} \overline{E_u}\Bigg] \leqslant \prod_{u\in U_t} \mathbb{P}[\overline{E_u}] \leqslant \exp\Bigg(-\sum_{u\in U_t} \mathbb{P}[E_u]\Bigg) = o(1). \qquad \square$$

3.2 *Proof of Theorem 1.2(c) –* **push&pull** *is fast on expanders*

As we are now in the case without edge deletions, we begin with a lemma that determines the expected number of informed vertices in one round. Intuitively we will show that *push* and *pull* do not interact badly, and therefore *push&pull* is given as a straightforward combination of *push* and *pull*.

Lemma 3.4. *Let $\mathcal{G}$ be an expander sequence and abbreviate $I_t = I_t^{(pp)}$.*

(a) *Let $|I_t| \leqslant n/\log n$. Then $\mathbb{E}_t[|I_{t+1} \setminus I_t|] = (2q + o(1))|I_t|$.*
(b) *Let $|U_t| \leqslant n/\log n$. Then $\mathbb{E}_t[|U_{t+1}|] = (1 + o(1))e^{-q}(1-q)|U_t|$.*

Proof. We begin with (a). The probability that $v \in U_t$ gets informed by *pull* is $q|N(v) \cap I_t|/|N(v)|$. Thus, using Lemma 2.7,

$$\begin{aligned}
\mathbb{E}_t[|I_{t+1}^{(pull)} \backslash I_t|] &= \sum_{u \in U_t} q \frac{|N(u) \cap I_t|}{|N(u)|} \\
&= q \sum_{u \in U_t} \frac{|N(u) \cap I_t|}{\Delta_n(1+o(1))} \\
&= (q+o(1)) \frac{e(U_t, I_t)}{\Delta_n} \\
&= (q+o(1)) \frac{|U_t|\,|I_t|\Delta_n/n + o(\Delta_n)|I_t|}{\Delta_n}. \qquad (3.2)
\end{aligned}$$

Since $|I_t| = o(n)$ we obtain that $|U_t| = (1-o(1))n$, and this expression simplifies to $(q+o(1))|I_t|$.

Before we switch our attention to *push* we make a simple observation. Let $a_1, \ldots, a_k, k \in \mathbb{N}$ be real numbers. Then, using the fact that for any $a = o(1)$ it is $e^{-a+o(a)} = 1-a$ and $e^{-a} = 1-a+o(a)$, we have

$$\prod_{1 \leqslant i \leqslant k} (1-a_i) = \exp\left(-(1+o(1)) \sum_{1 \leqslant i \leqslant k} a_i\right) = 1-(1+o(1)) \sum_{1 \leqslant i \leqslant k} a_i \quad \text{if} \sum_{1 \leqslant i \leqslant k} a_i = o(1). \qquad (3.3)$$

The probability that $v \in U_t$ gets informed by *push* is

$$1 - \prod_{i \in N(v) \cap I_t} (1 - q/|N(v)|).$$

According to Lemma 2.10(b) there is $B_t \subseteq U_t$ such that $|N(u) \cap I_t| = o(|N(u)|)$ for all $u \in B_t$ and $|U_t \setminus B_t| = o(|I_t|)$. Thus (3.3) is applicable, and in a similar fashion to (3.2) we get

$$\begin{aligned}
\mathbb{E}_t[|I_{t+1}^{(push)} \setminus I_t|] &= \sum_{u \in U_t} 1 - \prod_{i \in N(u) \cap I_t} \left(1 - \frac{q}{|N(i)|}\right) \\
&= q \sum_{u \in B_t} \frac{|N(u) \cap I_t|}{\Delta_n(1+o(1))} + o(|I_t|) \\
&= (q+o(1))|I_t|. \qquad (3.4)
\end{aligned}$$

We express the expected number of vertices informed by *push&pull* after one additional round in terms of the expected values we just calculated ((3.2) and (3.4)):

$$\begin{aligned}
\mathbb{E}_t[|I_{t+1} \backslash I_t|] &= \mathbb{E}_t\big[|I_{t+1}^{(pull)} \backslash I_t| + |I_{t+1}^{(push)} \backslash I_t| - |(I_{t+1}^{(push)} \backslash I_t) \cap (I_{t+1}^{(pull)} \backslash I_t)|\big] \\
&= (2q - o(1))|I_t| - \mathbb{E}_t\big[|(I_{t+1}^{(push)} \backslash I_t) \cap (I_{t+1}^{(pull)} \backslash I_t)|\big]. \qquad (3.5)
\end{aligned}$$

Lemma 2.10(a) gives a set

$$A_t \subseteq I_{t+1}^{(push)}, |A_t| = (1-o(1))|I_{t+1}^{(push)}|,$$

such that

$$|N(u) \cap I_{t+1}^{(push)}| = o(1)|N(u)| \quad \text{for all } u \in A_t.$$

Since push and pull happen independently,

$$\begin{aligned}\mathbb{E}_t\left[|(I_{t+1}^{(pull)}\backslash I_t)\cap(I_{t+1}^{(push)}\backslash I_t)|\mid I_{t+1}^{(push)}\right] &= \sum_{u\in I_{t+1}^{(push)}\backslash I_t}\mathbb{P}_t[u\in I_{t+1}^{(pull)}\backslash I_t]\\ &= \sum_{u\in I_{t+1}^{(push)}\backslash I_t} q\frac{|N(u)\cap I_t|}{|N(u)|}\\ &\leqslant \sum_{u\in A_t} q\frac{|N(u)\cap I_t|}{|N(u)|}+\sum_{u\in I_{t+1}^{(push)}\backslash A_t} q\frac{|N(u)\cap I_t|}{|N(u)|}.\end{aligned}$$

Using the fact that $|N(u)\cap I_t| = o(|N(u)|)$ for all $u\in A_t$, we obtain

$$\mathbb{E}_t\left[|(I_{t+1}^{(pull)}\backslash I_t)\cap(I_{t+1}^{(push)}\backslash I_t)|\right]\leqslant \mathbb{E}_t\left[o(|A_t|)+|I_{t+1}^{(push)}\setminus A_t|\right]=o(|I_t|),$$

as

$$|A_t|\leqslant |I_{t+1}^{(push)}|\leqslant 2|I_t| \quad\text{and}\quad |I_{t+1}^{(push)}\setminus A_t| = o(|I_{t+1}^{(push)}|)=o(|I_t|).$$

Combining this with (3.5) we get $\mathbb{E}_t[|I_{t+1}\setminus I_t|]=(2q+o(1))|I_t|$, as claimed.

Next we show (b). Let A_u be the event that an uninformed vertex u does not get informed by the *push* algorithm, let B_u be the corresponding event for *pull*. Then A_u and B_u are independent and $A_u\cap B_u$ is the event that u does not get informed in the current round. We obtain

$$\begin{aligned}\mathbb{P}_t[A_u] &= \prod_{i\in N(u)\cap I_t}\left(1-\frac{q}{|N(i)|}\right)\\ &\leqslant \left(1-\frac{q}{\Delta_n}\right)^{|N(u)\cap I_t|}\\ &\leqslant \exp\left(-q\frac{|N(u)\cap I_t|}{\Delta_n}\right)\\ &= \exp\left(\frac{-q|N(u)\cap I_t|}{(1+o(1))|N(u)|}\right)\end{aligned}$$

and

$$\mathbb{P}_t[B_u]=1-\frac{q|N(u)\cap I_t|}{|N(u)|}.$$

According to Lemma 2.10(a) there is a set $C_t\subseteq U_t$, $|C_t|=(1-o(1))|U_t|$ such that $|N(u)\cap I_t| = (1-o(1))|N(u)|$ for all $u\in C_t$. As $\mathbb{P}_t[A_u\cap B_u]\leqslant 1$, we therefore get

$$\mathbb{E}_t[|U_{t+1}|]=\sum_{u\in U_t}\mathbb{P}_t[A_u\cap B_u]\leqslant\sum_{u\in C_t}\mathbb{P}_t[A_u]\cdot\mathbb{P}_t[B_u]+|U_t\setminus C_t|\leqslant(1+o(1))e^{-q}(1-q)|U_t|.$$

For the lower bound we need to find a lower bound on the probability of a single uninformed vertex not getting informed in one round by *push*. Indeed, for any $u\in U_t$ and sufficiently large n,

$$\mathbb{P}_t[A_u]=\prod_{v\in N(u)\cap I_t}\left(1-\frac{q}{|N(v)|}\right)\geqslant\left(1-\frac{q}{\delta_n}\right)^{|N(u)\cap I_t|}\geqslant e^{-q\Delta_n/\delta_n}. \tag{3.6}$$

Combining this inequality with the trivial bound $\mathbb{P}[B_u] \geqslant 1-q$, we get a lower bound on the expected number of uninformed vertices after one round using *push&pull*:

$$\begin{aligned}\mathbb{E}_t[|U_{t+1}|] &= \sum_{u\in U_t} \mathbb{P}_t[A_u \cap B_u] \\ &= \sum_{u\in U_t} \mathbb{P}_t[A_u]\cdot \mathbb{P}_t[B_u] \\ &\geqslant e^{-q\Delta_n/\delta_n}(1-q)|U_t| \\ &= (1+o(1))e^{-q}(1-q)|U_t|. \qquad \square\end{aligned}$$

Next we show upper and lower bounds that together with Lemma 2.11 imply Theorem 1.2(c).

Lemma 3.5. *Let $\mathcal{G}$ be an expander sequence and abbreviate $I_t = I_t^{(pp)}$. Let $q \in (0, 1]$. Then the following statements hold w.h.p.*

(a) *Let $\sqrt{\log n} \leqslant |I_t| \leqslant n/\log n$. Then there are $\tau_1, \tau_2 = \log_{1+2q}(n/|I_t|) + o(\log n)$ such that $|I_{t+\tau_2}| < n/\log n < |I_{t+\tau_1}|$.*
(b) *Let $n/\log n \leqslant |I_t| \leqslant n - n/\log n$. Then there is $\tau = o(\log n)$ such that $|I_{t+\tau}| > n - n/\log n$.*
(c) *Let $|I_t| \geqslant n - n/\log n$.*
  (i) *Case $q = 1$. Then there is $\tau = o(\log n)$ such that $|I_{t+\tau}| = n$.*
  (ii) *Case $q \neq 1$. Then there is $\tau \leqslant \log n/(q - \log(1-q)) + o(\log n)$ such that $|I_{t+\tau}| = n$.*

Proof. Since $|I_t| \geqslant |I_t^{(pull)}|$, statements (b) and (c) for $q = 1$ follow immediately from Lemma 3.2. To see (a), note that by using Lemma 3.4 we get $\mathbb{E}_t[|I_{t+1}\backslash I_t|] = (2q + o(1))|I_t|$, and applying Lemma 2.6 implies the claim.

Finally we show (c) for $q \neq 1$. Let $|I_t| \geqslant n - n/\log n$. By Lemma 3.4, we obtain that, for any $\tau \in \mathbb{N}$,

$$\mathbb{E}_t[|U_{t+\tau}|] = ((1+o(1))e^{-q}(1-q))^{\tau}|U_t|.$$

Thus we may choose $\tau = \log n/(q - \log(1-q)) + o(\log n)$ such that, say, $\mathbb{E}_t[|U_{t+\tau}|] \leqslant |U_t|/n \leqslant 1/\log n$. Thus $\mathbb{P}_t[|U_{t+\tau}| \geqslant 1] \leqslant o(1)$ by Markov's inequality. $\square$

Note that for $q = 1$ this already implies Theorem 1.2(c). This leaves the case for $q \neq 1$.

Lemma 3.6. *Let $\mathcal{G}$ be an expander sequence and abbreviate $I_t = I_t^{(pp)}$, let $q \in (0, 1)$ and $|I_t| \leqslant n/2$. Then for $\tau = \log n/(q - \log(1-q))$ and all $c < 1$ w.h.p. $|I_{t+c\tau}| < n$.*

Proof. We consider a modified process in which vertices have a higher chance of getting informed. In particular, note that the probability that $u \in U_t$ gets informed by *pull* is at most $q|N(u) \cap I_t|/|N(u)| \leqslant q$ and that all these events are independent; according to (3.6) the probability that $u \in U_t$ gets informed by *push* is at most $1 - e^{-q\Delta_n/\delta_n}$. Now we assume that each such u gets independently informed with probability exactly $1 - e^{-q\Delta_n/\delta_n}(1-q)$. Then the runtime in this modified model constitutes a lower bound for the runtime in the original model. Let $u \in U_t$ and E_u be the event that u does not get informed in this modified model in $c\tau$ rounds. Thus, for $c < 1$,

$$\mathbb{P}[E_u] \geqslant ((1-q)e^{-q\Delta_n/\delta_n})^{c\tau} = \omega(n^{-1}),$$

and as the events E_u are independent and $|U_t| = \Theta(n)$,

$$\mathbb{P}\Big[\bigwedge_{u\in U_t} \overline{E_u}\Big] \leqslant \prod_{u\in U_t} \mathbb{P}[\overline{E_u}] \leqslant \exp\Big(-\sum_{u\in U_t} \mathbb{P}[E_u]\Big) = o(1). \qquad \square$$

3.3 Proof of Theorem 1.5(a) – push *informs almost all vertices fast in spite of edge deletions*

To shorten the notation, let us call the setting with deleted edges the 'new model' and the setting without deleted edges the 'old model', that is, the term 'new model' corresponds to the graphs in $\tilde{\mathcal{G}}$ while 'old model' refers to the (original) graphs in $\mathcal{G}$. We prove Lemma 3.7, which directly implies Theorem 1.5(a). We write $I_t = I_t^{(push)}$ throughout.

Lemma 3.7. *Under the assumptions of Theorem 1.5(a), the following holds for the new model:*

(a) *There are $\tau, \tilde{\tau} = \log_{1+q}(n) + o(\log n)$ such that w.h.p. $|I_{\tilde{\tau}}| < n/\log n < |I_\tau|$.*
(b) *Assume $|I_t| \geqslant n/\log n$. Then there is a $\tau = o(\log n)$ such that w.h.p. $|I_{t+\tau}| \geqslant n - n/\log n$.*

For the proof of Lemma 3.7 we will need the following statements, the first one taken from [28].

Lemma 3.8 (proof of Lemma 2.5 in [28])**.** *Consider the old model. Assume $|I_t| < n/\log n$ and $q = 1$. Then*

$$\mathbb{P}_t[|I_{t+1}| = |I_t| + (1 - o(1))|I_t|] = 1 - o(1). \tag{3.7}$$

Lemma 3.9. *Consider* push *on a sequence of graphs $(G_n)_{n\in\mathbb{N}}$, where G_n has n vertices. Assume that $|I_t| = \omega(1)$ and that (3.7) holds for $q = 1$, that is, assume that*

$$\mathbb{P}_t[|I_{t+1}| = |I_t| + (1 - o(1))|I_t|] = 1 - o(1) \quad \textit{for } q = 1.$$

Then, for $q \in (0, 1]$,

$$\mathbb{P}_t[|I_{t+1}| = |I_t| + (q - o(1))|I_t|] = 1 - o(1). \tag{3.8}$$

Moreover, assume that whenever $|I_t| < n/\log n$, for $q = 1$, (3.7) holds. Then there are $\tau, \tilde{\tau} = \log_{1+q}(n) + o(\log n)$ such that w.h.p.

$$|I_{\tilde{\tau}}| < n/\log n < |I_\tau|. \tag{3.9}$$

Proof. For a graph G and for $v \in I_t$, let $X_v(G)$ denote the vertex to which v pushes in round t. Let

$$N_{t+1} := \{X_v(G_n) \mid v \in I_t\} \cap U_t.$$

Note that whenever $|I_t| < n/\log n$, w.h.p. $|N_{t+1}| = (1 - o(1))|I_t|$ from (3.7). For $q \in (0, 1]$ each vertex in N_{t+1} has a probability at least q of being informed and all these events are independent; thus (3.8) follows directly by applying the Chernoff bounds whenever $|I_t| = \omega(1)$.

In order to prove the second statement we call a round t that does not satisfy (3.8) a *failed* round. Note that we just argued that the probability that a round fails is $o(1)$ whenever $|I_t| = \omega(1)$ and $|I_t| < n/\log n$, and the events that distinct rounds fail are independent. In particular, the number of failed rounds among the next R rounds, assuming that $|I_t|$ stays below $n/\log n$, is w.h.p. $o(R)$. Moreover, if a round does not fail, the number of informed vertices increases by a factor of $(1 + q + o(1))$ and otherwise it may increase by an arbitrary factor in the interval $[1, 2]$.

Finally, Lemma 2.11 yields that there is $t^* = o(\log n)$ such that w.h.p. $|I_{t^*}| = \omega(1)$, which implies that after $R + t^*$ rounds, the number of informed vertices is w.h.p. in the interval

$$[(1+q+o(1))^{R-o(R)}, (1+q+o(1))^{R-o(R)} \cdot 2^{o(R)}],$$

and choosing $R = \log_{1+q}(n) + o(\log n)$ in two ways establishes (3.9). □

In the subsequent proof of Lemma 3.7 we will use the simple observations that, for any $n \in \mathbb{N}_0$,

$$\mathbb{P}[\mathrm{Bin}(n, 1/2) \geqslant n/2] \geqslant 1/2 \quad \text{and} \quad \mathbb{P}[\mathrm{Bin}(n, 1/4) \geqslant n/4] \geqslant 1/4 \tag{3.10}$$

(see *e.g.* [22] when $n > 4$), and the other cases are checked easily.

Proof of Lemma 3.7. We first show (a). We assume $q = 1$ and prove that, for $|I_t| < n/\log n$, (3.7) also holds in the new model; then claim (a) follows directly from Lemma 3.9. Let $G = (V, E)$ be a graph. For $v \in I_t$ let $X_v(G)$ denote the vertex to which v pushes in round t. For $u \in V$ let $c_u(G) := |\{v \in I_t \mid X_v(G) = u\}|$ denote the number of times u is pushed in round t. Let

$$\mathcal{Y}_t(G) := \{v \in I_t \mid c_v(G) = 1\} \quad \text{and} \quad \mathcal{H}_t(G) := \{v \in I_t \mid c_v(G) \geqslant 1\}$$

denote the set of informed vertices that are being pushed exactly once in round t and the set of informed vertices that are being pushed at least once in round t respectively. Let

$$\mathcal{Z}_t(G) := \{v \in V \mid c_v(G) \geqslant 2\}$$

denote the set of vertices that are being pushed more than once in round t. Let $Y_t(G) := |\mathcal{Y}_t(G)|$ and $H_t(G) := |\mathcal{H}_t(G)|$ and, in a slight abuse of notation, let

$$Z_t(G) := \sum_{k \geqslant 2} (k-1) \cdot |\{v \in V \mid c_v(G) = k\}|$$

denote the number of vertices that are being pushed multiple times in round t counted with multiplicity. Note that the quantity $Y + Z$ denotes the number of pushes that have no effect in the respective round, that is, there are $Y + Z$ pushes that are useless in the sense that even without them, the same number of vertices would become informed in the respective round. In the following paragraphs we condition on I_t implicitly, that is, we write $\mathbb{P}[\ldots]$ instead of $\mathbb{P}_t[\ldots]$, *etc.*, to lighten the notation. We want to show that (3.7) does hold in the new model; for contradiction we assume that this is not the case. Hence we can infer that there is a constant $c > 0$ such that

$$\limsup_{n\to\infty} \mathbb{P}[Y_t(\tilde{G}_n) \geqslant c|I_t|] > 0 \quad \text{or} \quad \limsup_{n\to\infty} \mathbb{P}[Z_t(\tilde{G}_n) \geqslant c|I_t|] > 0.$$

Thus, without loss of generality, we can assume that there is $f^* > 0$ and $n_0 \in \mathbb{N}$ such that

$$\mathbb{P}[Y_t(\tilde{G}_n) \geqslant c|I_t|] > f^* \text{ for all } n \geqslant n_0 \quad \text{or} \quad \mathbb{P}[Z_t(\tilde{G}_n) \geqslant c|I_t|] > f^* \text{ for all } n \geqslant n_0.$$

If this is not the case we can restrict ourselves to a suitable subsequence of $(n)_{n\in\mathbb{N}}$ on which it is true. Next, we describe an explicit coupling between the new and the old model. For any vertex v consider $X_v(G_n)$. If $X_v(G_n) \in N_{\tilde{G}_n}(v)$, then set $X_v(\tilde{G}_n) := X_v(G_n)$ and otherwise choose $X_v(\tilde{G}_n)$ uniformly at random from $N_{\tilde{G}_n}(v)$. Note that $X_v(G_n), X_v(\tilde{G}_n)$ have by construction the correct marginal distribution. Moreover, note that by construction, the family

$$(X_v(G_n) \mid (X_u(\tilde{G}_n))_{u\in V_n})_{v\in V_n} \tag{3.11}$$

of random variables is independent, since $X_v(G_n)$ depends only on $X_v(\tilde{G}_n)$ for all $v \in V_n$.

We begin with the case that $\mathbb{P}[Y_t(\tilde{G}_n) \geqslant c|I_t|] > f^*$. We will show

$$\mathbb{P}[H_t(G_n) \geqslant Y_t(\tilde{G}_n)/2 \mid \mathcal{Y}_t(\tilde{G}_n)] \geqslant 1/2$$

and then, since by assumption $\mathbb{P}[Y_t(\tilde{G}_n) \geqslant c|I_t|] > f^*$, we can infer $\mathbb{P}[H_t(G_n) \geqslant c|I_t|/2] \geqslant f^*/2$, which contradicts Lemma 3.8. Let $\mathcal{Y}_t(\tilde{G}_n) = \{y_1, \ldots, y_{Y_t(\tilde{G}_n)}\}$. Then there are distinct vertices $v_1, \ldots, v_{Y_t(\tilde{G}_n)} \in I_t$ such that $X_{v_i}(\tilde{G}_n) = y_i$ for all $i \in \{1, \ldots, Y_t(\tilde{G}_n)\}$. Due to (3.11) the events $(\{X_{v_i}(G_n) = X_{v_i}(\tilde{G}_n)\})_{1 \leqslant i \leqslant Y_t}$ are independent. Moreover, for all $i \in \{1, \ldots, Y_t(\tilde{G}_n)\}$,

$$\mathbb{P}[X_{v_i}(G_n) = X_{v_i}(\tilde{G}_n) \mid \mathcal{Y}_t(\tilde{G}_n)] = \frac{d_{\tilde{G}_n}(v_i)}{d_{G_n}(v_i)} \geqslant 1/2 + \varepsilon$$

and therefore, given $\mathcal{Y}_t(\tilde{G}_n)$, $H_t(G_n)$ dominates a binomially distributed random variable $\mathrm{Bin}(Y_t(\tilde{G}_n), 1/2)$. In particular, this implies with (3.10) that

$$\mathbb{P}[H_t(G_n) \geqslant Y_t(\tilde{G}_n)/2 \mid \mathcal{Y}_t(\tilde{G}_n)] \geqslant 1/2,$$

as claimed.

We continue with the case $\mathbb{P}[Z_t(\tilde{G}_n) \geqslant c|I_t|] > f^*$. Let $\mathcal{Z}_t(\tilde{G}_n) = \{z_1, \ldots, z_{|\mathcal{Z}_t(\tilde{G}_n)|}\}$. Then, for any $i \in \{1, \ldots, |\mathcal{Z}_t(\tilde{G}_n)|\}$ let $n_i := c_{z_i}(\tilde{G}_n) \geqslant 2$, that is, there are distinct vertices $v_{i,1}, \ldots, v_{i,n_i}$ such that $X_v(\tilde{G}_n) = z_i$ for all $v \in \{v_{i,1}, \ldots, v_{i,n_i}\}$. We will show that

$$\mathbb{P}[Z_t(G_n) \geqslant Z_t(\tilde{G}_n)/8 \mid \mathcal{Z}_t(\tilde{G}_n), n_1, \ldots, n_{|\mathcal{Z}_t(\tilde{G}_n)|}] \geqslant 1/8 \tag{3.12}$$

and then, since by assumption $\mathbb{P}[Z_t(\tilde{G}_n) \geqslant c|I_t|] > f^*$, we obtain $\mathbb{P}[Z_t(G_n) \geqslant c/8|I_t|] \geqslant f^*/8$, which contradicts Lemma 3.8. Due to (3.11) the events

$$(\{X_{v_{i,j}}(G_n) = X_{v_{i,j}}(\tilde{G}_n)\})_{1 \leqslant i \leqslant |\mathcal{Z}_t(\tilde{G}_n)|, 1 \leqslant j \leqslant n_i} \tag{3.13}$$

are independent. Moreover, for all $1 \leqslant i \leqslant |\mathcal{Z}_t(\tilde{G}_n)|$, $1 \leqslant j \leqslant n_i$,

$$\mathbb{P}\big[X_{v_{i,j}}(G_n) = X_{v_{i,j}}(\tilde{G}_n) \mid \mathcal{Z}_t(\tilde{G}_n), n_1, \ldots, n_{|\mathcal{Z}_t(\tilde{G}_n)|}\big] = \frac{d_{\tilde{G}_n}(v_{i,j})}{d_{G_n}(v_{i,j})} \geqslant 1/2 + \varepsilon. \tag{3.14}$$

For $1 \leqslant i \leqslant |\mathcal{Z}_t(\tilde{G}_n)|$ let $B_i \sim \mathrm{Bin}(n_i, 1/2)$ be independent random variables. Moreover, let $M_1 := \{i \mid 1 \leqslant i \leqslant |\mathcal{Z}_t(\tilde{G}_n)|, n_i = 2\}$ and $M_2 := \{i \mid 1 \leqslant i \leqslant |\mathcal{Z}_t(\tilde{G}_n)|, n_i > 2\}$. Using (3.13) and (3.14), given $\mathcal{Z}_t(\tilde{G}_n), n_1, \ldots, n_{|\mathcal{Z}_t(\tilde{G}_n)|}$, we infer that $Z_t(G_n)$ dominates

$$\sum_{i=1}^{|\mathcal{Z}_t(\tilde{G}_n)|} \max\{B_i - 1, 0\} \geqslant \sum_{i \in M_1} \max\{B_i - 1, 0\} + \sum_{i \in M_2} B_i - |M_2|.$$

We treat the two sums individually. Note that $\sum_{i \in M_1} \max\{B_i - 1, 0\} \sim \mathrm{Bin}(|M_1|, 1/4)$; in particular,

$$\mathbb{P}\Bigg[\sum_{i \in M_1} \max\{B_i - 1, 0\} \geqslant |M_1|/4\Bigg] \geqslant 1/4$$

by (3.10). Regarding the second sum, since $\sum_{i \in M_2} B_i \sim \mathrm{Bin}(\sum_{i \in M_2} n_i, 1/2)$, we obtain

$$\mathbb{P}\Bigg[\sum_{i \in M_2} B_i \geqslant 1/2 \sum_{i \in M_2} n_i\Bigg] \geqslant 1/2.$$

Thus, given $\mathcal{Z}_t(\tilde{G}_n), n_1, \ldots, n_{|\mathcal{Z}_t(\tilde{G}_n)|}$ and using $2|M_1| = \sum_{i\in M_1} n_i$ and $\sum_{i\in M_2} n_i \geqslant 3|M_2|$, we infer that with probability at least $1/4 \cdot 1/2 = 1/8$

$$\begin{aligned}
Z_t(G_n) &\geqslant \frac{1}{4}|M_1| + \frac{1}{2}\sum_{i\in M_2} n_i - |M_2| \\
&= \frac{1}{8}\sum_{i\in M_1} n_i + \frac{1}{2}\sum_{i\in M_2} n_i - |M_2| \\
&\geqslant \frac{1}{8}\sum_{i\in M_1} n_i + \frac{1}{6}\sum_{i\in M_2} n_i \\
&\geqslant \frac{1}{8}\sum_{i=1}^{|\mathcal{Z}_t(\tilde{G}_n)|} n_i \\
&= \frac{1}{8}(Z_t(\tilde{G}_n) + |\mathcal{Z}_t(\tilde{G}_n)|) \\
&\geqslant \frac{1}{8}Z_t(\tilde{G}_n).
\end{aligned}$$

This establishes (3.12). All in all, for $q = 1$ we have shown that (3.7) also holds in the new model. Hence claim (a) follows directly from Lemma 3.9.

Next we prove claim (b). We write $\Delta_n := \Delta(G_n)$, $\tilde{\Delta}_n := \Delta(\tilde{G}_n)$, $\delta_n := \delta(G_n)$ and $\tilde{\delta}_n := \delta(\tilde{G}_n)$; moreover, we write $\tilde{N}(\cdot)$ instead of $N_{\tilde{G}_n}(\cdot)$. We assume that $|I_t| \in [n/\log n, n - n/\log n]$. We further distinguish two cases, namely $|I_t| \in [n/\log n, n/2]$ and $|I_t| \in [n/2, n - n/\log n]$. We start with the case $|I_t| \in [n/\log n, n/2]$. Using Lemmas 2.7 and 2.8 and the assumption that $\Delta_n/\delta_n = 1 + o(1)$ we obtain, for any $0 < \bar{\varepsilon} < \varepsilon/2$, for n sufficiently large,

$$e(I_t, U_t) > \bar{\varepsilon}\delta_n|I_t|. \tag{3.15}$$

Using the fact that $e^x \geqslant (1 + x/n)^n$ for $n \in \mathbb{N}$ and $|x| \leqslant n$, we obtain

$$\mathbb{E}_t[|I_{t+1}\backslash I_t|] \geqslant \sum_{u\in\tilde{N}(I_t)\backslash I_t}\left[1 - \prod_{v\in\tilde{N}(u)\cap I_t}\left(1 - \frac{q}{\tilde{\Delta}_n}\right)\right] \geqslant \sum_{u\in\tilde{N}(I_t)\backslash I_t} 1 - e^{-|\tilde{N}(u)\cap I_t|q/\tilde{\Delta}_n}.$$

Further, using the fact that $e^{-x} \leqslant 1 - x/2$ for any $x \in (0, 1)$ and (3.15) yields the bound

$$\mathbb{E}_t[|I_{t+1}\backslash I_t|] \geqslant \sum_{u\in\tilde{N}(I_t)\backslash I_t} \frac{q|\tilde{N}(u)\cap I_t|}{2\tilde{\Delta}_n} = \frac{qe(I_t, U_t)}{2\tilde{\Delta}_n} \geqslant \frac{\bar{\varepsilon}q\delta_n}{2\Delta_n}|I_t|.$$

For this case the claim follows by Lemma 2.5, when setting $f = n/\log n, g = \log n$ and $c = \bar{\varepsilon}q\delta_n/(2\Delta_n)$.

Finally we consider the case $|I_t| \in [n/2, n - n/\log n]$; here we examine the shrinking of U_t. Using Lemmas 2.7 and 2.8 we obtain, for any $0 < \bar{\varepsilon} < \varepsilon/2$, for n sufficiently large, $e(I_t, U_t) > \bar{\varepsilon}\delta_n|U_t|$. Hence, again using the fact that for any $x \in (0, 1)$ it holds that $e^{-x} \leqslant 1 - x/2$, and that for $n \in \mathbb{N}$ and $|x| \leqslant n$ it is $e^x \geqslant (1 + x/n)^n$, we obtain

$$\begin{aligned}
\mathbb{E}_t[|U_{t+1}|] &= \sum_{u\in U_t}\prod_{v\in\tilde{N}(u)\cap I_t}\left(1 - \frac{q}{d_{\tilde{G}_n}(v)}\right) \\
&\leqslant \sum_{u\in U_t} e^{-|\tilde{N}(u)\cap I_t|q/\tilde{\Delta}_n}
\end{aligned}$$

$$\leqslant \sum_{u\in U_t} 1 - \frac{q|\tilde{N}(u)\cap I_t|}{2\tilde{\Delta}_n}$$
$$\leqslant |U_t| - \frac{\bar{\varepsilon} q\delta_n}{2\tilde{\Delta}_n}|U_t|$$
$$\leqslant \left(1 - \frac{\bar{\varepsilon} q\delta_n}{2\Delta_n}\right)|U_t|.$$

Using the tower property of conditional expectation, we immediately get

$$\mathbb{E}_t[|U_{t+\tau}|] \leqslant \left(1 - \frac{\bar{\varepsilon} q\delta_n}{2\Delta_n}\right)^{\tau}|U_t|, \qquad \tau \in \mathbb{N}.$$

Thus, for

$$\tau := -2\log\log(n)/\log(1-\bar{\varepsilon} q\delta_n/(2\Delta_n)) = o(\log n),$$

we have $\mathbb{E}_t[|U_{t+\tau}|] = o(n/\log n)$. Hence by Markov's inequality, $\mathbb{P}[|U_{t+\tau}| \geqslant n/\log n] = o(1)$. □

3.4 *Proof of Theorem 1.4(a) – edge deletions slow down* push

Let $I_t^{(push)} := I_t$. In order to show the claim we construct an explicit sequence of graphs that has the desired property. More precisely, for any $\varepsilon > 0$, each $q \in (0, 1]$ and $n \in \mathbb{N}$ we will define a graph $G_n(\varepsilon)$ that is obtained by deleting edges from the complete graph on n vertices such that each vertex keeps at least a $(1-\varepsilon)$ fraction of its edges and such that *push* slows down significantly.

We define $G_n(\varepsilon) = (V_1 \cup V_2, E)$ with vertex set $V = V_1 \cup V_2$, where $V_1 := \{1, \ldots, \lfloor n/2\rfloor\}$ and $V_2 := \{\lfloor n/2\rfloor + 1, \ldots, n\}$, as follows. We include in E all pairs of vertices that intersect V_1 and, moreover, we add edges (that now have endpoints only in V_2) such that all vertices in V_2 have degree $\lceil(1-\varepsilon)n\rceil + 1 \pm 1$. According to Lemma 3.7(a) there is a $t = \log_{1+q}(n) + o(\log n)$ such that w.h.p. $|I_t| < n/\log n$. It thus suffices to show that it takes w.h.p. at least $(1+\varepsilon/2)q^{-1}\log n$ more rounds to inform all remaining vertices.

Let $U_t' := U_t^{(push)} \cap V_2$. As $|I_t| < n/\log n$ we have $|U_t'| \geqslant n/4$ with plenty of room to spare. In the remainder of this proof we will consider a modified process in which vertices have a higher chance of getting informed; in particular we assume that in each round, *all* vertices choose a neighbour independently and uniformly at random and after this round the chosen vertices are informed. Let E_u denote the event that $u \in U_t'$ does not get informed within the next $\tau := (1+\varepsilon/2)q^{-1}\log n$ rounds in this modified model. Each vertex $u \in U_t'$ has $\lfloor n/2\rfloor$ neighbours that have degree $n-1$, at most $\lceil(1-\varepsilon)n\rceil + 1 \pm 1 - \lfloor n/2\rfloor \leqslant (1/2-\varepsilon)n + 4$ neighbours that have at least degree $(1-\varepsilon)n$ and no further neighbours. Therefore, using the fact that for any $a \in \mathbb{R}$ we have $(1+a/n)^n = e^a + O(1/n)$, we obtain for each $u \in U_t'$

$$\mathbb{P}_t[E_u] \geqslant \left(\left(1-\frac{q}{n-1}\right)^{n/2}\left(1-\frac{q}{(1-\varepsilon)n}\right)^{(1/2-\varepsilon)n+4}\right)^{\tau}$$
$$= (1+o(1))(e^{-q(1/2+(1/2-\varepsilon)/(1-\varepsilon))})^{\tau}$$
$$= (1+o(1))\exp\left(-\frac{4-4\varepsilon-3\varepsilon^2}{4-4\varepsilon}\log n\right)$$
$$= \omega(n^{-1}).$$

In this modified model the events $\{\overline{E_u} \mid u \in U_t'\}$ also satisfy $\mathbb{P}_t[\overline{E_u} \mid \{\overline{E_v}\colon v \in U\}] \leqslant 1-p$ for all $u \in V_2$ and $U \subseteq V \setminus \{u\}$ and for some $p = \omega(n^{-1})$. This follows immediately from the previous calculation, as conditioning on an event like $\{\overline{E_v}\colon v \in U\}$ only decreases the number of vertices that can push to u. Thus, as $|U'_t| = \Theta(n)$,

$$\mathbb{P}_t\left[\bigwedge_{u\in U'_t} \overline{E_u}\right] \leqslant \prod_{u\in U'_t} (1-p) \leqslant \exp\left(-\sum_{u\in U'_t} p\right) = o(1).$$

3.5 Proof of Theorems 1.3(b) and 1.5(a) – **push&pull** *informs almost all vertices fast in spite of edge deletions*

Before we show the actual proof we will first present an informal argument that contains all relevant ideas and important observations. Let $\sqrt{\log n} \leqslant |I_t| \leqslant n/\log n$ and assume $q = 1$. In Section 3.3 we proved that for *push* the informed vertices nearly double in every round for an arbitrary expander sequence with edge deletions and an otherwise arbitrary set I_t. For *pull* this is not true; however, we proved in Section 3.1 that the number of edges between the informed and the uninformed vertices nearly doubles in every round. The first attempt towards the proof of Theorems 1.3(b) and 1.5(b) then seems obvious: one would try to show that either the vertices triple every round, or the the edges do so, or for example that the product of the two quantities increases by a factor of 9. As it turns out, this is in general not the case; indeed, it is possible to choose an expander sequence, to delete edges such that each vertex keeps at least an $(1/2+\varepsilon)$-fraction of its neighbours, and to choose a (large) set of informed vertices I_t such that after one round w.h.p. either $|I_{t+1}| < c|I_t|$ or $e(I_{t+1}, U_{t+1}) < ce(I_t, U_t)$ or $|I_{t+1}|e(I_{t+1}, U_{t+1}) < c^2|I_t|e(I_t, U_t)$ for some $c < 3$. On the other hand, and although we have no explicit description of these 'malicious' sets, it seems rather unlikely that such sets will occur several times during the execution of *push&pull*.

In order to show the claimed running time of *push&pull* we will impose some additional structure. Let $\varepsilon > 0$. In the subsequent exposition we assume that our graph G – obtained from an expander by deleting edges such that each vertex keeps at least a $(1/2+\varepsilon)$ fraction of the edges – has a *very special* structure. In particular, we assume that there is a partition $\Pi = (V_i)_{i\in[k]}$ of the vertex set of G into a bounded number k of equal parts such that $E_G(V_i) = \emptyset$ for all $1 \leqslant i \leqslant k$ and such that the induced subgraph (V_i, V_j) looks like a random regular bipartite graph for all $1 \leqslant i < j \leqslant k$. Of course, not every relevant G admits such a partition; however, Szemerédi's Regularity Lemma guarantees that every sufficiently large graph has a partition that is in a well-defined sense *almost* like the one described previously, and a substantial part of our proof is concerned with showing that being 'almost special' does not hurt significantly.

Assuming that G is very special, let us collect some easy facts. Denote the degree of $u \in V_i$ in the induced subgraph (V_i, V_j) by d_{ij}; this immediately gives $d_G(u) = \sum_{\ell=1}^{k} d_{i\ell}$, and note that $d_{ii} = 0$ as there are no edges in V_i. Moreover, regular bipartite random graphs satisfy an expander property, that is,

$$\begin{aligned} e(W_i, W_j) &= d_{i,j}|W_i|\,|W_j|/|V_j| + o(d_{i,j})|W_i| \\ &\approx |W_i|\,|W_j| d_{ij} k/n \quad \text{for all } W_i \subseteq V_i,\, W_j \subseteq V_j,\, 1 \leqslant i < j \leqslant k, \end{aligned}$$

where we used the fact that all $|V_i|$ are of equal size. This is quite similar to the property that we used in our preceding analysis on expander sequences; see Lemma 2.7. As a pair in Π behaves like a bipartite expander sequence, we can easily compute the expected number of informed vertices like we did in Section 3.2. We do so now for *pull*. Let $|I_{t+1}^{i,j}|$ be the number of vertices in V_i informed after round $t+1$ by *pull* from vertices only in V_j and set $I_t^i := I_t \cap V_i$, $U_t^i := U_t \cap V_i$ for all $1 \leqslant i \leqslant k$. Thus, as long as I_t^i is much smaller than V_i (and thus also $U_t^i \approx |V_i| = n/k$), we get

$$\mathbb{E}_t[|I_{t+1}^{(pull),i,j}\setminus I_t|] = \sum_{u\in U_t^i} \frac{|N(u)\cap I_t^j|}{d(u)} = \frac{e(U_t^i, I_t^j)}{\sum_{1\leqslant \ell\leqslant k} d_{i\ell}} \approx \frac{d_{ij}}{\sum_{1\leqslant \ell\leqslant k} d_{i\ell}}|I_t^j|.$$

A similar calculation, which we do not perform in detail, yields for *push*

$$\mathbb{E}_t[|I_{t+1}^{(push),i,j}\setminus I_t|] \approx \frac{d_{ij}}{\sum_{1\leqslant \ell\leqslant k} d_{\ell j}}|I_t^j|.$$

Moreover, as in previous proofs it turns out that the number of vertices informed simultaneously by *push* as well as *pull* is negligible; compare with the proof of Lemma 3.4. Thus we obtain that more or less

$$\mathbb{E}_t[|I_{t+1}^{(pp),i,j}|] \approx |I_t^i| + \left(\frac{d_{ij}}{\sum_{1\leqslant \ell\leqslant k} d_{i\ell}} + \frac{d_{ij}}{\sum_{1\leqslant \ell\leqslant k} d_{\ell j}}\right)|I_t^j|,$$

and by linearity of expectation

$$\mathbb{E}_t[|I_{t+1}^{(pp),i}|] \approx |I_t^i| + \sum_{1\leqslant j\leqslant k}\left(\frac{d_{ij}}{\sum_{1\leqslant \ell\leqslant k} d_{i\ell}} + \frac{d_{ij}}{\sum_{1\leqslant \ell\leqslant k} d_{\ell j}}\right)|I_t^j|.$$

Set $X_t = (|I_t^i|)_{i\in[k]}$ and $A = (A_{ij})_{1\leqslant i,j\leqslant k}$, the matrix with entries

$$A_{ij} = \frac{d_{ij}}{\sum_{1\leqslant \ell\leqslant k} d_{i\ell}} + \frac{d_{ij}}{\sum_{1\leqslant \ell\leqslant k} d_{\ell j}} \quad \text{for } 1\leqslant i\neq j\leqslant k$$

and $A_{ii} = 1$ for $1\leqslant i\leqslant k$. With this notation we obtain the recursive relation

$$\mathbb{E}_t[X_{t+1}] \approx A\cdot X_t, \tag{3.16}$$

that is, we may expect that $X_t \approx \mathbb{E}_t[X_t] \approx A^t X_0$. If we then let $\lambda_{\max}$ denote the greatest eigenvalue of A, then we obtain that to leading order

$$|I_t| \approx \lambda_{\max}^t.$$

Our aim is to show that *push&pull* is (at least) as fast as on the complete graph, that is, $|I_t| \gtrsim 3^t$, so we take a closer look at the eigenvalues of A. By construction A is symmetric, so the largest eigenvalue equals $\sup_{\|x\|=1} \|x^T Ax\|$, and the simple choice $x = k^{-1/2}\mathbf{1}$ yields

$$\lambda_{\max} \geqslant \frac{\sum_{(i,j)} A_{i,j}}{k} = \frac{\sum_{j=1}^k 1 + \sum_{i=1}^k \sum_{j=1}^k d_{ij}/(\sum_{\ell=1}^k d_{i\ell}) + \sum_{j=1}^k \sum_{i=1}^k d_{ij}/(\sum_{\ell=1}^k d_{\ell j})}{k} = 3.$$

This neat property leads us to the expected result

$$T_{pp}(G) = (1+o(1))\log_{\lambda_{\max}} n \leqslant (1+o(1))\log_3 n,$$

and it also completes the informal argument that justifies the claim made in Theorems 1.3(b) and 1.5(b). In the rest of this section we will turn this argument step by step into a formal proof by filling in all missing pieces.

Obtaining an appropriate regular partition. An important ingredient in the previous sketch was the assumption that the given graph has a partition into a bounded number of equal parts, such that the bipartite graph induced by any two different parts looks like a random regular graph. This assumption is quite strong and very much not true in general. However, restricting ourselves to dense graphs we can actually come quite close to that. Let us begin with some definitions; the statements are taken from [30].

Definition 3.1 (density). Given a graph $G=(V,E)$ and two disjoint non-empty sets of vertices $X, Y \subseteq V$, we define the *density* of the pair (X, Y) as

$$d_G(X, Y) = \frac{e_G(X, Y)}{|X|\,|Y|}.$$

As usual, if the graph is clear from the context the index will be omitted. The next definition gives a partition that is close to the previously described properties; all sets in the partition have nearly the same size and nearly all pairs behave in a well-defined sense like regular bipartite random graphs.

Definition 3.2 ((ε, k_0, K_0)-Szemerédi partition). Let $G=(V,E)$ and $k \in \mathbb{N}$. We call $\Pi = \{V_i\}_{i\in[k]}$ an (ε, k_0, K_0)-*Szemerédi partition* of G if the following conditions are fulfilled.

(a) $V_1 \dot\cup \cdots \dot\cup V_k = V$.
(b) $k_0 \leqslant k \leqslant K_0$.
(c) $|V_1| \leqslant \cdots \leqslant |V_k| \leqslant |V_1| + 1$.
(d) For all but at most εk^2 pairs (V_i, V_j) of Π with $i<j$, we have that for all subsets $U_i \subseteq V_i$ and $U_j \subseteq V_j$ with $|U_i| \geqslant \varepsilon|V_i|$ and $|U_j| \geqslant \varepsilon|V_j|$,

$$|d(U_i, U_j) - d(V_i, V_j)| \leqslant \varepsilon.$$

A pair (V_i, V_j) satisfying the last condition is called ε-*regular*. For pairs (V_i, V_j) in Π we will abbreviate $d(V_i, V_j)$ to d_{ij}.

Next we state Szemerédi's Regularity Lemma. It guarantees that we will have a Szemerédi partition if the underlying graph is large enough.

Lemma 3.10 ([30], The Regularity Lemma). *For every $\varepsilon > 0$ and every $k_0 \in \mathbb{N}$ there exist $K_0 = K_0(\varepsilon, k_0)$ and n_0 such that every graph $G=(V,E)$ with at least $|V| = n \geqslant n_0$ vertices admits an* (ε, k_0, K_0)-Szemerédi partition.

The next lemma gives a useful property of regular pairs. In particular, with the exception of a small set only, all other vertices have a degree that is close to dN, where d is the density of the pair and N is the number of vertices in each part. In fact the statement also is true for arbitrary but not too small subsets of the parts.

Lemma 3.11. *Let $G=(V,E)$ be a graph, $\varepsilon > 0$ and $U, U' \subseteq V$. Suppose that (U, U') is an ε-regular pair, and let $W \subseteq U'$, $|W| \geqslant \varepsilon|U'|$. Furthermore, let $\mathcal{E}(U, W) \subseteq U$ be the largest set such that $|d(u, W) - d(U, U')| \geqslant \varepsilon$ for all $u \in \mathcal{E}(U, W)$. Then $|\mathcal{E}(U, W)| \leqslant 2\varepsilon|U|$.*

Proof. We will prove this by contradiction. Assume that $|\mathcal{E}(U, W)| \geqslant 2\varepsilon|U|$. Let us write $\mathcal{E}(U, W) = S \cup L$, where

$$S = \{u \in \mathcal{E}(U, W) \colon d(u, W) < d(U, U') - \varepsilon\}, \quad L = \{u \in \mathcal{E}(U, W) \colon d(u, W) > d(U, U') + \varepsilon\}.$$

Then $|S| \geqslant \varepsilon|U|$ or $|L| \geqslant \varepsilon|U|$. In the former case

$$d(S, W) = \frac{\sum_{u\in S} e(u, W)}{|S|\,|W|} = \frac{\sum_{u\in S} d(u, W)}{|S|} < d(U, U') - \varepsilon.$$

As $|S| \geqslant \varepsilon|U|$, $|W| \geqslant \varepsilon|U'|$, this contradicts the assumption that (U, U') is an ε-regular pair. The case $|L| \geqslant \varepsilon|U|$ follows analogously by showing that $d(L, W) > d(U, U') + \varepsilon$. $\square$

We call the set $\mathcal{E}(U, W)$ in Lemma 3.11 the exceptional set of U with respect to W. In particular, Lemma 3.11 implies that for every ε-regular pair (U, U') and *all* $W \subseteq U'$, $|W| \geqslant (1 - c\varepsilon)|U'|$, $c > 0$ we have

$$\begin{aligned} |d(u, W) - d(U, U')| &\leqslant |d(u, W) - d(u, U')| + |d(u, U') - d(U, U')| \\ &\leqslant (c+1)\varepsilon \quad \text{for all } u \in U \backslash \mathcal{E}(U, U'). \end{aligned} \tag{3.17}$$

Having done these preparations we can now determine a partition that comes close to the initially described properties.

Lemma 3.12. *Let $G_n = (V, E)$ be a graph on n vertices such that $\delta_{G_n} \geqslant \alpha n$ for some $\alpha > 0$. Then for all $\eta > 0$ and $k_0 > 1/\sqrt{\eta}$ there exists $n_0, K_0 \in \mathbb{N}$ such that for all G_n with $n \geqslant n_0$ there is an (η, k_0, K_0)-Szemerédi partition $\Pi = \{V_i\}_{i\in[k]}$ of G_n with the following property. There is $F \subseteq \Pi$ with $|F| \leqslant \eta k$ such that, for all $V_i \in \Pi \backslash F$,*

- *there are at most ηk non-η-regular pairs (V_i, V_j), $j \in [k]$, and*
- *there exists an exceptional set N_i, $|N_i| \leqslant \eta |V_i|$ such that*

$$d(u) \leqslant (1 + \eta)\frac{n}{k} \sum_{1\leqslant j\leqslant k} d(V_i, V_j) \quad \text{for all } u \in V_i \backslash N_i.$$

Proof. According to Lemma 3.10, for all $\xi > 0$ and $k_0 > 1/\sqrt{\xi}$, there are $n_0, K_0 \in \mathbb{N}$ such that for all G_n with $n \geqslant n_0$ there is a $k \in \mathbb{N}$ and a (ξ, k_0, K_0)-Szemerédi partition $\Pi = \{V_i\}_{i\in[k]}$ of G_n. Let $F \subseteq \Pi$ contain the parts $V_i \in \Pi$ such that there are at least $\sqrt{\xi} k$ other parts $V_j \in \Pi$ such that the pair (V_i, V_j) is not ξ-regular. As there are at most ξk^2 non-ξ-regular pairs, we infer that $|F| \leqslant \sqrt{\xi} k$. Let $V_i \in \Pi \setminus F$. Further, let $A_i \subseteq \Pi$ be such that (V_i, V_j) is a ξ-regular pair for all $V_j \in \Pi \setminus A_i$ and (V_i, V_j) is not ξ-regular for all $V_j \in A_i$. The definition of F implies that $|A_i| \leqslant \sqrt{\xi} k$. For these $V_j \in \Pi \setminus A_i$ let $\mathcal{E}_i(V_j) = \mathcal{E}(V_i, V_j)$ be the exceptional set of V_i with respect to V_j. On top of that let $N_i \subseteq V_i$ be the set of points in V_i that are in at least $\sqrt{\xi} k$ exceptional sets with respect to parts in $\Pi \setminus A_i$. As there are at most k exceptional sets and by Lemma 3.11 each exceptional set has at most $2\xi |V_i|$ vertices, we get that $|N_i| \leqslant 2\sqrt{\xi} |V_i|$. Let $V_i \in \Pi \backslash F$, $u \in V_i \backslash N_i$ and let $B(u) \subseteq \Pi \backslash A_i$ be the set of parts such that $u \in \mathcal{E}_i(V_j)$ for all $V_j \in B$. Then $|B| \leqslant \sqrt{\xi} k$ and

$$\begin{aligned} d(u) &= \sum_{1\leqslant j\leqslant k} |V_j| d(u, V_j) \\ &= \left(\sum_{V_j \in A_i \cup B} |V_j| d(u, V_j) + \sum_{V_j \in \Pi \backslash (A_i \cup B)} |V_j| d(u, V_j) \right) \\ &\leqslant \left| N(u) \cap \left(\bigcup_{V_j \in A_i \cup B \cup \{V_i\}} V_j \right) \right| + \sum_{1\leqslant j\leqslant k} |V_j| (d(V_i, V_j) + \xi). \end{aligned}$$

By the definition of F and as $u \in V_i \backslash N_i$, we get that

$$\left| \bigcup_{V_j \in A_i \cup B \cup \{V_i\}} V_j \right| \leqslant (\sqrt{\xi} k + \sqrt{\xi} k + 1)(n/k + 1) \leqslant 3\sqrt{\xi} n.$$

With that at hand and by using $d(u) \geqslant \alpha n$ and the fact that the sizes of the parts in Π differ by at most one, we obtain

$$d(u) \leqslant 3\sqrt{\xi} n + \frac{n}{k} \sum_{1\leqslant j\leqslant k} d(V_i, V_j) + 2\xi n \leqslant \frac{n}{k} \sum_{1\leqslant j\leqslant k} d(V_i, V_j) + 5\sqrt{\xi} d(u)/\alpha.$$

Let $\eta > 0$. Choosing ξ small enough such that

$$\max\{\xi, 2\sqrt{\xi}, 1/(1-5\sqrt{\xi}/\alpha)-1\} \leqslant \eta$$

implies the claim. □

The recursion relation. In this section we exploit the properties of the partition to study the expected number of informed vertices after one additional round; our aim is to establish a precise version of (3.16). In the remainder let $\|A\|_F = (\sum_{1\leqslant i\leqslant n}\sum_{1\leqslant j\leqslant n}|a_{i,j}|^2)^{1/2}$ denote the Frobenius norm of a matrix $A \in \mathbb{R}^{n\times n}$.

For the next lemma consider the setting of Theorems 1.3(b) and 1.5(b), that is, we are given an expander sequence $(G_n)_{n\in\mathbb{N}}$ with minimal degree $\delta_n \geqslant \alpha n$ for some $\alpha > 0$ and an $\varepsilon > 0$. We obtain a sequence of graphs $(\tilde{G}_n)_{n\in\mathbb{N}}$ by deleting up to a $1/2-\varepsilon$ fraction of the edges at each vertex in G_n. Further, let $\eta > 0$, $k_0 \in \mathbb{N}$ and $\Pi = \{V_i\}_{i\in[k]}$ be the (η, k_0, K_0)-Szemerédi partition of $\tilde{G}_n$ as given by Lemma 3.12. For that partition define $\mathcal{E}_{i,j} := \mathcal{E}(V_i, V_j)$ as the exceptional set of V_i with respect to V_j given by Lemma 3.11, $i \neq j \in [k]$, F and N_i as the exceptional sets from Lemma 3.12, $i \in \Pi \setminus F$. Moreover, let $\Pi_i = \{V_j \in \Pi \setminus F\colon (V_i, V_j) \text{ is } \eta\text{-regular}\}$ and note that

$$|\Pi_i| \geqslant (1-2\eta)k, \quad |N_i| \leqslant \eta|V_i| \quad \text{and} \quad |\mathcal{E}_{i,j}| \leqslant 2\eta|V_i| \quad \text{for all } i \in \Pi\setminus F, j \in \Pi_i. \tag{3.18}$$

Finally, define

$$\mathcal{H}_{i,j'} = N_i \cup \mathcal{E}_{i,j'}, \quad i \in \Pi\setminus F \text{ and } j \in \Pi_i$$

and

$$X_{t,i,j} = |I_t^{(pp)} \cap (V_i \setminus (N_i \cup \mathcal{E}_{i,j}))|, \quad i \in \Pi \setminus F \text{ and } j \in \Pi_i$$

as well as

$$X_{t,i} = \min_{j\in\Pi_i} X_{t,i,j}, \quad i \in \Pi\setminus F.$$

This definition guarantees that $|I_t^{(pp)}| \geqslant \|X_t\|_1$. The cornerstone of our proof is the following lemma, which bounds the growth of $X_t = (X_{t,i})_{i\in\Pi\setminus F}$ after one round.

Lemma 3.13. *Consider the situation as described above and assume additionally that $|X_{t,i}| \geqslant \log\log n$ for all $i \in \Pi\setminus F$ and that $|I_t^{(pp)}| \leqslant n/\log n$. Then, for all $\nu > 0$ and n large enough, there exists a symmetric matrix A with biggest eigenvalue $\lambda_{\max} \geqslant 1+2q-\nu$ and an error matrix ΔA with $\|\Delta A\|_F \leqslant \nu$ such that w.h.p.*

$$X_{t+1} \geqslant (A + \Delta A)X_t.$$

Proof. We set $I_t^{\mathcal{P},i} = I_t^{\mathcal{P}} \cap V_i$, $U_t^{\mathcal{P},i} = U_t^{\mathcal{P}} \cap V_i$ for $\mathcal{P} \in \{push, pull, pp\}$ and let

$$I_{t+1}^{\mathcal{P},i,j} \backslash I_t = \{u \in U_t \cap V_i \mid \text{there is } v \in I_t \cap V_j \text{ such that } u \text{ gets informed by } v \text{ using } \mathcal{P}\}$$

be the vertices in V_i newly informed in round $t+1$ by operations involving only vertices from V_i and V_j. Let $(i,j) \in \Pi\setminus F$. For all $u \in U_t^i$ we know that $d(u) \geqslant \alpha n/2$. Moreover, $|I_t^i| \leqslant |I_t| \leqslant n/\log n$. Thus the probability of $u \in U_t^i$ being informed by vertices in I_t^j via *pull* is $q|N(u) \cap I_t^j|/|N(u)| = o(1)$. As the events of u being informed by *push* and *pull* are independent,

$$\mathbb{P}[u \in I_{t+1}^{(push),i,j} \cap I_{t+1}^{(pull),i,j}] = o(1)\mathbb{P}[u \in I_{t+1}^{(push),i,j}].$$

Thus, for any set $S \in V$,

$$\mathbb{E}[|(I_{t+1}^{(pp),i,j} \setminus I_t) \cap S|] = (1-o(1))(\mathbb{E}[|(I_{t+1}^{(pull),i,j} \setminus I_t) \cap S|] + \mathbb{E}[|(I_{t+1}^{(push),i,j} \setminus I_t) \cap S|]). \tag{3.19}$$

Let $i \in \Pi \backslash F$ and $j \in \Pi_i$. We start by determining the expected number of vertices informed by *pull*. Further, set

$$D_i = (1+\eta)\frac{n}{k} \sum_{1 \leqslant \ell \leqslant k} d_{i\ell}.$$

According to Lemma 3.12, all $v \in U_t^i \setminus N_i$ have degree less than D_i. Let $j' \in \Pi_i$. Then

$$\mathbb{E}_t[|I_{t+1}^{(pull),i,j} \backslash (I_t \cup \mathcal{H}_{i,j'})|] = \sum_{u \in U_t^i \backslash \mathcal{H}_{i,j'}} q\frac{|N(u) \cap I_t^j|}{|N(u)|} \geqslant q\frac{e(U_t^i \setminus \mathcal{H}_{i,j'}, I_t^j)}{D_i}.$$

Since $|I_t^i| \leqslant |I_t| \leqslant n/\log n$, we get with room to spare that $|U_t^i \setminus \mathcal{H}_{i,j'}| \geqslant (1-5\eta)n/k$ for n large enough and all $j' \in \Pi_i$. Applying (3.17), where we choose $W = U_t^i \setminus \mathcal{H}_{i,j'}$, yields $|d(U_t^i \setminus \mathcal{H}_{i,j'}, u) - d_{ij}| \leqslant 6\eta$ for all $u \in V_j \setminus \mathcal{E}_{j,i}$. Thus

$$\begin{aligned}\mathbb{E}_t[|I_{t+1}^{(pull),i,j} \backslash (I_t \cup \mathcal{H}_{i,j'})|] &\geqslant q\frac{(d_{ij}-6\eta)|U_t^i \setminus \mathcal{H}_{i,j'}|\,|I_t^j \backslash \mathcal{E}_{j,i}|}{D_i} \\ &\geqslant (1-5\eta)q\frac{(d_{ij}-6\eta)|I_t^j \backslash (\mathcal{E}_{j,i} \cup N_j)|}{D_i k/n}.\end{aligned}$$

As $D_i = (1+\eta)n/k \sum_{1 \leqslant \ell \leqslant k} d_{i\ell}$, we get for

$$c_1 := (1-6\eta)(1+\eta)^{-1}$$

with $X_{t,j,i} = |I_t^j \backslash (\mathcal{E}_{j,i} \cup N_j)|$ that

$$\mathbb{E}_t[|I_{t+1}^{(pull),i,j} \backslash (I_t \cup \mathcal{H}_{i,j'})|] \geqslant c_1 \cdot q\frac{(d_{ij}-6\eta)X_{t,j,i}}{\sum_{1 \leqslant \ell \leqslant k} d_{i\ell}} \quad \text{for all } i \in \Pi \setminus F \text{ and } j, j' \in \Pi_i. \tag{3.20}$$

We continue with *push*. Let $i \in \Pi \backslash F$ and $j, j' \in \Pi_i$, and set (as before)

$$D_j = (1+\eta)\frac{n}{k} \sum_{1 \leqslant \ell \leqslant k} d_{\ell j}.$$

Then

$$\mathbb{E}_t[|I_{t+1}^{(push),i,j} \backslash (I_t \cup \mathcal{H}_{i,j'})|] = \sum_{u \in U_t^i \backslash \mathcal{H}_{i,j'}} \left(1 - \prod_{v \in N(u) \cap I_t^j} \left(1 - \frac{q}{|N(v)|}\right)\right).$$

According to Lemma 3.12 all $v \in I_t^j \backslash N_j$ have degree less than D_j and furthermore $|I_t| = o(n) = o(D_j)$. Thus (3.3) yields the estimate

$$\begin{aligned}\mathbb{E}_t[|I_{t+1}^{(push),i,j} \backslash (I_t \cup \mathcal{H}_{i,j'})|] &\geqslant \sum_{u \in U_t^i \backslash \mathcal{H}_{i,j'}} \left(1 - \left(1 - \frac{q}{D_j}\right)^{|N(u) \cap (I_t^j \backslash N_j)|}\right) \\ &\geqslant (1-o(1)) \sum_{u \in U_t^i \backslash \mathcal{H}_{i,j'}} q\frac{|N(u) \cap (I_t^j \backslash N_j)|}{D_j}.\end{aligned} \tag{3.21}$$

The remaining steps are similar to the previously considered case of *pull*. By assumption we have that $|I_t^j \backslash \mathcal{H}_{j,i}| = X_{t,j,i}$ and as $|I_t^i| \leqslant |I_t| \leqslant n/\log n$ we obtain that $|U_t^i \backslash \mathcal{H}_{i,j'}| \geqslant (1-5\eta)n/k$ for n large enough and all $j' \in \Pi_i$. Using (3.17) we obtain that $|d(U_t^i \setminus \mathcal{H}_{i,j'}, u) - d_{ij}| \leqslant 6\eta$ for all $u \in V_j \setminus \mathcal{E}_{j,i}$. Thus

$$\mathbb{E}_t[|I_{t+1}^{(push),i,j} \backslash (I_t \cup \mathcal{H}_{i,j'})|] \geqslant (q - o(1)) \frac{e(U_t^i \setminus \mathcal{H}_{i,j'}, I_t^j \backslash (N_j \cup \mathcal{E}_{j,i}))}{D_j}$$
$$\geqslant (q - o(1)) \frac{(d_{ij} - 6\eta)|U_t^i \setminus \mathcal{H}_{i,j'}| X_{t,j,i}}{D_j}.$$

Using the fact that $D_j = (1+\eta)n/k \sum_{1 \leqslant \ell \leqslant k} d_{\ell j}$, we get for the same constant c_1 as in (3.20) and n large enough

$$\mathbb{E}_t[|I_{t+1}^{(push),i,j} \backslash (I_t \cup \mathcal{H}_{i,j'})|] \geqslant c_1 \cdot q \frac{(d_{ij} - 6\eta) X_{t,j,i}}{\sum_{1 \leqslant \ell \leqslant k} d_{\ell j}} \quad \text{for all } i \in \Pi \setminus F \text{ and } j, j' \in \Pi_i. \tag{3.22}$$

With (3.19), we can combine the results for *pull*, (3.20), and *push*, (3.22), to get for $c_2 := c_1 - \eta$

$$\mathbb{E}_t[|I_{t+1}^{(pp),i,j} \backslash (I_t^i \cup \mathcal{H}_{i,j'})|] \geqslant c_2 \cdot q \left(\frac{d_{ij} - 6\eta}{\sum_{1 \leqslant \ell \leqslant k} d_{i\ell}} + \frac{d_{ij} - 6\eta}{\sum_{1 \leqslant \ell \leqslant k} d_{\ell j}} \right) X_{t,j,i} \quad \text{for all } i \in \Pi \setminus F, j, j' \in \Pi_i. \tag{3.23}$$

Next we will show how we can exploit (3.23) to obtain (a lower bound for) $\mathbb{E}_t[|(I_{t+1}^{(pp),i} \setminus I_t)|]$. Let $i \in \Pi \setminus F$ and $u \in U_t^i$. Using $|I_t| = o(n)$ and (3.3) we obtain

$$\mathbb{P}_t[u \in I_{t+1}^{(push),i} \setminus I_t] = 1 - \prod_{i \in N(u) \cap I_t} \left(1 - \frac{1}{|N(i)|}\right) = (1 - o(1)) \sum_{i \in N(u) \cap I_t} \frac{1}{|N(i)|}.$$

Let $W \subseteq V$. Using (3.19), the previous equation and that Π is a partition, we get

$$\mathbb{E}_t[|(I_{t+1}^{(pp),i} \setminus I_t) \cap W|] = (1 - o(1)) \sum_{u \in U_t^i \cap W} \left(\frac{|N(u) \cap I_t|}{|N(u)|} + \sum_{i \in N(u) \cap I_t} \frac{1}{|N(i)|} \right)$$
$$= (1 - o(1)) \sum_{u \in U_t^i \cap W} \left(\sum_{j \in [k]} \left(\frac{|N(u) \cap I_t \cap V_j|}{|N(u)|} + \sum_{i \in N(u) \cap I_t \cap V_j} \frac{1}{|N(i)|} \right) \right)$$
$$= (1 - o(1)) \sum_{j \in [k]} \mathbb{E}_t[|(I_{t+1}^{(pp),i,j} \setminus I_t) \cap W|].$$

Choose $W = V \setminus \mathcal{H}_{i,j'}$. Then the previous equation implies

$$\mathbb{E}_t[|I_{t+1}^{(pp),i} \backslash (I_t \cup \mathcal{H}_{i,j'})|] \geqslant (1 - o(1)) \sum_{j \in \Pi \setminus F} \mathbb{E}_t[|I_{t+1}^{(pp),i,j} \backslash (I_t \cup \mathcal{H}_{i,j'})|] \quad \text{for all } i \in \Pi \setminus F, j' \in \Pi_i,$$

which in turn, using (3.23) and $X_{t,j,i} \geqslant X_{t,j}$ for all $j \in \Pi \setminus F$ and $i \in \Pi_j$, implies for $c := c_2 - \eta$

$$\mathbb{E}_t[X_{t+1,i,j'}] \geqslant X_{t,i} + c \cdot q \sum_{j \in \Pi_i} \left(\frac{d_{ij} - 6\eta}{\sum_{1 \leqslant \ell \leqslant k} d_{i\ell}} + \frac{d_{ij} - 6\eta}{\sum_{1 \leqslant \ell \leqslant k} d_{\ell j}} \right) X_{t,j} \quad \text{for all } i \in \Pi \setminus F, j' \in \Pi_i. \tag{3.24}$$

Assume that (3.24) holds not only in expectation but also for a slightly smaller c, say $c - \eta$, with high probability. We are going to show this at the end of the proof. Using this assumption and a union bound over $j' \in \Pi_i$ gives w.h.p.

$$X_{t+1,i} = \min_{j' \in \Pi_i} X_{t+1,i,j'} \geqslant \langle a_i, (X_{t,j})_{j \in \Pi_i} \rangle \quad \text{for all } i \in \Pi \setminus F, \tag{3.25}$$

where for $i \in \Pi \setminus F$ and $j \in \Pi_i$ we have

$$a_{ij} = \mathbb{1}[i=j] + c \cdot q\left(\frac{d_{ij} - 6\eta}{\sum_{1 \leqslant \ell \leqslant k} d_{i\ell}} + \frac{d_{ij} - 6\eta}{\sum_{1 \leqslant \ell \leqslant k} d_{\ell j}}\right). \tag{3.26}$$

Let A be the $|\Pi \setminus F| \times |\Pi \setminus F|$ matrix with entries as in the previous equation, that is, $A = (a_{ij})_{(i,j)\in(\Pi\setminus F)^2}$ is given by (3.26) for all $(i,j) \in (\Pi \setminus F)^2$. Note that A is symmetric. Then we obtain from (3.25)

$$X_{t+1} \geqslant B \cdot X_t,$$

with $B = A + \Delta A$, where

$$(\Delta A)_{ij} = \begin{cases} 0 & i \in \Pi \backslash F \text{ and } j \in \Pi_i, \\ -a_{ij} & i \in \Pi \backslash F \text{ and } j \in \Pi \setminus (F \cup \Pi_i). \end{cases}$$

Set

$$F' := \{(i,j) \in (\Pi \setminus F)^2 \mid j \in \Pi \setminus (F \cup \Pi_i)\}.$$

As $d(u) \geqslant \alpha n/2$ for all $u \in V$ and some $\alpha > 0$, we also know that $\sum_{1 \leqslant \ell \leqslant k} d_{\ell j} \geqslant k\alpha/2$. Together with $0 \leqslant d_{i,j} \leqslant 1$ for all $(i,j) \in [k]^2$, we get that

$$\left|\frac{d_{ij} - 6\eta}{\sum_{1 \leqslant \ell \leqslant k} d_{i\ell}}\right| \leqslant \frac{2}{\alpha k}.$$

Using the fact that $|F'| \leqslant 2\eta k^2$ (see (3.18)), we obtain

$$\|\Delta A\|_F^2 = \sum_{(i,j)\in F'} a_{ij}^2 \leqslant \sum_{(i,j)\in F'} \left(\frac{4}{\alpha k}\right)^2 \leqslant 2\eta k^2 \left(\frac{4}{\alpha k}\right)^2 = \frac{4^2 \cdot 2 \cdot \eta}{\alpha^2}$$

and thus $\|\Delta A\|_F \leqslant 4\sqrt{2\eta}/\alpha$. This leaves us with bounding the biggest eigenvalue $\lambda_{\max}$ of A. Using the well-known inequality for symmetric matrices,

$$\lambda_{\max} \geqslant \sum_{(i,j)\in(\Pi\setminus F)^2} A_{ij}/|\Pi \setminus F|,$$

we obtain

$$\begin{aligned} \lambda_{\max} &\geqslant \frac{1}{|\Pi \setminus F|} \sum_{(i,j)\in(\Pi\setminus F)^2} a_{ij} \\ &\geqslant \frac{1}{|\Pi \setminus F|}\left(\sum_{(i,i)\in(\Pi\setminus F)^2} 1 + \sum_{(i,j)\in[k]^2} \frac{cq(d_{ij} - 6\eta)}{\sum_{1 \leqslant \ell \leqslant k} d_{i\ell}}\right. \\ &\qquad \left. + \sum_{(i,j)\in[k]^2} \frac{cq(d_{ij} - 6\eta)}{\sum_{1 \leqslant \ell \leqslant k} d_{\ell j}} - 2 \sum_{i\in[k]\setminus(\Pi\setminus F)} \sum_{j\in[k]} \frac{cq}{\sum_{1 \leqslant \ell \leqslant k} d_{\ell j}}\right). \end{aligned}$$

Note that $|\Pi \setminus F| \geqslant (1-\eta)k$, $|[k] \setminus (\Pi \setminus F)| \leqslant \eta k$. Moreover, $\sum_{1 \leqslant \ell \leqslant k} d_{\ell j} \geqslant \alpha k/2$ for all $j \in [k]$. Thus

$$\lambda_{\max} \geqslant 1 + \frac{1}{k}\left(cqk + cqk - 12cq \sum_{(i,j)\in[k]^2} \frac{\eta}{\sum_{1 \leqslant \ell \leqslant k} d_{\ell j}} - 2cq\frac{\eta k^2}{\alpha k/2}\right) \geqslant 1 + 2cq(1 - 8\eta/\alpha).$$

Choosing η small enough such that $2q(1 - c(1 - 8\eta/\alpha)), 4\sqrt{2\eta}/\alpha \leqslant \nu$ implies the claim of this lemma.

This leaves us with proving that (3.24) also holds with high probability. As $|I_{t+1}^{(pp)}|$ conditioned on I_t is a self-bounding function, so is $|I_{t+1}^{(pp),i} \setminus I_t^i|$ for all $i \in \Pi \setminus F$ and therefore also $|I_{t+1}^{(pp),i} \setminus (I_t \cup \mathcal{H}_{i,j'})| =: Y_{t+1,i,j'}$ for all $i \in \Pi \setminus F$ and $j' \in \Pi_i$. Note that $Y_{t+1,i,j'} = X_{t+1,i,j'} - X_{t,i,j'}$. Lemma 2.2 yields that

$$\mathbb{P}_t[Y_{t+1,i,j'} \geqslant (1 - \mathbb{E}_t[Y_{t+1,i,j'}]^{-1/3})\mathbb{E}_t[Y_{t+1,i,j'}]] \geqslant 1 - \mathbb{E}_t[Y_{t+1,i,j'}]^{-1/3}$$

and therefore, setting

$$Z_{t,i} = c \cdot q \sum_{j \in \Pi_i} \left(\frac{d_{ij} - 6\eta}{\sum_{1 \leqslant \ell \leqslant k} d_{i\ell}} + \frac{d_{ij} - 6\eta}{\sum_{1 \leqslant \ell \leqslant k} d_{\ell j}} \right) X_{t,j} \quad \text{for all } i \in \Pi \setminus F,$$

and using (3.24), *i.e.* $\mathbb{E}_t[Y_{t+1,i,j'}] \geqslant Z_{t,i}$ for all $i \in \Pi \setminus F$ and $j' \in \Pi_i$, we get with probability at least $1 - k^3 Z_{t,i}^{-1/3}$

$$Y_{t+1,i,j'} \geqslant (1 - Z_{t,i}^{-1/3}) Z_{t,i} \quad \text{for all } i \in \Pi \setminus F \text{ and } j' \in \Pi_i.$$

This and $|I_t^i| \geqslant X_{t,i}$ for all $i \in \Pi \setminus F$ implies that (3.24) also holds with high probability for a marginally smaller c, as claimed. □

Extension. We now solve the linear recurrence relation above and extend it to more than one round to get an upper bound on the runtime of *push&pull*. We first state a Chernoff bound that will be very useful in the next lemma.

Lemma 3.14 ([26]). *Let $\varepsilon, \delta > 0$. Suppose that $X_1, \ldots, X_n$ are independent geometric random variables with parameter δ, so $\mathbb{E}[X_i] = 1/\delta$ for each i. Let $X := \sum_{1 \leqslant i \leqslant n} X_i$, $\mu = \mathbb{E}[X] = n/\delta$. Then*

$$\mathbb{P}[X \geqslant (1+\varepsilon)\mu] \leqslant e^{-n(\varepsilon - \log(1+\varepsilon))} \leqslant e^{-\varepsilon^2 n/2(1+\varepsilon)}$$

Together with Lemma 2.11 the following lemma implies Theorems 1.3(b) and 1.5(b).

Lemma 3.15. *Consider the setting of Theorems 1.3(b) and 1.5(b) and let $I_t = I_t^{(pp)}$. The following statements hold w.h.p.*

(a) *Let $S \subseteq V_n$, $|S| = \Theta(n)$. Then there is $t = \Theta(\log\log n)$ such that w.h.p. $|I_t| \geqslant |I_t \cap S| \geqslant \log\log n$.*
(b) *Let $\log\log n \leqslant |I_t| \leqslant n/\log n$. Then there is $\tau \leqslant \log_{1+2q}(n/|I_t|) + o(\log n)$ such that $|I_{t+\tau}| > n/\log n$.*
(c) *Let $n/\log n \leqslant |I_t| \leqslant n - n/\log n$. Then there is $\tau = o(\log n)$ such that $|I_{t+\tau}| > n - n/\log n$.*
(d) *Let $|I_t| \geqslant n - n/\log n$ and $q = 1$. Then there is $\tau = o(\log n)$ such that $|I_{t+\tau}| = n$.*

Proof. As $|I_t^{(pp)}| \geqslant |I_t^{(pull)}|$ clearly (c) and (d) follow from Lemma 3.2. We show (a) by determining a lower bound for the probability that an arbitrary vertex gets informed after a constant number of rounds. Set $\beta = \min\{\alpha, \epsilon\}$, let $S_0 = \{u\}$ and choose $w \in V$, $w \neq u$. By Lemma 2.9 there is $d \leqslant 8/\beta^2 + 2$ and $c = (\beta^4/64)^{8/\beta^2+3} \in (0,1)$ such that there are at least cn^{d-1} paths of (edge) length d from u to w. Let $\gamma = (u, v_1, \ldots, v_{d-1}, w)$ be such a path from u to w, and let A_γ denote the event that w is informed via γ after exactly d rounds performing only *push* operations, that is, A_γ is the event that in the first round the randomly selected neighbour of u is v_1, in the second round the randomly selected neighbour of v_1 is v_2 and so forth, until in the dth round the randomly selected neighbour of v_{d-1} is w. Obviously, the probability of A_γ is bounded from below by n^{-d}. Further, let $\gamma' \neq \gamma$ be another path from u to w with length d. As γ and γ' differ by at least one edge we readily obtain that $\mathbb{P}[A_\gamma \cap A_{\gamma'}] = 0$. Let Γ denote the set of all paths with length d from u to w. Having done these preparations we use them to conclude for all $w \in V$ and $t \geqslant 0$

$$\mathbb{P}_t[w \in I_{t+d}] \geqslant \mathbb{P}_t\left[\bigcup_{\gamma\in\Gamma} A_\gamma\right] \geqslant \sum_{\gamma\in\Gamma} \mathbb{P}_t[A_\gamma] \geqslant \sum_{\gamma\in\Gamma} n^{-d} \geqslant \frac{c}{n}. \tag{3.27}$$

We define a modified protocol as follows. Wait $d := \lceil 8/\beta^2 + 2\rceil$ rounds, after that with probability c choose one uninformed vertex uniformly at random and set it as informed. Repeat. Call the vertices informed by this algorithm $I_t^\star$. Then the probability of any vertex being informed after d rounds is

$$\mathbb{P}_t[v \in I_{t+d}^\star | v \notin I_t^\star] = c/n.$$

Thus, for any $t \geqslant 0$,

$$\mathbb{P}_t[v \in I_{t+d} | v \in U_t] \geqslant \mathbb{P}_t[v \in I_{t+d}^\star | v \notin I_t^\star] = c/n.$$

Note that for any $s \in \mathbb{N}$ the set $I_{sd}^\star$ is generated by a very simple procedure: s times independently, with probability c, we choose a random vertex and put it into $I_{sd}^\star$. Thus $|I_{sd}^\star \cap S|$ is binomially distributed with s trials, where each one has success probability $c|S|/n = \Theta(c)$; it follows readily that $|I_{sd}^\star \cap S|$ concentrates around a multiple of s for large s, and the claim follows by choosing $s = \Theta(\log\log n)$.

This leaves (b) to be shown. Part (a) implies that there is some $t_0 = o(\log n)$ such that $X_{t_0,i} = \Theta(\log\log n)$ for all $i \in \Pi \setminus F$ by choosing $S = V_i \setminus (N_i \cup \mathcal{E}_{i,j}), j \in \Pi_i$ and applying a union bound over i and j. Thus we can apply Lemma 3.13. It gives w.h.p., say with probability $1 - g(n) = 1 - o(1)$, that $X_{t+1} \geqslant (A + \Delta A)X_t$, A has maximal eigenvalue $\lambda_{\max}(A) \geqslant 1 + 2q - \nu$ and $\|\Delta A\|_F \leqslant \nu$. Then $B := A + \Delta A$ has maximal eigenvalue

$$\lambda_{\max}(B) \geqslant \lambda_{\max}(A) - \|\Delta A\|_F \geqslant 1 + 2q - 2\nu$$

(Theorem of Wielandt and Hoffmann; see *e.g.* [24]).

Set $f(n) := (\log(n/\log n))^{2/3}$. Our assumptions guarantee that $f(n) = \omega(1)$ and $f(n) = o(\log n)$. Moreover, set

$$\tau := \frac{1}{1-g(n)} \cdot \frac{\log(n/\log n)}{\log(\lambda_{\max}(B))} + f(n) = \frac{\log n}{\log(\lambda_{\max}(B))} + o(\log n).$$

Let $(X_i)_{i\in\mathbb{N}}$ be independent and identically distributed geometric random variables with expectation $1 - g(n)$. Set $X = X_1 + X_2 + \cdots + X_{\mathcal{T}}$ with $\mathcal{T} = \log(n/\log n)/\log(\lambda_{\max}(B))$. We show that $\mathbb{P}[X \leqslant \tau] = 1 - o(1)$. To see this, note first that by linearity of expectation $\mathbb{E}[X] = \tau - f(n)$. Then, by Lemma 3.14,

$$\mathbb{P}[X \leqslant \tau] = \mathbb{P}\left[X \leqslant \left(1 + \frac{f(n)}{\tau + f(n)}\right)\mathbb{E}[X]\right] \geqslant 1 - \exp\left(-\Theta\left(\frac{f(n)^2}{\tau}\right)\right) = 1 - o(1).$$

Thus we have w.h.p.

$$|I_{t+\tau}| \geqslant \|X_{t+\tau}\|_1 \geqslant \|B^{\mathcal{T}} X_{t_0}\|_1.$$

Let v be an eigenvector of B to $\lambda_{\max}(B)$. As $v \neq 0$ there is an index ℓ such that $v_\ell \neq 0$. Without loss of generality we can assume that $v_\ell = 1$, as v/v_ℓ is also an eigenvector to $\lambda_{\max}(B)$. Thus $(B^{\mathcal{T}} v)_\ell = \lambda_{\max}(B)^{\mathcal{T}}$, $(B^{\mathcal{T}}(X_{t_0} - v))_i \geqslant 0$ for all $1 \leqslant i \leqslant k$ and therefore

$$|I_{t+\tau}| \geqslant (B^{\mathcal{T}} X_{t_0})_\ell \geqslant (B^{\mathcal{T}}(v + X_{t_0} - v))_\ell = (B^{\mathcal{T}} v)_\ell + (B^{\mathcal{T}}(X_{t_0} - v))_\ell \geqslant (B^{\mathcal{T}} v)_\ell \geqslant \lambda_{\max}(B)^{\mathcal{T}}.$$

Our choice of $\mathcal{T}$ yields w.h.p. $|I_{t+\tau}| \geqslant \lambda_{\max}(B)^{\mathcal{T}} \geqslant n/\log n$. Note that, since $\nu > 0$ was chosen arbitrarily, we in fact have that $\tau \leqslant \log_{1+2q}(n) + o(\log n)$, and the proof is completed. □

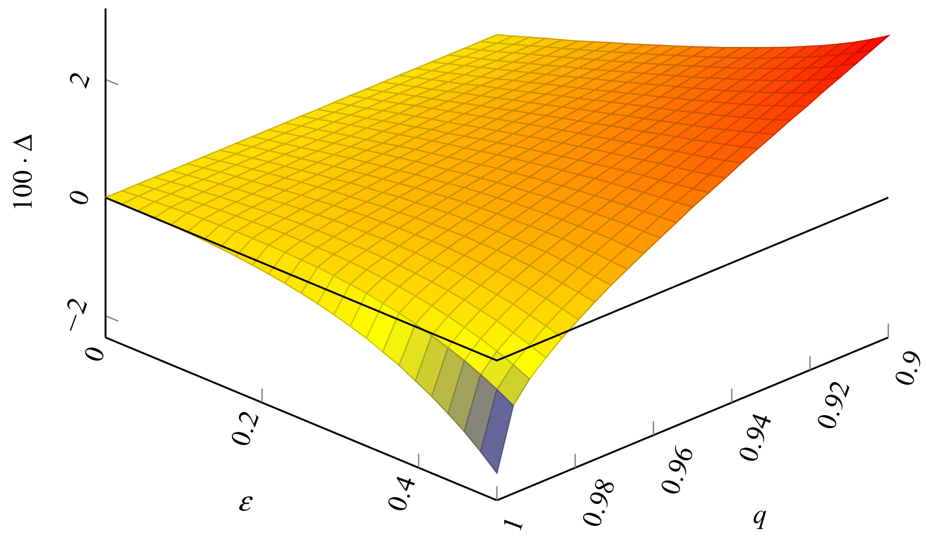

Figure 1. Plotted values of Δ in $T_{pp}(G_n(\varepsilon), q) - c_{pp} \log n = \Delta \log n + o(\log n)$, for $0.9 < q < 1$ and $0 < \varepsilon < 1/2$.

3.6 Proof of Theorem 1.4(b) – edge deletions may slow down push&pull

For any $0 < \varepsilon < 1/2$, $q \in (0, 1)$ we consider a sequence of graphs $(G_n(\varepsilon))_{n\in\mathbb{N}} = ((V_n, E_n))_{n\in\mathbb{N}}$ that is similar to the one studied in the proof of Theorem 1.4(a). Let $V_n = A_n \cup B_n$ with $A_n := \{1, \ldots, \lfloor n/2 \rfloor\}$, $B_n := \{\lfloor n/2 \rfloor + 1, \ldots, n\}$ and $\deg(v) = n - 1$ for all $v \in A_n$. Let the induced subgraph of B_n be a random graph in which each edge is included independently with probability $p = 1 - 2\varepsilon$. We know and it is easy to show (see *e.g.* [15, Section IV]) that w.h.p. this subgraph is almost regular, that is,

$$d_{B_n}(v) = (1 + o(1))(1 - 2\varepsilon)n/2 \quad \text{for all } v \in B_n, \tag{3.28}$$

and is an expander, which means that for every $S_n \subseteq B_n$, $1 \leqslant |S_n| \leqslant n/4$ and $d_{B_n} := (1 - 2\varepsilon)n/2$ we have

$$e(S_n, B_n \backslash S_n) = (1 + o(1)) \frac{d_{B_n}|S_n|\,|B_n \setminus S_n|}{|B_n|} = (1 - 2\varepsilon + o(1))|S_n|\,|B_n \setminus S_n|. \tag{3.29}$$

First we give a statement that describes the expected number of informed vertices after performing one round of *push&pull*.

Lemma 3.16. *Let $G_n(\varepsilon) = (A_n \cup B_n, E_n)$ be as above.*

(a) *Let $\sqrt{\log n} \leqslant |I_t| \leqslant n/\log n$ and set*

$$X_t = (|I_t^{(pp),(A)}|, |I_t^{(pp),(B)}|) := (|I_t^{(pp)} \cap A_n|, |I_t^{(pp)} \cap B_n|).$$

Then $\mathbb{E}_t[X_{t+1}] = (1 + o(1))MX_t$, where

$$M = \begin{pmatrix} 1 + q & q\big(1 + \varepsilon/(2 - 2\varepsilon)\big) \\ q\big(1 + \varepsilon/(2 - 2\varepsilon)\big) & 1 + q\big(1 - 2\varepsilon/(2 - 2\varepsilon)\big) \end{pmatrix}.$$

(b) *Let $|U_t^{(pp)}| \leqslant n/\log n$. Then*

$$\mathbb{E}_t[|U_{t+1}^{(pp)}|] \leqslant (1 + o(1))e^{-q(1/2 + (1/2 - \varepsilon)/(1 - \varepsilon))}(1 - q)|U_t|.$$

Proof. For $J \in \{A, B\}$, $J_n \in \{A_n, B_n\}$ set $U_t^{(J)} := U_t \cap J_n$, $I_t^{(J)} := I_t \cap J_n$ and $I_{t+1}^{(pp),(J)} = I_{t+1}^{(pp)} \cap J_n$. We first prove (a) by computing the expected number of informed vertices after a single round. Since

$d(u) = \Theta(n)$ for all $u \in V_n$ and $|I_t| \leqslant n/\log n$, the probability of $u \in U_t$ being informed by *pull* is

$$\mathbb{P}_t[u \in I_{t+1}^{(pull)} \setminus I_t] = \frac{q|N(u) \cap I_t|}{|N(u)|} = o(1).$$

As the events of u being informed by *push* and *pull* are independent, we have

$$\mathbb{P}_t[u \in (I_{t+1}^{(push)} \cap I_{t+1}^{(pull)}) \setminus I_t] = o(1)\mathbb{P}_t[u \in I_{t+1}^{(push)} \setminus I_t].$$

Thus

$$\mathbb{E}_t[|I_{t+1}^{(pp)} \setminus I_t|] = (1 + o(1))(\mathbb{E}_t[|I_{t+1}^{(push)} \setminus I_t|] + \mathbb{E}_t[|I_{t+1}^{(pull)} \setminus I_t|]).$$

We look at *pull* in detail first. Recall that

$$\deg(v) = n - 1 \text{ for all } v \in A_n \quad \text{and} \quad \deg(v) = (1 + o(1))(1 - \varepsilon)n \text{ for all } v \in B_n.$$

Moreover, using (3.29), we obtain

$$\begin{aligned}
\mathbb{E}_t[|I_{t+1}^{(pull)} \setminus I_t|] &= \sum_{u \in U_t} q \frac{|N(u) \cap I_t|}{|N(u)|} \\
&= \sum_{u \in U_t^{(A)}} q \frac{|N(u) \cap I_t|}{|N(u)|} + \sum_{u \in U_t^{(B)}} q \frac{|N(u) \cap I_t|}{|N(u)|} \\
&= (q + o(1)) \frac{n}{2} \left(\frac{|I_t^{(A)}| + |I_t^{(B)}|}{n} + \frac{|I_t^{(A)}| + (1 - 2\varepsilon)|I_t^{(B)}|}{(1 - \varepsilon)n} \right)
\end{aligned}$$

and thus

$$\begin{aligned}
\mathbb{E}_t[|I_{t+1}^{(pull),(A)} \setminus I_t|] &= (q + o(1)) \frac{|I_t^{(A)}| + |I_t^{(B)}|}{2}, \\
\mathbb{E}_t[|I_{t+1}^{(pull),(B)} \setminus I_t|] &= (q + o(1)) \frac{|I_t^{(A)}| + (1 - 2\varepsilon)|I_t^{(B)}|}{2(1 - \varepsilon)}.
\end{aligned}$$

Next we consider *push*. By using $|I_t| = o(n) = o(\delta_{G_n(\varepsilon)})$ and (3.3), we obtain

$$\begin{aligned}
\mathbb{E}_t[|I_{t+1}^{(push)} \setminus I_t|] &= \sum_{u \in U_t} 1 - \prod_{i \in N(u) \cap I_t} \left(1 - \frac{q}{|N(i)|}\right) \\
&= \sum_{u \in U_t} (1 + o(1)) \sum_{i \in N(u) \cap I_t} \frac{q}{|N(i)|} \\
&= (q + o(1)) \sum_{u \in U_t} \left(\frac{|I_t^{(A)}|}{n} + \frac{\mathbb{1}[u \in U_t^{(A)}]|I_t^{(B)}| + \mathbb{1}[u \in U_t^{(B)}]|N(u) \cap I_t^{(B)}|}{(1 - \varepsilon)n} \right)
\end{aligned}$$

and thus, with $|U_t^{(A)}|, |U_t^{(B)}| = (1 - o(1))n/2$ and (3.29),

$$\begin{aligned}
\mathbb{E}_t[|I_{t+1}^{(push),(A)} \setminus I_t|] &= (q + o(1)) \left(\frac{|I_t^{(A)}|}{2} + \frac{|I_t^{(B)}|}{2} + \frac{\varepsilon|I_t^{(B)}|}{2(1 - \varepsilon)} \right), \\
\mathbb{E}_t[|I_{t+1}^{(push),(B)} \setminus I_t|] &= (q + o(1)) \left(\frac{|I_t^{(A)}|}{2} + \frac{|I_t^{(B)}|}{2} - \frac{\varepsilon|I_t^{(B)}|}{2(1 - \varepsilon)} \right).
\end{aligned}$$

Accumulating the calculated expectations for *pull* and *push* yields the claim.

Next we show (b). The assumption implies that $|I_t| = (1 - o(1))n$ and therefore $|I_t^{(A)}| = |I_t^{(B)}| = (1 - o(1))n/2$. Let D_u be the event that an uninformed vertex u does not get informed by the *push* algorithm; let E_u be the corresponding event for *pull*. Then D_u and E_u are independent and $D_u \cap E_u$ is the event that u does not get informed in the current round. Let $u \in U_t^{(A)}$, and then

$$\begin{aligned}\mathbb{P}_t[D_u] &= \prod_{v\in I_t^{(A)}}\left(1-\frac{q}{|N(v)|}\right)\prod_{v\in I_t^{(B)}}\left(1-\frac{q}{|N(v)|}\right)\\ &= (1-o(1))\left(1-\frac{q}{n}\right)^{|I_t^{(A)}|}\left(1-\frac{q}{(1-\varepsilon)n}\right)^{|I_t^{(B)}|}\\ &= e^{-q(1/2+1/(2(1-\varepsilon)))}+o(1)\\ &\leqslant e^{-q(1/2+(1-2\varepsilon)/(2(1-\varepsilon)))}+o(1)\end{aligned}$$

and

$$\mathbb{P}_t[E_u] = 1-\frac{q|N(u)\cap |I_t||}{|N(u)|} = 1-\frac{q|I_t|}{n-1} = 1-q+o(1).$$

Now consider $u \in U_t^{(B)}$; then according to (3.28) we have

$$|N(u)\cap I_t^{(B)}| = |N(u)\cap B_n| - |N(u)\cap U_t^{(B)}| = (1+o(1))(1-2\varepsilon)n/2.$$

Therefore

$$\begin{aligned}\mathbb{P}_t[D_u] &= \prod_{v\in I_t^{(A)}}\left(1-\frac{q}{|N(v)|}\right)\prod_{v\in N(u)\cap I_t^{(B)}}\left(1-\frac{q}{|N(v)|}\right)\\ &= (1-o(1))e^{-q/2}\left(1-\frac{q}{(1-\varepsilon)n}\right)^{|N(u)\cap I_t^{(B)}|}\\ &= e^{-q(1/2+(1-2\varepsilon)/(2(1-\varepsilon)))}+o(1)\end{aligned}$$

and

$$\mathbb{P}_t[E_u] = 1-\frac{q|N(u)\cap |I_t||}{|N(u)|} = 1-(1+o(1))\frac{q(|I_t^{(A)}|+|N(u)\cap I_t^{(B)}|)}{(1-\varepsilon)n} = 1-q+o(1).$$

Combining the results for $u \in U_t^{(A)}$ and $u \in U_t^{(B)}$, we get

$$\mathbb{E}_t[|U_{t+1}|] = \sum_{u\in U_t}\mathbb{P}_t[D_u]\mathbb{P}_t[E_u] \leqslant (1+o(1))e^{-q(1/2+(1/2-\varepsilon)/(1-\varepsilon))}(1-q)|U_t|. \qquad \square$$

Remark 3.1. Let $\lambda_{\max}$ be the greatest eigenvalue of M as defined in Lemma 3.16(a). Then

$$\lambda_{\max} = 1+2q+\left(2q\left(\sqrt{(\varepsilon^2/2-\varepsilon+1)}-1\right)+q\varepsilon\right)/(2-2\varepsilon) > 1+2q.$$

Next comes a lemma that bounds the runtime of *push&pull* on $G_n(\varepsilon)$. In particular, (a) and (c) of Lemma 3.17 provide a lower bound on the runtime, and (a), (b) and (d) of Lemma 3.17 together with Lemma 3.15(a) provide an upper bound.

Lemma 3.17. *Let $I_t = I_t^{(pp)}$, $\varepsilon > 0$ and $\lambda = \lambda_{\max}(M)$ be the greatest eigenvalue of M as given in Lemma 3.16(a). Consider $G_n(\varepsilon)$.*

(a) *Let $\sqrt{\log n} \leqslant |I_t| \leqslant n/\log n$. Then there are $\tau_1, \tau_2 = \log_\lambda(n/|I_t|) + o(\log n)$ such that $|I_{t+\tau_1}| < n/\log n < |I_{t+\tau_1}|$.*

(b) *Let $n/\log n \leqslant |I_t| \leqslant n - n/\log n$. Then there is $\tau = o(\log n)$ such that $|I_{t+\tau}| > n - n/\log n$.*
(c) *Let $|I_t| \leqslant n/\log n$. Then there is*

$$\tau \geqslant \log n/\log\left((1-q)^{-1}\exp\left(q(1/2+(1/2-\varepsilon)/(1-\varepsilon))\right)\right) - o(\log n)$$

such that $|I_{t+\tau}| < n$.
(d) *Let $|I_t| \geqslant n - n/\log n$ and $q \in (0,1)$. Then there is*

$$\tau \leqslant \log n/\log\left((1-q)^{-1}\exp\left(q(1/2+(1/2-\varepsilon)/(1-\varepsilon))\right)\right) + o(\log n)$$

such that $|I_{t+\tau}| = n$.

Proof. We do not give a proof for (b) as it follows immediately from Lemma 3.15(a). For $J \in \{A, B\}$ set $U_t^{(J)} := U_t \cap J_n, I_t^{(J)} := I_t \cap J_n$. We prove (a) first. Let $t_0 > 0$ be the first round such that $|I_{t_0}| \geqslant \log\log n$ and set X_t and M as in Lemma 3.16(a); note that Lemma 3.15(a) also gives that $(X_{t_0})_i \geqslant \log\log n/2$ for $i \in \{1,2\}$. Then, for all $t \geqslant t_0$ such that $|I_t| \leqslant n/\log n$, we obtain from Lemma 3.16(a) that $\mathbb{E}_t[X_{t+1}] = (1+o(1))MX_t$ and, in particular, $\mathbb{E}_t[(X_{t+1})_i] = \Theta(|I_t|)$ for $i \in \{1,2\}$. As every component of X_t is self-bounding, Lemma 2.1 applies and we get for $i \in \{1,2\}$

$$\mathbb{P}_t[|(X_{t+1})_i - \mathbb{E}_t[(X_{t+1})_i]| \geqslant \mathbb{E}_t[(X_{t+1})_i]^{2/3}] = O(|I_t|^{-1/3})$$

and by the union bound, provided that $|I_t| \leqslant n/\log n$,

$$\mathbb{P}_t\left[\bigcap_{i\in\{1,2\}} (|(X_{t+1})_i - \mathbb{E}_t[(X_{t+1})_i]| \leqslant \mathbb{E}_t[(X_{t+1})_i]^{2/3})\right] = 1 - O(|I_t|^{-1/3}). \tag{3.30}$$

Using (3.30) we want to find a bound on $|I_{t+1}|$. As long as $|I_t| \leqslant n/\log n$, we get

$$((1 - O(|I_{t_0}|^{-1/3}))M)^{t+1-t_0} X_{t_0} \leqslant X_{t+1} \leqslant ((1 + O(|I_{t_0}|^{-1/3}))M)^{t+1-t_0} X_{t_0}.$$

As seen in Remark 3.1, M has maximal eigenvalue $\lambda_{\max} > 1$, and as M is a positive matrix there is a positive eigenvector v to $\lambda_{\max}$; see [32]. This gives constants $c_1, c_2 > 0$ such that $c_1 v \log\log n \leqslant X_{t_0} \leqslant c_2 v \log\log n$, and for t large enough

$$\frac{c_1}{c_2}((1 - O(|I_{t_0}|^{-1/3}))\lambda_{\max})^{t+1-t_0} X_{t_0} \leqslant X_{t+1} \leqslant \frac{c_2}{c_1}((1 + O(|I_{t_0}|^{-1/3}))\lambda_{\max})^{t+1-t_0} X_{t_0},$$

and therefore

$$|I_{t+1}| \leqslant \frac{c_1}{c_2}((1+o(1))\lambda_{\max})^{t-t_0}|I_{t_0}|,$$

as long as the right-hand side is bounded by $n/\log n$. For all these t we also get

$$|I_{t+1}| \geqslant \frac{c_2}{c_1}((1-o(1))\lambda_{\max})^{t-t_0}|I_{t_0}|.$$

Proceeding as in Lemmas 2.5 and 2.6 and their proofs, where we replace the events

$$|I_t| \geqslant \mathbb{E}_{t-1}[|I_t|] - \mathbb{E}_{t-1}[|I_t|]^{2/3} \quad \text{or} \quad |I_t| \geqslant n/g(n)$$

and

$$||I_t| - \mathbb{E}_{t-1}[|I_t|]| \leqslant \mathbb{E}_{t-1}[|I_t|]^{2/3}$$

with

$$\bigcap_{i\in\{1,2\}} ((X_{t+1})_i \geqslant (1 - \mathbb{E}_t[(X_{t+1})_i]^{-1/3})\mathbb{E}_t[(X_{t+1})_i]) \quad \text{or} \quad |I_t| \geqslant n/\log n$$

and

$$\bigcap_{i\in\{1,2\}} (|(X_{t+1})_i - \mathbb{E}_t[(X_{t+1})_i]| \leqslant \mathbb{E}_t[(X_{t+1})_i]^{2/3})$$

we obtain the statement. Next we show (c). The assumption guarantees that less than $n/\log n$ vertices are informed. Thus $|U_t^{(B)}| \geqslant n/2 - |I_t| \geqslant (1/2 - 1/\log n)n$. We consider a modified dissemination process, where in each round, each uninformed vertex always chooses an informed neighbour (but does not necessarily get informed as the message transmission may fail), and additionally each vertex chooses a neighbour independently and uniformly at random and after this round the chosen vertex is informed with probability q; in other words, we assume that uninformed vertices can inform other vertices. In this modified process the probability of an uninformed vertex $u \in U_t^{(B)}$ staying uninformed after performing one round is given by the product of the probabilities of not being informed by *pull* or via *push* by a vertex in A_n or B_n. Using (3.29) and $(1-1/n)^n = e^{-1+o(1)}$, we get $g(n) = o(1)$ such that

$$\mathbb{P}_t[u \in U_{t+1}^{(B)}] = (1-q)\left(1-\frac{q}{n}\right)^{n/2}\left(1-\frac{q}{(1-\varepsilon)n}\right)^{|N(u)\cap B_n|}$$
$$= (1-q)\exp\left(-q\left(\frac{1}{2}+\frac{1/2-\varepsilon}{1-\varepsilon}\right)+g(n)\right).$$

As we have seen in the proof of Lemma 3.16(b), the probability of being informed by *push&pull* is greater for a vertex in A_n than for a vertex in B_n. Therefore it is sensible to expect that some vertices in B_n will be the last to be informed. Consequently let E_u denote the event that a currently uninformed vertex $u \in U_t^{(B)}$ does not get informed in this modified version within the next

$$\tau := \frac{1}{\log\left((1-q)^{-1}\exp\left(q(1/2+(1/2-\varepsilon)/(1-\varepsilon)-g(n)\right)\right)}\log(n) - h(n)$$

rounds, where $h = o(\log n)$ and $h = \omega(1)$. Therefore we have

$$\mathbb{P}_t[E_u] = \left((1-q)\exp\left(-q\left(\frac{1}{2}+\frac{1/2-\varepsilon}{1-\varepsilon}\right)+g(n)\right)\right)^{\tau} = \frac{1}{n}e^{\omega(1)}.$$

In this modified model the events $\{E_u \mid u \in U_t^{(B)}\}$ satisfy that there is $p = \omega(n^{-1})$ such that

$$\mathbb{P}_t[E_u \mid \{\overline{E_v}\colon v \in U\}] \geqslant p \quad \text{for all } u \in B_n \text{ and } U \subseteq V \setminus \{u\}.$$

This follows immediately by the above calculations. Thus, as $|U_t^{(B)}| = \Theta(n)$,

$$\mathbb{P}_t\left[\bigwedge_{u\in U_t^{(B)}} \overline{E_u}\right] \leqslant \prod_{u\in U_t^{(B)}}(1-p) \leqslant \exp\left(-\sum_{u\in U_t^{(B)}} p\right) = o(1).$$

Finally we show (d). By Lemma 3.16(b), we obtain that for any $\tau \in \mathbb{N}$,

$$\mathbb{E}_t[|U_{t+\tau}|] \leqslant ((1+o(1))e^{-q(1/2+(1/2-\varepsilon)/(1-\varepsilon))}(1-q))^{\tau}|U_t|.$$

Then, for some

$$\tau := \frac{\log(n)}{\log\left((1-q)^{-1}\exp\left(q(1/2+(1/2-\varepsilon)/(1-\varepsilon))\right)\right)} + o(\log n),$$

we obtain that, say, $\mathbb{E}_t[|U_{t+\tau}|] \leqslant |U_t|/n \leqslant 1/\log n$. Thus $\mathbb{P}_t[|U_{t+\tau}| \geqslant 1] \leqslant o(1)$ by Markov's inequality. □

Lemma 3.17 together with Lemma 2.11 gives that

$$T_{pp}(G_n(\varepsilon), q) = \log_\lambda n + \frac{1}{q(1-1.5\varepsilon)/(1-\varepsilon) - \log(1-q)}\log n + o(\log n),$$

where

$$\lambda = 1 + 2q + \Big(2q\Big(\sqrt{(\varepsilon^2/2 - \varepsilon + 1)} - 1\Big) + q\varepsilon\Big)/(2 - 2\varepsilon) > 1 + 2q.$$

To see whether *push&pull* actually slowed down (in terms of order $\log n$) one has to compare the runtime on this sequence of graphs to $c_{pp} \log n$, the runtime on expander sequences. In Figure 1 we can see that it slows down for nearly all values of ε and q in question; however, there are admissible values of ε and q such that the process even speeds up.

References

[1] Acan, H., Collevecchio, A., Mehrabian, A. and Wormald, N. (2017) On the push&pull protocol for rumour spreading. In *Extended Abstracts Summer 2015* (J. Díaz *et al.*, eds), pp. 3–10. Springer.

[2] Angel, O., Mehrabian, A. and Peres, Y. (2017) The string of diamonds is tight for rumor spreading. In *Approximation, Randomization, and Combinatorial Optimization: Algorithms and Techniques (APPROX/RANDOM 2017)*, pp. 26:1–26:9. Schloss Dagstuhl-Leibniz-Zentrum für Informatik.

[3] Boucheron, S., Lugosi, G. and Bousquet, O. (2004) Concentration inequalities. In *Advanced Lectures on Machine Learning*, Vol. 3176 of Lecture Notes in Computer Science, pp. 208–240. Springer.

[4] Boyd, S., Ghosh, A., Prabhakar, B. and Shah, D. (2006) Randomized gossip algorithms. *IEEE/ACM Trans. Inform. Theory* **52** 2508–2530.

[5] Censor-Hillel, K., Haeupler, B., Kelner, J. and Maymounkov, P. (2012) Global computation in a poorly connected world: fast rumor spreading with no dependence on conductance. In *Proceedings of the Forty-Fourth Annual ACM Symposium on Theory of Computing (STOC '12)*, pp. 961–970. ACM.

[6] Chierichetti, F., Giakkoupis, G., Lattanzi, S. and Panconesi, A. (2018) Rumor spreading and conductance. *J. Assoc. Comput. Mach.* **65** 17.

[7] Daum, S., Kuhn, F. and Maus, Y. (2016) Rumor spreading with bounded in-degree. In *Structural Information and Communication Complexity: 23rd International Colloquium (SIROCCO 2016)*, pp. 323–339.

[8] Dellamonica, D., Kohayakawa, Y., Marciniszyn, M. and Steger, A. (2008) On the resilience of long cycles in random graphs. *Electron. J. Combin.* **15** #R321.

[9] Demers, A., Greene, D., Houser, C., Irish, W., Larson, J., Shenker, S., Sturgis, H., Swinehart, D. and Terry, D. (1988) Epidemic algorithms for replicated database maintenance. *ACM SIGOPS Operating Systems Review* **22** 8–32.

[10] Doerr, B., Fouz, M. and Friedrich, T. (2011) Social networks spread rumors in sublogarithmic time. In *Proceedings of the Forty-Third Annual ACM Symposium on Theory of Computing (STOC '11)*, pp. 21–30. ACM.

[11] Doerr, B. and Kostrygin, A. (2017) Randomized rumor spreading revisited. In *44th International Colloquium on Automata, Languages, and Programming (ICALP 2017)*, pp. 138:1–138:14.

[12] Doerr, B. and Künnemann, M. (2014) Tight analysis of randomized rumor spreading in complete graphs. In *Proceedings of the Meeting on Analytic Algorithmics and Combinatorics (ANALCO 2014)*, pp. 82–91. SIAM.

[13] Elsässer, R. and Sauerwald, T. (2009) On the runtime and robustness of randomized broadcasting. *Theoret. Comput. Sci.* **410** 3414–3427.

[14] Feige, U., Peleg, D., Raghavan, P. and Upfal, E. (1990) Randomized broadcast in networks. *Random Struct. Algorithms* **1** 447–460.

[15] Fountoulakis, N., Huber, A. and Panagiotou, K. (2010) Reliable broadcasting in random networks and the effect of density. In *2010 Proceedings IEEE INFOCOM*, pp. 1–9. IEEE.

[16] Fountoulakis, N. and Panagiotou, K. (2010) Rumor spreading on random regular graphs and expanders. In *Approximation, Randomization, and Combinatorial Optimization: Algorithms and Techniques (RANDOM 2010)*, Vol. 6302 of Lecture Notes in Computer Science, pp. 560–573. Springer.

[17] Fountoulakis, N., Panagiotou, K. and Sauerwald, T. (2012) Ultra-fast rumor spreading in social networks. In *Proceedings of the Twenty-Third Annual ACM–SIAM Symposium on Discrete Algorithms (SODA '12)*, pp. 1642–1660. SIAM.

[18] Friedrich, T., Sauerwald, T. and Stauffer, A. (2013) Diameter and broadcast time of random geometric graphs in arbitrary dimensions. *Algorithmica* **67** 65–88.

[19] Frieze, A. M. and Grimmett, G. R. (1985) The shortest-path problem for graphs with random arc-lengths. *Discrete Appl. Math.* **10** 57–77.

[20] Giakkoupis, G. (2011) Tight bounds for rumor spreading in graphs of a given conductance. In *28th International Symposium on Theoretical Aspects of Computer Science (STACS 2011)*, pp. 57–68. Schloss Dagstuhl–Leibniz-Zentrum für Informatik

[21] Giakkoupis, G. (2014) Tight bounds for rumor spreading with vertex expansion. In *Proceedings of the Twenty-Fifth Annual ACM–SIAM Symposium on Discrete Algorithm (SODA 2014)*, pp. 801–815.

[22] Greenberg, S. and Mohri, M. (2014) Tight lower bound on the probability of a binomial exceeding its expectation. *Statist. Probab. Lett.* **86** 91–98.
[23] Haeupler, B. (2015) Simple, fast and deterministic gossip and rumor spreading. *J. Assoc. Comput. Mach.* **62** 47.
[24] Hoffman, A. J. and Wielandt, H. W. (1953) The variation of the spectrum of a normal matrix. *Duke Math. J.* **20** 37–39.
[25] Hoory, S., Linial, N. and Wigderson, A. (2006) Expander graphs and their applications. *Bull. Amer. Math. Soc.* **43** 439–561.
[26] Janson, S. (2018) Tail bounds for sums of geometric and exponential variables. Statistics & Probability Letters, **135** 1–6.
[27] Karp, R. M., Schindelhauer, C., Shenker, S. and Vöcking, B. (2000) Randomized rumor spreading. In *41st Annual Symposium on Foundations of Computer Science (FOCS 2000)*, pp. 565–574.
[28] Panagiotou, K., Pérez-Giménez, X., Sauerwald, T. and Sun, H. (2015) Randomized rumour spreading: the effect of the network topology. *Combin. Probab. Comput.* **24** 457–479.
[29] Panagiotou, K. and Speidel, L. (2017) Asynchronous rumor spreading on random graphs. *Algorithmica* **78** 968–989.
[30] Rödl, V. and Schacht, M. (2010) Regularity lemmas for graphs. In *Fete of Combinatorics and Computer Science*, Vol. 20 of Bolyai Society Mathematical Studies, pp. 287–325. Springer.
[31] Sudakov, B. and Vu, V. H. (2008) Local resilience of graphs. *Random Struct. Algorithms* **33** 409–433.
[32] Wielandt, H. (1950) Unzerlegbare, nicht negative Matrizen. *Math. Z.* **52** 642–648.

Cite this article: Daknama R, Panagiotou K and Reisser S (2021). Robustness of randomized rumour spreading. *Combinatorics, Probability and Computing* **30**, 37–78. https://doi.org/10.1017/S0963548320000310

6 Asymptotics for Pull on the Complete Graph

This chapter is a reprint of

> Panagiotou, K., & Reisser, S. (2021). Asymptotics for pull on the complete graph. Submitted to *Stochastic Processes and their Applications*

A preprint of this version is online at `https://arxiv.org/abs/2110.09044`.

My own contribution. This paper was written together with my supervisor Konstantinos Panagiotou. I contributed substantially to all results. The fundamental ideas were worked out in collaboration with my supervisor. The detailed proofs and their presentations were developed by me. The final version was edited under guidance of Konstantinos Panagiotou.

Asymptotics for Pull on the Complete Graph

Konstantinos Panagiotou Simon Reisser

Ludwig-Maximilians-Universität München
Monday 18^{th} Oktober, 2021

Abstract

We study the randomized rumor spreading algorithm *pull* on complete graphs with n vertices. Starting with one informed vertex and proceeding in rounds, each vertex yet uninformed connects to a neighbor chosen uniformly at random and receives the information, if the vertex it connected to is informed. The goal is to study the number of rounds needed to spread the information to everybody, also known as the *runtime*.

In our main result we provide a description, as n gets large, for the distribution of the runtime that involves a martingale limit. This allows us to establish that in general there is no limiting distribution and that convergence occurs only on suitably chosen subsequences $(n_i)_{i\in\mathbb{N}}$ of $\mathbb{N}$, namely when the fractional part of $(\log_2 n_i + \log_2 \ln n_i)_{i\in\mathbb{N}}$ converges.

1 Introduction

Randomized rumour spreading has applications in replicated databases [7], mobile networks [16], epidemic modelling [1] and even crypto-currency [20]. Given a graph, the algorithm/protocol *pull* works as follows. We start by selecting a vertex and equipping it with some piece of information. Then we proceed in rounds, in which every vertex yet uninformed connects to a neighbor chosen uniformly at random and receives the information, if the vertex it connected to is informed. We will study the number of rounds needed to spread the information to all vertices, also known as *runtime*. For a graph $G = (V, E)$ and any vertex $v \in V$ we denote the (random) runtime of *pull* on G with starting vertex v by $X(G, v)$. The most basic case, and the one studied here, is to set $G = K_n$, the complete graph on n vertices. In that case specifying the initial vertex is not necessary; we therefore just write X_n for the runtime of *pull* on K_n.

Related Work Randomized rumour spreading has been researched intensively since its introduction and popularization in [13, 18]. One direction of research describes the runtime of randomized rumour spreading protocols using only general graph parameters like conductance [2], diameter [11] or expansion [14]. A different direction is to obtain ever more precise bounds on the runtime on specific graph classes [12, 3, 5]. One major step in that direction was achieved in [8], where the authors studied the runtime of *pull* on the complete graph. They showed that

$$\mathbb{E}[X_n] = \log_2 n + \log_2 \ln n + O(1),$$

as well as the related large deviation bound

$$P\big(|X_n - \mathbb{E}[X_n]| \geq r\big) \leq Ae^{-\alpha r} \quad \text{for suitable } A, \alpha > 0 \text{ and all } r \in \mathbb{N}. \tag{1.1}$$

Actually, in [8] qualitatively similar results for several other rumor spreading protocols were shown. In particular they studied the protocol *push*, which differs from *pull* in the way the information is spread from vertex to vertex: in *push*, each *informed* vertex chooses a uniformly random vertex and passes the information forward if the targeted vertex is uninformed.

Regarding *push* we have by now a much more precise picture. In [6] the distribution of the runtime was described for large n. Let X_n^{push} be the runtime of *push* on the complete graph, γ the

Euler-Mascheroni constant, G a Gumble distributed random variable with parameter γ and c a specific 1-periodic function with amplitude about 10^{-9}. Then in [6] it was shown that, as $n \to \infty$,

$$\sup_{k\in\mathbb{N}} \left| P(X_n^{push} \geq k) - P\big(\lceil \log_2 n + \ln n + G + \gamma + c(\log_2 n - \lfloor \log_2 n \rfloor)\rceil \geq k\big) \right| = o(1).$$

This result implies, see [6], that there is a limiting distribution only on suitable subsequences, that is on sequences $(n_i)_{i\in\mathbb{N}}$ such that for some $x, y \in [0, 1)$ $\log_2 n_i - \lfloor \log_2 n_i \rfloor \to x$ and $\ln n_i - \lfloor \ln n_i \rfloor \to y$. Moreover, it shows that the expected runtime of *push* converges only on such sequences as well, with the consequence that

$$\log_2 + \ln n + 1.18242 \leq \mathbb{E}[X_n^{push}] \leq \log_2 + \ln n + 1.18263, \quad n \in \mathbb{N},$$

where both bounds are (essentially) achieved for appropriate subsequences of natural numbers. This improved upon a longish list of previous papers [13, 21, 9, 8], where sharper and sharper results for the runtime of *push* were derived.

All randomized rumor spreading protocols described here proceed in rounds. In [17] minimal path lengths on graphs with random edge weights were studied. If the weights are exponentially distributed, this problem is equivalent to the so-called *asynchronous pull*, where instead of having rounds each uninformed vertex chooses independently neighbours according to a rate-1 Poisson process. In [17] it was shown that, after appropriate normalization, the limiting distribution of the runtime of asynchronous *pull* on complete graphs converges and the limit is the sum of two independent Gumbel distributed random variables. Some more recent results on asynchronous rumor spreading are [15, 19, 22].

Results Our main result describes the distribution of the runtime of *pull* on complete graphs.

Theorem 1.1. *There is a continuous random variable X such that, as $n \to \infty$,*

$$\sup_{k\in\mathbb{N}} \left| P(X_n \geq k) - P(\lceil \log_2 n + \log_2 \ln n + X \rceil \geq k) \right| = o(1).$$

Actually, we can provide some information about X. To this end, let us write I_t for the set of informed vertices at the start of round t. In particular, $|I_0| = 1$. By definition of the protocol, every uninformed vertex becomes informed in round t independently with probability $|I_t|/n$. That is, in distribution

$$|I_{t+1}| = |I_t| + \mathrm{Bin}\big(n - |I_t|, |I_t|/n\big).$$

If $|I_t| = o(n)$, then the binomial distribution is very close to being Poisson, and we thus may approximate $|I_t|$ by the sequence of random variables given by

$$J_0 = 1, \quad \text{and} \quad J_{t+1} = J_t + \mathrm{Po}(J_t), \quad t \in \mathbb{N}_0.$$

This sequence doubles every round in expectation, $\mathbb{E}[J_t] = 2^t$. Moreover, it is fairly easy to establish that $(H_t)_{t\in\mathbb{N}_0}$ with $H_t = 2^{-t} J_t$ is a martingale and uniformly integrable. Thus, the Martingale Convergence Theorem guarantees the existence of a random variable H such that H_t converges almost surely to H. In essence, $2^{-t}|I_t|$ is 'close' to H for large t (and n); we formalize this statement in Lemma 2.1 below. The random variable X is then given as $X = -\log_2 H$.

From this construction of H we may obtain further information about its distribution. For example, the characteristic function φ of H has the property $\varphi = \lim_{t\to\infty} \varphi_t$, where φ_t is the characteristic function of H_t, by Levy's Continuity Theorem. Using this we establish in Section 2.1 that H is continuous and almost surely positive, two main ingredients in the proof of Theorem 1.1. However, finding more properties of H, like a handy expression for its density or expressions for its moments, turned out to be a tough challenge that we leave as an open problem.

Let us denote with $(X + x)|_{\mathbb{Z}}$ the distribution of X translated by x and restricted to integers only, that is, $(X + x)|_{\mathbb{Z}}$ is the random variable with domain $\mathbb{Z}$ and distribution

$$P\big((X + x)|_{\mathbb{Z}} \leq k\big) := P(X \leq k - x), \quad k \in \mathbb{Z}.$$

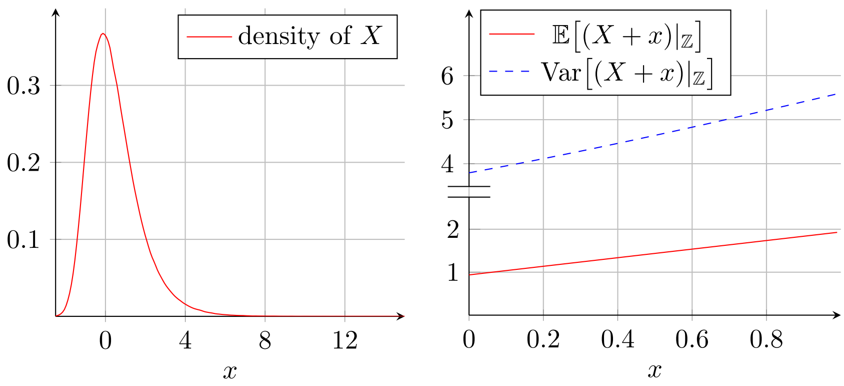

Figure 1: The left plot shows an estimate of the density of the random variable X from Theorem 1.1. The right plot shows, as a function of $x \in [0,1]$, the estimated expectation and variance of the random variable $(X+x)|_{\mathbb{Z}}$ defined in Corollary 1.2.

With this definition at hand and by choosing a suitable subsequence we can derive a limiting distribution from Theorem 1.1.

Corollary 1.2. *Let $x \in [0,1)$ and let n_i be a strictly increasing sequence of natural numbers such that $\log_2 n_i + \log_2 \ln n_i - \lfloor \log_2 n_i + \log_2 \ln n_i \rfloor \to x$. Then, as $i \to \infty$, in distribution*

$$X_{n_i} - \lfloor \log_2 n_i + \log_2 \ln n_i \rfloor \to (X+x)|_{\mathbb{Z}} \,.$$

This corollary warrants some further remarks. First of all, it is not immediately clear that a sequence with the required properties exists, at least it was not to us. Luckily it requires only moderate effort to find a suitable one. For example, we may choose

$$n_i = \big\lfloor \exp(W(2^{i+x})) \big\rfloor, \quad i \in \mathbb{N},$$

where W is the principal branch of the Lambert W function (or product logarithm). A key property of W is that $W(x)e^{W(x)} = x$ for all $x > -1/e$. As $W(z) = (1+o(1)) \ln z$ for large z, see for example [4], the sequence $(n_i)_{i\in\mathbb{N}}$ is strictly increasing. Thus, as i gets large,

$$2^{\log_2 n_i + \log_2 \ln n_i} = n_i \ln n_i = \big(1+o(1)\big) \cdot \exp\big(W(2^{i+x})\big) W(2^{i+x}) = 2^{i+x+o(1)}.$$

Secondly, we can actually say more. Large deviation bounds for *pull*, see (1.1), yield that $(X_n)^k$ is absolutely integrable for all $k \in \mathbb{N}$ and thus convergence in distribution also implies convergence of all moments. In particular, for sequences as in Corollary 1.2

$$\mathbb{E}\left[\left(X_{n_i} - \lfloor \log_2 n_i + \log_2 \ln n_i \rfloor\right)^k\right] \to \mathbb{E}\left[\left((X+x)|_{\mathbb{Z}}\,\right)^k\right] \qquad \forall\, k \in \mathbb{N}.$$

However, as already mentioned, extracting more information from this statement requires more detailed knowledge about the moments/the distribution of X that we do not have. On a positive side, we can provide some preliminary numerical results, see Fig. 1. To get these numbers, we have drawn 10^6 instances of the random variable $-\log_2 H_{28}$ as a substitute for the random variable $X = -\log_2 H$. To approximate the density we used the `gaussian_kde` function of Pythons Scipy package. To estimate first and second moments of $(X+x)|_{\mathbb{Z}}$ we used the formulas

$$\mathbb{E}\big[(X+x)|_{\mathbb{Z}}\big] = \sum_{k\ge1} \Big(P(X \ge k-x-1) - P(X \le -k-x)\Big)$$

and

$$\mathbb{E}\big[\big((X+x)|_{\mathbb{Z}}\big)^2 + (X+x)|_{\mathbb{Z}}\big] = 2 \cdot \sum_{k\ge1} k\Big(P(X \ge k-x-1) - P(X \le -k-x)\Big),$$

where we again substituted X by $-\log_2 H_{28}$.

Outline The paper is structured as follows. In the next section we give the proof of the main results, which is based on three key lemmas characterising the different phases of *pull*. At first, as long as less than $n^{1/3}$ vertices are informed, *pull* is best described by a branching process, which, suitably normalized, has limiting distribution H, see Lemma 2.1. After that, the protocol follows essentially a deterministic recurrence relation as described in Lemma 2.2. Once there are only $o(\sqrt{n})$ uninformed vertices remaining the behaviour changes once more, in that all these vertices will be informed in one additional round, see Lemma 2.3. The proof of Theorem 1.1, based on these lemmas as outlined, is given in Subsection 2.1. After that, we give the short proof for Corollary 1.2 in Subsection 2.2. The proofs of Lemmas 2.1-2.3 can be found in Subsections 2.3-2.5.

2 Proofs

2.1 Proof of Theorem 1.1

Our first auxiliary lemma establishes that initially – as long as there are not too many informed vertices – the number of informed vertices essentially doubles in each round and the deviation from perfect doubling can be described in terms of a non-trivial random variable.

Lemma 2.1. *There is a continuous and almost surely positive random variable H such that for all $\varepsilon > 0$ there are constants $n_0, t_0 \in \mathbb{N}$ such that for all $n \geq n_0$ and $t_0 \leq t \leq \log_2(n^{1/3})$*

$$\sup_{x\in\mathbb{R}} \left| P\left(2^{-t}|I_t| \geq x\right) - P(H \geq x) \right| \leq \varepsilon.$$

The proof is in Section 2.3. From now on we fix some $\varepsilon > 0$ and set for the remainder

$$t_1 := \lfloor \log_2(n^{1/3}) \rfloor,$$

where $n \geq n_0$ is given by the previous lemma. For rounds $t \geq t_1$ it turns out that the behaviour of *pull* can be best described by a (deterministic) recurrence relation. We have roughly $2^{t_1} \approx n^{1/3}$ informed vertices, enough so that it is reasonable to assume that the number of newly informed vertices in the following rounds is strongly concentrated around its expectation.

Denote by U_t the set of uniformed vertices at the start of round t. Then it is easy to see that $\mathbb{E}[|U_{t+1}| \mid I_t] = (|U_t|/n)^2 n$, see also Lem. 2.14 below, and thus we expect $|U_{t_1+t}|$, given $|U_{t_1}|$, to be close to $(|U_{t_1}|/n)^{2^t} n$. The only thing that we have to take care of is that (small) deviations from the expectation are not blown out of proportions when considering multiple rounds. The next lemma, that we prove in Subsection 2.4, does exactly that. With high probability, or abbreviated as whp, means with probability tending to 1 as n tends to infinity.

Lemma 2.2. *With high probability*

$$\bigcap_{t\geq 0} \left\{ \left| |U_{t_1+t}| - \left(|U_{t_1}|/n\right)^{2^t} n \right| \leq |U_{t_1+t}| \cdot n^{-1/50} + n^{1/4} \right\}.$$

This lemma will enable us to track the process all the way until there are fewer than $\sqrt{n}$ uniformed vertices remaining. Indeed, as we will argue shortly, $\sqrt{n}$ is an important threshold in the following sense. On the one hand, if there are substantially more than $\sqrt{n}$ uninformed vertices, very likely the process will not terminate in the next round. In contrast, if there are much less than $\sqrt{n}$ uninformed vertices, the process will likely terminate in the next round. Furthermore, we will see that it is very unlikely that $|U_t| = \Theta(\sqrt{n})$ for some t, where the process terminates only with constant probability. The next lemma summarizes our findings.

Lemma 2.3. *Let $T = \min\left\{t \in \mathbb{N} : |U_t| < \sqrt{n}\right\}$, then with high probability*

$$|U_{T-1}| = \omega\left(\sqrt{n}\right), \quad |U_T| = o\left(\sqrt{n}\right), \quad |U_T| > 0 \quad \textit{and} \quad |U_{T+1}| = 0.$$

The last statement implies immediately that whp $X_n = T+1$ and it thus provides a handy way to compute X_n. To that end we utilize Lemma 2.2 that guarantees whp for all $t \in \mathbb{N}$

$$|U_{t_1+t}| = \big(1+o(1)\big)\big(|U_{t_1}|/n\big)^{2^t} n + O\big(n^{1/4}\big).$$

Together with Lemma 2.3 this implies whp

$$\big(|U_{t_1}|/n\big)^{2^{T-t_1-1}} n = \omega\big(\sqrt{n}\big) \quad \text{and} \quad \big(|U_{t_1}|/n\big)^{2^{T-t_1}} n = o\big(\sqrt{n}\big)$$

and therefore also whp

$$T = \min\Big\{t \in \mathbb{N} : \big(|U_{t_1}|/n\big)^{2^{t-t_1}} n < \sqrt{n}\Big\} = \min\Big\{t \in \mathbb{N} : (1-|I_{t_1}|/n)^{2^{t-t_1}} n < \sqrt{n}\Big\}.$$

Let T' be the real number such that $\big(1-|I_{t_1}|/n\big)^{2^{T'}} n = \sqrt{n}$. Then $T = \lfloor T'+t_1+1 \rfloor$ and it is straightforward to verify that

$$T' = \log_2 n - \log_2 |I_{t_1}| + \log_2 \ln n - 1 + o(1).$$

Therefore, as Lemma 2.3 yields whp $X_n = T+1 = \lfloor T'+t_1+2 \rfloor$,

$$\sup_{k\in\mathbb{N}} \Big| P(X_n \geq k) - P\Big(\big\lfloor \log_2 n + \log_2 \ln n - \log_2\big(2^{-t_1}|I_{t_1}|\big) + 1 + o(1)\big\rfloor \geq k\Big)\Big| = o(1).$$

In Lemma 2.1 we showed that $2^{-t_1}|I_{t_1}|$ converges in distribution to a random variable H; since H is continuous so is $X = -\log_2 H$ and the claim in Theorem 1.1 follows readily.

2.2 Proof of Corollary 1.2

Theorem 1.1 states that there is a continuous random variable X such that for all $x \in [0,1)$ and strictly increasing sequences n_i such that $\log_2 n_i + \log_2 \ln n_i - \lfloor \log_2 n_i + \log_2 \ln n_i \rfloor \to x$

$$\sup_{k\in\mathbb{N}} \Big| P\big(X_{n_i} \geq k\big) - P\big(\log_2 n_i + \log_2 \ln n_i + X + 1 \geq k\big)\Big| = o(1).$$

Setting $\{y\} = y - \lfloor y \rfloor$ for all $y \in \mathbb{R}$ and substituting $k = \lfloor \log_2 n_i + \log_2 \ln n_i \rfloor + t + 1$ we obtain

$$\sup_{t\in\mathbb{Z}} \Big| P\big(X_{n_i} \geq \lfloor \log_2 n_i + \log_2 \ln n_i \rfloor + t + 1\big) - P\big(\{\log_2 n_i + \log_2 \ln n_i\} + X \geq t\big)\Big| = o(1).$$

Thus X being a continuous random variable

$$\sup_{t\in\mathbb{Z}} \big| P\big(X_{n_i} \geq \lfloor \log_2 n_i + \log_2 \ln n_i \rfloor + 1 + t\big) - P(x + X \geq t)\big| = o(1).$$

Thus $P(X_{n_i} - \lfloor \log_2 n_i + \log_2 \ln n_i \rfloor \leq t) \overset{i\to\infty}{\longrightarrow} P(X \leq t - x)$, as claimed.

2.3 Proof of Lemma 2.1

To prove Lemma 2.1 first recall that

$$J_0 = 1 \quad \text{and} \quad J_{t+1} = J_t + \mathrm{Po}(J_t),\ H_t = 2^{-t} J_t,\ t \in \mathbb{N}_0.$$

We show three claims in order to prove Lemma 2.1, namely that $|I_t|$ is close to J_t, then that $(H_t)_{t\in\mathbb{N}_0}$ is a martingale that converges (to H) and finally that the limit is absolutely continuous.

$|I_t|$ and J_t are close. We begin with a simple lemma that determines the first and second moment of J_t.

Lemma 2.4. *For all $t \in \mathbb{N}_0$ and J_t as defined above*

$$\mathbb{E}[J_t] = 2^t \quad \text{and} \quad \mathbb{E}[J_t^2] = 2^{t-1}(3 \cdot 2^t - 1).$$

Proof. We compute both moments inductively, starting with the base case

$$\mathbb{E}[J_0] = 1 \quad \text{and} \quad \mathbb{E}[J_0^2] = 1.$$

Moreover, using the tower property of the expectation and that $\mathbb{E}[\mathrm{Po}(\lambda)] = \lambda$ for any $\lambda > 0$ we obtain by induction

$$\mathbb{E}[J_{t+1}] = \mathbb{E}[J_t + \mathrm{Po}(J_t)] = \mathbb{E}\Big[\mathbb{E}\big[J_t + \mathrm{Po}(J_t) \bigm| J_t\big]\Big] = 2 \cdot \mathbb{E}[J_t] = 2^{t+1}.$$

We compute the second moment similarly. Since $\mathbb{E}[\mathrm{Po}(\lambda)^2] = \lambda + \lambda^2$ we obtain that

$$\begin{aligned}\mathbb{E}[J_{t+1}^2] &= \mathbb{E}\left[\big(J_t + \mathrm{Po}(J_t)\big)^2\right] = \mathbb{E}\left[J_t^2 + 2 \cdot J_t \mathrm{Po}(J_t) + \mathrm{Po}(J_t)^2\right] = 4 \cdot \mathbb{E}[J_t^2] + \mathbb{E}[J_t] \\ &= 4 \cdot 2^{t-1}(3 \cdot 2^t - 1) + 2^t = 2^t(3 \cdot 2^{t+1} - 1).\end{aligned}$$

□

The next (well-known) statement bounds the distance between two Poisson distributed random variables and furthermore quantifies the distance in the Poisson limit theorem. Recall that the *total variation distance* for two integer valued random variables X, Y can be defined as

$$\mathrm{d}(X,Y) := \frac{1}{2} \sum_{k \in \mathbb{Z}} \big|P(X = k) - P(Y = k)\big|. \tag{2.1}$$

Lemma 2.5 ([23], Eq. 3.6 and Thm. 4.1)**.**

a) Let $\lambda_1, \lambda_2 \in \mathbb{N}$ and $X \sim \mathrm{Po}(\lambda_1)$ and $Y \sim \mathrm{Po}(\lambda_2)$ be independent Poisson-distributed random variables. Then $\mathrm{d}(X,Y) \leq |\lambda_1 - \lambda_2|$.

b) Let $X \sim \mathrm{Bin}(n,p)$ and $Y \sim \mathrm{Po}(np)$. Then $\mathrm{d}(X,Y) \leq np^2$.

With these ingredients at hand we can give a bound on the distance of $|I_t|$ and J_t that is quite strong as long as t is not too large.

Lemma 2.6. *For all $t \in \mathbb{N}$*

$$\mathrm{d}_t := \mathrm{d}\big(|I_t|, J_t\big) \leq 2 \cdot 4^t/n.$$

Proof. Note that it suffices to consider only the case $t \leq \log_4 n$, as otherwise the claimed bound is greater than one and consequently trivially true. There are $|I_t|$ informed vertices in round t. Then the probability of any vertex $v \in U_t$ to be informed in that round is $|I_t|/n$ and furthermore it is independent of all other uninformed vertices, that is, the number of newly informed vertices is binomially distributed with $|U_t|$ tries and success probability $|I_t|/n$. Thus, in distribution,

$$|I_{t+1}| = |I_t| + \mathrm{Bin}\big(|U_t|, |I_t|/n\big), \tag{2.2}$$

an equation that we have already encountered in the introduction. We prove the statement of the lemma by induction over t. The base case is obvious as $|I_0| = 1 = J_0$. For the induction step, we use (2.1) together with $|I_t|$ and J_t only taking values on the positive integers, to get

$$\begin{aligned}2 \cdot \mathrm{d}_{t+1} &= \sum_{k \geq 1} \big|P\big(|I_{t+1}| = k\big) - P(J_{t+1} = k)\big| \\ &= \sum_{k=1}^{n} \left| \sum_{\ell=1}^{k} P\big(|I_{t+1}| = k \bigm| |I_t| = \ell\big) P\big(|I_t| = \ell\big) - P\big(J_{t+1} = k \bigm| J_t = \ell\big) P(J_t = \ell) \right| + P(J_{t+1} > n).\end{aligned}$$

To simplify this expression we consider the auxiliary calculation

$$\begin{aligned}
&\sum_{k=1}^{n}\Big|\sum_{\ell=1}^{k} P\big(|I_{t+1}|=k \mid |I_t|=\ell\big)\big(P\big(|I_t|=\ell\big)-P(J_t=\ell)\big)\Big|\\
\leq&\sum_{\ell=1}^{n}\Big|P\big(|I_t|=\ell\big)-P(J_t=\ell)\Big|\sum_{k=\ell}^{n}P\big(|I_{t+1}|=k \mid |I_t|=\ell\big)\\
\leq&\sum_{\ell=1}^{n}\Big|P\big(|I_t|=\ell\big)-P(J_t=\ell)\Big|\leq 2\cdot \mathrm{d}_t.
\end{aligned}$$

In order to obtain a bound for the tail probability of J_{t+1} we use Lemma 2.4 as well as the assumption $t \leq \log_4 n$ so that by Chebyshev's inequality and plenty of room to spare

$$P(J_{t+1}>n)\leq P\big(|J_{t+1}-\mathbb{E}[J_{t+1}]|>n-\mathbb{E}[J_{t+1}]\big)\leq \frac{\mathrm{Var}[J_{t+1}]}{(n-\mathbb{E}[J_{t+1}])^2}\leq 4^t/n.$$

Applying these bounds to d_{t+1} we get

$$2\cdot \mathrm{d}_{t+1}\leq \sum_{k=1}^{n}\sum_{\ell=1}^{k}P(J_t=\ell)\Big|P\big(|I_{t+1}|=k \mid |I_t|=\ell\big)-P\big(J_{t+1}=k \mid J_t=\ell\big)\Big|+2\cdot \mathrm{d}_t+4^t/n.$$

Next we plug in the distributions for $|I_{t+1}|-|I_t|$ (binomial) and J_{t+1} (Poisson) to get

$$2\cdot \mathrm{d}_{t+1}\leq \sum_{k=1}^{n}\sum_{\ell=1}^{k}P(J_t=\ell)\Big|P\big(\mathrm{Bin}(n-\ell,\ell/n)=k-\ell\big)-P\big(\mathrm{Po}(\ell)=k-\ell\big)\Big|+2\cdot \mathrm{d}_t+4^t/n$$

and shifting indices yields

$$\begin{aligned}
2\cdot \mathrm{d}_{t+1}\leq&\sum_{\ell=1}^{n}P(J_t=\ell)\sum_{k=1}^{n}\Big|\big(P(\mathrm{Bin}(n-\ell,\ell/n)=k\big)-P\big(\mathrm{Po}(\ell)=k\big)\Big|+2\cdot \mathrm{d}_t+4^t/n\\
\leq&\sum_{\ell=1}^{n}P(J_t=\ell)\cdot 2\cdot \mathrm{d}\big(\mathrm{Bin}(n-\ell,\ell/n),\mathrm{Po}(\ell)\big)+2\cdot \mathrm{d}_t+4^t/n.
\end{aligned}$$

As d is a metric we can use the triangle inequality and with Lemma 2.5 we get for all $0\leq \ell\leq n$

$$\begin{aligned}
\mathrm{d}\big(\mathrm{Bin}(n-\ell,\ell/n),\mathrm{Po}(\ell)\big)&\leq \mathrm{d}\big(\mathrm{Bin}(n-\ell,\ell/n),\mathrm{Po}((n-\ell)\ell/n)\big)+\mathrm{d}\big(\mathrm{Po}((n-\ell)\ell/n),\mathrm{Po}(\ell)\big)\\
&\leq (n-\ell)(\ell/n)^2+\big|(n-\ell)\ell/n-\ell\big|,
\end{aligned}$$

which is at most $2\ell^2/n$. By plugging this into the previous inequality we get

$$\mathrm{d}_{t+1}\leq \sum_{\ell=1}^{n}P(J_t=\ell)\frac{2\ell^2}{n}+\mathrm{d}_t+4^t/n\leq 2\cdot\frac{\mathbb{E}[J_t^2]}{n}+\mathrm{d}_t+4^t/n.$$

Lemma 2.4 determines the second moment of J_t. By using the induction hypothesis we conclude

$$\mathrm{d}_{t+1}\leq 3\cdot 4^t/n+2\cdot 4^t/n+4^t/n\leq 2\cdot 4^{t+1}/n.$$

□

$(H_t)_{t\in\mathbb{N}_0}$ is a martingale that converges to H. Next we show that the sequence $(H_t)_{t\in\mathbb{N}_0}$ is a martingale and converges almost surely and in $\mathcal{L}^2$ to the random variable H.

Lemma 2.7. *There is a random variable H such that $H_t\to H$ almost surely and in $\mathcal{L}^2$. Furthermore H has mean 1 and variance* $1/2$.

Proof. First we show that H_t is a martingale. Let $\mathcal{F}_t$ be the filtration induced by the random variables J_t, then

$$\mathbb{E}\big[H_{t+1} \bigm| \mathcal{F}_t\big] = \mathbb{E}\big[2^{-t-1}\big(J_t + \mathrm{Po}(J_t)\big) \bigm| \mathcal{F}_t\big] = 2^{-t}J_t = H_t.$$

Next, we show that $H_t \in \mathcal{L}^2$. Therefore we compute

$$\mathbb{E}\big[|H_t|^2\big] = 2^{-2t}\mathbb{E}\big[J_t^2\big] = 2^{-2t} \cdot 2^t(3 \cdot 2^{t+1} - 1) \le 6$$

using Lemma 2.4 and consequently

$$\sup_{t\in\mathbb{N}_0} \mathbb{E}\big[|H_t|^2\big] < \infty. \tag{2.3}$$

This yields the integrability of H_t and as H_t is obviously measurable with respect to $\mathcal{F}_t$ we conclude that it is indeed a martingale. Thus (2.3) and $\mathcal{L}^p$ convergence of martingales implies the first claim, see e.g. [10, Thm. 4.4.6]. The values for the expectation and the variance follow immediately from Lemma 2.4 by scaling with 2^{-t} and 2^{-2t} respectively and then taking the limit. $\square$

An even stronger version of this lemma could be shown, i.e., the martingale converges in $\mathcal{L}^p$ for all $p > 1$. In any case, the version stated here suffices for our purposes.

In order to show the properties of H claimed in Lemma 2.1 we need to describe its characteristic function. The next lemma does exactly that, but we need a definition first. Let

$$h(x) = h^{(1)}(x) = xe^{x-1}, \quad h^{(t+1)} = h^{(t)} \circ h,\ t \in \mathbb{N}. \tag{2.4}$$

This function is not new in the context of rumor spreading, it plays an important role in the closely related context of [6], where it describes the evolution of the number of uninformed vertices of *push* on complete graphs.

Lemma 2.8. *The characteristic functions φ_t of H_t and φ of H satisfy*

$$\varphi_t(x) = h^{(t)}\Big(e^{ix2^{-t}}\Big) \quad \textit{and} \quad \varphi(x) = \lim_{t\to\infty} \varphi_t(x).$$

Proof. To prove the claim we first compute the probability generating function $\tilde{J}_t(x)$ of J_t. Note that $\tilde{J}_0(x) = x$, as $P(J_0 = 1) = 1$. For $t \in \mathbb{N}_0$ and $|x| \le 1$ we get

$$\begin{aligned}
\tilde{J}_{t+1}(x) = \sum_{k\ge0} P(J_{t+1} = k)x^k &= \sum_{k\ge0} P\big(J_t + \mathrm{Po}(J_t) = k\big)x^k \\
&= \sum_{k\ge0} \sum_{0\le\ell\le k} P\big(\mathrm{Po}(\ell) = k - \ell\big)P(J_t = \ell)x^k \\
&= \sum_{\ell\ge0} P(J_t = \ell) \sum_{k\ge0} \frac{\ell^k}{k!}e^{-\ell}x^{k+\ell} \\
&= \sum_{\ell\ge0} P(J_t = \ell)e^{-\ell+x\ell}x^\ell \\
&= \sum_{\ell\ge0} P(J_t = \ell)\big(xe^{x-1}\big)^\ell = \tilde{J}_t\big(xe^{x-1}\big) = \big(\tilde{J}_t \circ h\big)(x).
\end{aligned}$$

Thus J_t has characteristic function $x \mapsto h^{(t)}\big(e^{ix}\big)$ and as $H_t = 2^{-t}J_t$ we immediately obtain that H_t has characteristic function $x \mapsto h^{(t)}\big(e^{ix2^{-t}}\big)$. With Levy's continuity theorem we infer that the characteristic function of H_t converges to the characteristic function of H, as Lemma 2.7 guarantees that H_t converges to H almost surely and therefore also in distribution. $\square$

Properties of H. In the last part of this section we show that H is absolutely continuous and almost surely positive. To show absolute continuity, we will argue that the characteristic function of H is integrable. To achieve this we first find a recurrence relation for real and imaginary parts of φ_t, the characteristic function of H_t. We then use this description to find a second order approximation of φ_t that eventually allows us to uniformly bound the absolute value of φ_t by an integrable function.

The mapping that we will use to describe the real and imaginary parts of φ_t is given by

$$F = F^{(1)} : \mathbb{R}^2 \to \mathbb{R}^2, F\left(\begin{pmatrix} R \\ I \end{pmatrix}\right) = e^{-1+R}\begin{pmatrix} \cos I & -\sin I \\ \sin I & \cos I \end{pmatrix}\begin{pmatrix} R \\ I \end{pmatrix}, \text{ and } F^{(t+1)} = F^{(t)} \circ F, \ \ t \in \mathbb{N}.$$

Moreover, we set $F^{(0)}$ to be the identity on $\mathbb{R}^2$.

Lemma 2.9. *Let φ_t be the characteristic function of H_t. Set $I_t(x) = \mathrm{Im}\big(\varphi_t(x)\big)$ (the imaginary part), $R_t(x) = \mathrm{Re}\big(\varphi_t(x)\big)$ (the real part) and $a_t(x) = |\varphi_t(x)|$. Then for all $t \in \mathbb{N}_0$*

$$\begin{pmatrix} R_t(x) \\ I_t(x) \end{pmatrix} = F^{(t)}\left(\begin{pmatrix} \cos(x2^{-t}) \\ \sin(x2^{-t}) \end{pmatrix}\right).$$

and

$$a_{t+1}(x) = a_t(x/2)\exp\big(-1 + R_t(x/2)\big).$$

Proof. Using Lemma 2.8 we obtain for $t \in \mathbb{N}_0$

$$\varphi_{t+1}(x) = h^{(t+1)}\left(e^{ix2^{-t-1}}\right) = h\left(h^{(t)}\left(e^{i(x/2)2^{-t}}\right)\right) = h\big(\varphi_t(x/2)\big). \tag{2.5}$$

We continue with a simple observation. For two complex numbers z, w the imaginary and real parts of their product satisfy

$$\mathrm{Re}(z\cdot w) = \mathrm{Re}(z)\mathrm{Re}(w) - \mathrm{Im}(z)\mathrm{Im}(w) \quad \text{and} \quad \mathrm{Im}(z\cdot w) = \mathrm{Re}(z)\mathrm{Im}(w) + \mathrm{Im}(z)\mathrm{Re}(w).$$

Using this observation and (2.5) we obtain

$$\begin{aligned} I_{t+1}(x) &= \mathrm{Im}\big(\varphi_{t+1}(x)\big) = \mathrm{Im}\Big(\varphi_t(x/2)\exp\big(-1+\varphi_t(x/2)\big)\Big) \\ &= R_t(x/2)\mathrm{Im}\Big(\exp\big(-1+\varphi_t(x/2)\big)\Big) + I_t(x/2)\mathrm{Re}\Big(\exp\big(-1+\varphi_t(x/2)\big)\Big) \\ &= \Big(R_t(x/2)\sin\big(I_t(x/2)\big) + I_t(x/2)\cos\big(I_t(x/2)\big)\Big)\exp\big(-1+R_t(x/2)\big) \end{aligned}$$

and similarly for the real part

$$\begin{aligned} R_{t+1}(x) &= \mathrm{Re}\big(\varphi_{t+1}(x)\big) = \mathrm{Re}\Big(\varphi_t(x/2)\exp\big(-1+\varphi_t(x/2)\big)\Big) \\ &= R_t(x/2)\mathrm{Re}\Big(\exp\big(-1+\varphi_t(x/2)\big)\Big) - I_t(x/2)\mathrm{Im}\Big(\exp\big(-1+\varphi_t(x/2)\big)\Big) \\ &= \Big(R_t(x/2)\cos\big(I_t(x/2)\big) - I_t(x/2)\sin\big(I_t(x/2)\big)\Big)\exp\big(-1+R_t(x/2)\big). \end{aligned}$$

Applying these two equations repeatedly and remembering that $\varphi_0(x) = e^{ix}$ and therefore $R_0(x) = \cos(x)$ as well as $I_0(x) = \sin(x)$ implies the first claim. To show the second claim in the lemma (about a_{t+1}) we use again (2.5) and $|e^z| = e^{\mathrm{Re}(z)}$ for all $z \in \mathbb{C}$

$$a_{t+1}(x) = \big|\varphi_t(x/2)\cdot\exp\big(\varphi(x/2)-1\big)\big| = a_t(x/2)\cdot\exp\big(-1+R_t(x/2)\big).$$

□

With that recursive description at hand we can derive a (first) handy approximation for φ_t.

Lemma 2.10. *Let $t \in \mathbb{N}_0$ and $x \in \mathbb{R}$. For all $0 \le j \le \max\{j \in \mathbb{N}_0 : |x2^{-t+j}| \le 1/16\}$*

$$\left| F^{(j)}\left(\begin{pmatrix} \cos(x2^{-t}) \\ \sin(x2^{-t}) \end{pmatrix}\right) - \begin{pmatrix} 1 - x^2 2^{-2t+j-2}(3 \cdot 2^j + 1) \\ x2^{-t+j} \end{pmatrix} \right| \le \begin{pmatrix} |x^3| 2^{3(-t+j)} \\ |x^2| 2^{2(-t+j)} \end{pmatrix}.$$

Proof. If $\{j \in \mathbb{N}_0 : |x2^{-t+j}| \le 1/16\} = \emptyset$ we have nothing to show, thus we assume that $j_{\max} := \max\{j \in \mathbb{N}_0 : |x2^{-t+j}| \le 1/16\} \ge 0$. We will show the claim by induction over all $0 \le j \le j_{\max}$. Very important ingredients in the forthcoming arguments are the following estimates for smallish x that are rather easy to show:

$$\left| \cos(x) - \left(1 - \frac{x^2}{2}\right) \right| \le \frac{x^4}{24} \quad \text{for all} \quad |x| \le 7 \tag{2.6}$$

and

$$\big| \sin(x) - x \big| \le \frac{x^3}{5} \quad \text{for all} \quad |x| \le 2. \tag{2.7}$$

These estimates yield (with quite some room to spare) immediately the induction start ($j = 0$), as by convention $F^{(0)}$ is the identity on $\mathbb{R}^2$. We proceed with the induction step. For the following computations abbreviate

$$\alpha_j = 2^{-2t+j-2}(3 \cdot 2^j + 1), \quad \beta_j = 2^{-t+j} \quad \text{and} \quad \Delta_j = \begin{pmatrix} \Delta_{j,1} \\ \Delta_{j,2} \end{pmatrix} = \begin{pmatrix} |x^3| 2^{-3t+3j} \\ |x^2| 2^{-2t+2j} \end{pmatrix}.$$

In the remainder of this proof we use the following notation. For real numbers a, b we write $a \pm b$ to denote *some* real number c that satisfies $|a - c| \le b$. In particular, if we apply a function, e.g., F, to $a \pm b$ we understand that as F applied to some number c in the designated interval. This notation is useful as we are only interested in upper and lower bounds on $F(a \pm b)$ that we can deduce from a and b only.

Let $0 \le j \le j_{\max} - 1$. By applying the induction hypothesis we get

$$F^{(j+1)}\left(\begin{pmatrix} \cos(x2^{-t}) \\ \sin(x2^{-t}) \end{pmatrix}\right) = F\left(\begin{pmatrix} 1 - x^2\alpha_j \\ x\beta_j \end{pmatrix} \pm \Delta_j\right) =: \begin{pmatrix} F_1 \\ F_2 \end{pmatrix}. \tag{2.8}$$

Using the definition of F we obtain for the first component

$$F_1 = (F_{11} - F_{12})F_{13},$$

where we abbreviated

$$\begin{aligned} F_{11} &= \big(1 - x^2\alpha_j \pm \Delta_{j,1}\big) \cos\big(x\beta_j \pm \Delta_{j,2}\big) \\ F_{12} &= \big(x\beta_j \pm \Delta_{j,2}\big) \sin\Big(x\beta_j \pm \Delta_{j,2}\Big) \\ F_{13} &= \exp\big(- x^2\alpha_j \pm \Delta_{j,1}\big). \end{aligned}$$

To study these expressions we look at three recurring components first. Note that, as $(\Delta_{j,2})^2 \le \Delta_{j,1}/16$ by our assumption on $j \le j_{\max}$,

$$\big(x\beta_j \pm \Delta_{j,2}\big)^2 = (x\beta_j)^2 \pm 2|x|\beta_j\Delta_{j,2} \pm (\Delta_{j,2})^2 = (x\beta_j)^2 \pm \frac{33}{16}\Delta_{j,1}. \tag{2.9}$$

Furthermore, using again $|x|2^{-t+j} \le 1/16$ guaranteed by $j \le j_{\max}$,

$$\big|\big(x\beta_j \pm \Delta_{j,2}\big)^3\big| \le \sum_{i=0}^{3} \binom{3}{i} (|x|\beta_j)^i (\Delta_{j,2})^{3-i} \le \Delta_{j,1}\left(1 + \frac{3}{16} + \frac{3}{16^2} + \frac{1}{16^3}\right) \le \frac{6}{5}\Delta_{j,1} \tag{2.10}$$

and similarly

$$\left|\left(x\beta_j \pm \Delta_{j,2}\right)^4\right| \leq \sum_{i=0}^{4} \binom{4}{i} (|x|\beta_j)^i (\Delta_{j,2})^{4-i} \leq \frac{4}{50}\Delta_{j,1}. \tag{2.11}$$

Lastly, we use again $|x|2^{-t+j} \leq 1/16$ and $x^4 a_j^2 \leq \Delta_{j,1}/16$ to bound

$$\left|\left(x^2\alpha_j \pm \Delta_{j,1}\right)^2\right| \leq |x|^4\alpha_j^2 + |x|^2\alpha_j\Delta_{j,1} + \Delta_{j,1}^2 \leq \Delta_{j,1}\left(\frac{1}{16} + \frac{1}{16^2} + \frac{1}{16^3}\right) \leq \frac{1}{8}\Delta_{j,1}. \tag{2.12}$$

Combining these bounds with (2.6) and (2.7) we will obtain estimates for the sin and cos terms in F_{11} and F_{12}. By (2.6)

$$\cos\left(x\beta_j \pm \Delta_{j,2}\right) = 1 - \frac{(x\beta_j \pm \Delta_{j,2})^2}{2} \pm \frac{(x\beta_j \pm \Delta_{j,2})^4}{24}$$

and by combining this with (2.9) and (2.11) we obtain that

$$\cos\left(x\beta_j \pm \Delta_{j,2}\right) = 1 - \frac{(x\beta_j)^2}{2} \pm \frac{17}{16}\Delta_{j,1}. \tag{2.13}$$

In the same way, using (2.7) and (2.10),

$$\sin\left(x\beta_j \pm \Delta_{j,2}\right) = (x\beta_j \pm \Delta_{j,2}) \pm \frac{\left(x\beta_j \pm \Delta_{j,2}\right)^3}{5} = x\beta_j \pm \Delta_{j,2} \pm \frac{\Delta_{j,1}}{4}. \tag{2.14}$$

Having done these preparations we proceed with deriving bounds for F_{11}, F_{12}, F_{13}. We begin with F_{11} and using (2.13) as well as $|x|2^{-x+j} \leq 1/16$ we obtain that

$$F_{11} = \left(1 - x^2\alpha_j \pm \Delta_{j,1}\right)\left(1 - \frac{(x\beta_j)^2}{2} \pm \frac{17}{16}\Delta_{j,1}\right) = 1 - x^2\left(\alpha_j + \frac{\beta_j^2}{2}\right) \pm \frac{17\Delta_{j,1}}{8}.$$

Furthermore, by making use of (2.14) and $|x|2^{-x+j} \leq 1/16$ we obtain

$$F_{12} = (x\beta_j \pm \Delta_{j,2})\left(x\beta_j \pm \Delta_{j,2} \pm \frac{\Delta_{j,1}}{4}\right) = (x\beta_j)^2 \pm \frac{17\Delta_{j,1}}{8}$$

and using the estimate $|\exp(x) - (1+x)| \leq x^2$, valid for all $|x| \leq 1$, as well as (2.12), we get

$$F_{13} = 1 - x^2\alpha_j \pm \Delta_{j,1} \pm \left(x^2\alpha_j \pm \Delta_{j,1}\right)^2 = 1 - x^2\alpha_j \pm \frac{9\Delta_{j,1}}{8}.$$

Thus, putting F_{11}, F_{12} and F_{13} together and using oce more $|x|2^{-t+j} \leq 1/16$, we get that

$$\begin{aligned} F_1 &= \left(1 - x^2\left(\alpha_j + \frac{3\beta_j^2}{2}\right) \pm \frac{34\Delta_{j,1}}{8}\right)\left(1 - x^2\alpha_j \pm \frac{9\Delta_{j,1}}{8}\right) \\ &= 1 - x^2\left(2\alpha_j + \frac{3\beta_j^2}{2}\right) \pm 6\Delta_{j,1} = 1 - x^2\alpha_{j+1} \pm \Delta_{j+1,1} \end{aligned}$$

confirming the induction step on the first component in (2.8). Going forward we switch our attention to the second component, which we again split into three parts

$$F_2 = (F_{21} - F_{22})F_{13},$$

where

$$\begin{aligned} F_{21} &= \left(1 - x^2\alpha_j \pm \Delta_{j,1}\right)\sin\left(x\beta_j \pm \Delta_{j,2}\right) \\ F_{22} &= \left(x\beta_j \pm \Delta_{j,2}\right)\cos\left(x\beta_j \pm \Delta_{j,2}\right). \end{aligned}$$

Similarly, as above using (2.14), (2.13) and $|x|2^{-t+j} \le 1/16$ we extend these expressions. We start with F_{21}

$$F_{21} = \left(1 - x^2\alpha_j \pm \Delta_{j,1}\right)\left(x\beta_j \pm \Delta_{j,2} \pm \frac{\Delta_{j,1}}{4}\right) = x\beta_j \pm \frac{9\Delta_{j,2}}{8}.$$

Continuing with F_{22}, again applying (2.13) and $|x|2^{-t+j} \le 1/16$,

$$F_{22} = (x\beta_j \pm \Delta_{j,2})\left(1 - \frac{(x\beta_j)^2}{2} \pm \frac{17\Delta_{j,1}}{16}\right) = x\beta_j \pm \frac{17\Delta_{j,2}}{16}.$$

Finally, we combine F_{21}, F_{22} and F_{13} and with $|x|2^{-t+j} \le 1/16$ we get

$$F_2 = \left(2x\beta_j \pm \frac{33\Delta_{j,2}}{16}\right)\left(1 - x^2\alpha_j \pm \frac{9\Delta_{j,1}}{8}\right) = 2x\beta_j \pm 3\Delta_{j,2} = x\beta_{j+1} \pm \Delta_{j+1,2}.$$

Thus we confirmed the induction step and conclude the proof. □

The next lemma bounds the absolute value of $\varphi_t(x)$ for large values of x and t implying that φ is integrable, a sufficient condition for the absolute continuity of H. We did not make any effort to optimize the involved constants.

Lemma 2.11. *For all $x \in \mathbb{R},\ |x| \ge 2^{2^{17}}$ and $t \in \mathbb{N},\ t \ge \log_2(16|x|)$,*

$$|\varphi_t(x)| = a_t(x) \le |x|^{-1.2}.$$

Proof. Let $\delta = 1/32$. As $|x| \ge 1$ there is some $t_0 \ge 0$ such that $x2^{-t_0} \in [\delta, 2\delta]$. To be completely explicit,

$$t_0 := \lceil \log_2(16|x|) \rceil.$$

Thus $t \ge t_0$ and set $j^\star = t - t_0 \ge 0$. Then with Lemma 2.10 and $\sigma(x)$ denoting the sign of x

$$\begin{pmatrix} F_1 \\ F_2 \end{pmatrix} := F^{(j^\star)}\left(\begin{pmatrix} \cos(x2^{-t}) \\ \sin(x2^{-t}) \end{pmatrix}\right) \le \begin{pmatrix} 1 - 5\delta^2/8 \\ \sigma(x)\cdot 2\delta \end{pmatrix} + \begin{pmatrix} (2\delta)^3 \\ (2\delta)^2 \end{pmatrix} \le \begin{pmatrix} 1 - \delta^2/2 \\ \sigma(x)\cdot 2\delta + (2\delta)^2 \end{pmatrix}. \tag{2.15}$$

Furthermore observe that by the definition of F and the facts that $0 \le \cos i \le 1$ and $0 \le i \sin i$ for all $i \in [-\pi/2, \pi/2]$

$$F\left(\begin{pmatrix} r \\ i \end{pmatrix}\right) = e^{-1+r}\begin{pmatrix} r\cos i - i\sin i \\ r\sin i + i\cos i \end{pmatrix} \le \begin{pmatrix} re^{-1+r} \\ \pi/2 \end{pmatrix} \quad \text{for all } r \in [0,1],\ i \in [-\pi/2, \pi/2]. \tag{2.16}$$

Moreover, recall from Lemma 2.9 that

$$\begin{pmatrix} R_i(x2^{-t+i}) \\ I_i(x2^{-t+i}) \end{pmatrix} = F^{(i)}\left(\begin{pmatrix} \cos(x2^{-t}) \\ \sin(x2^{-t}) \end{pmatrix}\right) \quad \text{for} \quad R_t(x) = \mathrm{Re}(\varphi_t(x)) \text{ and } I_t(x) = \mathrm{Im}(\varphi_t(x)).$$

Thus we can bound $R_i(x2^{-t+i}),\ i \ge j^\star$ by using (2.16) for the first $i - j^\star$ applications of F and (2.15) for the remaining $j^\star$ to infer that

$$\begin{pmatrix} R_i(x2^{-t+i}) \\ I_i(x2^{-t+i}) \end{pmatrix} \le \begin{pmatrix} F_1 \cdot e^{(i-j^\star)\cdot(-1+F_1)} \\ \pi/2 \end{pmatrix} \le \begin{pmatrix} (1-\delta^2/2)\cdot e^{-(i-j^\star)\cdot\delta^2/2} \\ \pi/2 \end{pmatrix}, \quad i \ge j^\star.$$

In particular

$$R_i(x2^{-t+i}) \le 1 \quad \text{and} \quad R_{i^\star}\left(x2^{-t+i^\star}\right) \le e^{-2} \quad \text{for all} \quad i \ge 0 \text{ and } i^\star \ge j^\star + 4/\delta^2. \tag{2.17}$$

Now we switch our focus to $a_t(x)$. By applying Lemma 2.9 and (2.17)

$$a_t(x) = a_{t-1}(x/2) \cdot e^{-1+R_{t-1}(x/2)} \leq \exp\left(\sum_{i=0}^{t-1}\big(-1+R_i(x2^{-t+i})\big)\right)$$
$$\leq \exp\left(\sum_{i=j^\star+4/\delta^2}^{t-1}\big(-1+R_i(x2^{-t+i})\big)\right) \leq \exp\left(\sum_{i=j^\star+4/\delta^2}^{t-1}\big(-1+e^{-2}\big)\right).$$

Note that the sum in the exponential is non-empty, as the definition of t_0 implies that $j^\star \leq t - \log_2\big(|x|/(2\delta)\big)$ and as $|x| > 2^{2^{17}}$ we have $\log_2\big(|x|/(2\delta)\big) > 4/\delta^2 + 1$. Thus

$$a_{t+1}(x) \leq \exp\Big(\big(\log_2(|x|/(2\delta)) - 4/\delta^2\big)\big(-1+e^{-2}\big)\Big)$$
$$\leq \exp\Big(-\big(4/\delta^2 - \log_2(2\delta)\big)\big(-1+e^{-2}\big)\Big) \cdot |x|^{(-1+e^{-2})/\ln 2}.$$

This implies that $a_t(x) \leq |x|^{-1.2}$, since numerically $(-1+e^{-2})/\ln 2 \leq -1.24$ and for all x with $|x| > 2^{2^{17}}$ additionally $\exp(-(4/\delta^2 - \log_2(2\delta))(-1+e^{-2})) \cdot |x|^{-0.04} \leq 1$. □

A close inspection of the previous proof suggests that actually $a_t(x) \sim |x|^{-c}$ with $c = 1/\ln 2 \approx 1.44$. This would be interesting (and it is harder) to prove and it may have important consequences, but we will not need that; for our purpose it is enough to know that $|\varphi|$ is integrable. Next we prove the last remaining claim in Lemma 2.1.

Lemma 2.12. *H is absolute continuous and almost surely positive.*

Proof. φ is a characteristic function and therefore bounded by 1. Thus by Lemma 2.11 it is integrable and this implies the absolute continuity of H. Next we argue that H is almost surely positive. We have shown that $H_t \geq 0$ for all t and as $H_t \overset{t\to\infty}{\longrightarrow} H$ it follows that $H \geq 0$. Furthermore we have just shown that H is indeed a continuous random variable and therefore $P(H = 0) = 0$ and consequently $H > 0$ almost surely.

□

Conclusion. We have shown that $2^{-t}|I_t|$ is close to H_t, which is a martingale that converges to H, a continuous random variable. Concluding this subsection we infer Lemma 2.1 from these statements.

Proof of Lemma 2.1. In Lemma 2.6 we have shown that for all $t \in \mathbb{N}$

$$\mathrm{d}(|I_t|, J_t) = \mathrm{d}(2^{-t}|I_t|, 2^{-t}J_t) \leq 2 \cdot 4^t/n.$$

Moreover in Lemma 2.12 we have shown convergence of $2^{-t}J_t$ to H in $\mathcal{L}^2$ and therefore also in distribution. Thus for all $\varepsilon > 0$

$$\sup_{x\in\mathbb{R}} \big|P\big(2^{-t}|I_t| \geq x\big) - P\big(2^{-t}J_t \geq x\big)\big| \leq \varepsilon/2$$

as well as

$$\sup_{x\in\mathbb{R}} \big|P\big(2^{-t}J_t \geq x\big) - P(H \geq x)\big| \leq \varepsilon/2.$$

and therefore the claimed convergence follows by the triangle inequality. Lemma 2.12 shows the final claim: H is continuous and almost surely positive. □

Now that we have proven Lemma 2.1, we state and prove a simple corollary for later reference.

Corollary 2.13. *Let $t_1 = \lfloor (1/3)\log_2 n \rfloor$. Then with high probability $|I_{t_1}| = \Theta\big(n^{1/3}\big)$.*

Proof. Lemma 2.1 yields that

$$P\big(|I_{t_1}| = o(n^{1/3})\big) \le P\big(H = 2^{-t_1} o(n^{1/3})\big) + o(1) = P\big(H = o(1)\big) + o(1).$$

Since H has a density this is $o(1)$. Moreover, by applying again Lemma 2.1

$$P\big(|I_{t_1}| = \omega(n^{1/3})\big) \le P\big(H = 2^{-t_1} \omega(n^{1/3})\big) + o(1) = P\big(H = \omega(1)\big) + o(1).$$

However, since $P(H \ge h) \to 0$ when $h \to \infty$ the proof is completed. □

2.4 Proof of Lemma 2.2

We begin with a simple lemma that determines the expected number of informed and uninformed vertices after a given round.

Lemma 2.14. *For any $t \in \mathbb{N}_0$*

$$\mathbb{E}\big[|U_{t+1}| \mid I_t\big] = \big(|U_t|/n\big)^2 n \quad \textit{and} \quad \mathbb{E}\big[|I_{t+1}| \mid I_t\big] = 2|I_t| - |I_t|^2/n.$$

Proof. From the definition of *pull* we know $|I_{t+1}| = |I_t| + \mathrm{Bin}\big(n - |I_t|, |I_t|/n\big)$, see also (2.2), thus

$$\mathbb{E}\big[|I_{t+1}| \mid I_t\big] = |I_t| + \big(n - |I_t|\big) \cdot \frac{|I_t|}{n} = 2|I_t| - \frac{|I_t|^2}{n}.$$

Using the relation $|I_t| = n - |U_t|$ yields directly the second claim. □

A key property that simplifies greatly the computations in this section is the following observation, in a similar form introduced in [5] and also applied in [6].

Lemma 2.15. *For any $t \in \mathbb{N}_0$*

$$\mathrm{Var}\big[|I_{t+1}| \mid I_t\big] \le \min\big\{\mathbb{E}\big[|I_{t+1}| \mid I_t\big], \mathbb{E}\big[|U_{t+1}| \mid I_t\big]\big\}.$$

Proof. As $|I_{t+1}| = |I_t| + \mathrm{Bin}\big(|U_t|, |I_t|/n\big)$ and $|U_t| = n - |I_t|$,

$$\mathrm{Var}\big[|I_{t+1}| \mid I_t\big] = \mathrm{Var}\big[|I_t| + \mathrm{Bin}(|U_t|, |I_t|/n) \mid I_t\big] = \mathrm{Var}\big[\mathrm{Bin}(|U_t|, |I_t|/n) \mid I_t\big] = \frac{|U_t|^2 \cdot |I_t|}{n^2}.$$

This is obviously bounded from above by $\mathbb{E}\big[|U_{t+1}| \mid I_t\big] = |U_t|^2/n$ as well as by $\mathbb{E}\big[|I_{t+1}| \mid I_t\big] = |I_t| + |U_t| \cdot |I_t|/n$. □

Lemma 2.15 and Chebychev's inequality ensure that the number of informed vertices is highly concentrated around its expectation as soon as enough vertices are informed. Compare the next lemma to [6, Lem. 3.4] for a similar statement for *push*.

Lemma 2.16. *Let $t_1 = \lfloor \log_2(n^{1/3}) \rfloor$. For $t \in \mathbb{N}$, $0 < \varepsilon < 1/4$ let C_t denote the event*

$$\Big||I_{t+1}| - \mathbb{E}\big[|I_{t+1}| \mid I_t\big]\Big| \le M(I_t)^{1/2+\varepsilon} + n^{\varepsilon}, \ \textit{where}\ M(I_t) = \min\big\{\mathbb{E}\big[|I_{t+1}| \mid I_t\big], \mathbb{E}\big[|U_{t+1}| \mid I_t\big]\big\}.$$

Then

$$P\left(\bigcap_{t \ge t_1} C_t \mid I_{t_1}\right) = 1 - o(1).$$

Proof. Observe that Corollary 2.13 implies whhp

$$\mathbb{E}\big[|I_{t+1}| \mid I_t\big] \geq |I_t| \geq |I_{t_1}| \geq n^{1/4} \quad \text{for all} \quad t \geq t_1. \tag{2.18}$$

Set $n' := n - n^{1/2+\varepsilon/3}$. Then

$$P\left(\overline{C_t} \mid I_t\right) = P\left(\overline{C_t}\mathbf{1}_{|I_t|>n'} \mid I_t\right) + P\left(\overline{C_t}\mathbf{1}_{|I_t|\leq n'} \mid I_t\right).$$

If $|I_t| > n'$, then $|U_t| < n^{1/2+\varepsilon/3}$ and so $M(I_t) \leq \mathbb{E}[|U_{t+1}| \mid I_t] = |U_t|^2/n \leq n^{2\varepsilon/3}$. In that case $\overline{C_t}$ thus implies that $|U_{t+1}| \geq n^{\varepsilon} \geq n^{\varepsilon/3}\mathbb{E}[|U_{t+1}| \mid I_t]$. By Markov's inequality

$$P\left(\overline{C_t}\mathbf{1}_{|I_t|>n'} \mid I_t\right) \leq P\big(|U_{t+1} \geq n^{\varepsilon/3}\mathbb{E}[|U_{t+1}| \mid I_t] \mid I_t\big) \leq n^{-\varepsilon/3}.$$

If $|I_t| \leq n'$, then $|U_t| \geq n^{1/2+\varepsilon/3}$ and using (2.18) also $M(I_t) \geq \min\{n^{1/4}, |U_t|^2/n\} \geq n^{2\varepsilon/3}$. In this case, using Lemma 2.15, $\overline{C_t}$ implies that

$$\Big||I_{t+1}| - \mathbb{E}\big[|I_{t+1}| \mid I_t\big]\Big| > M(I_t)^{1/2+\varepsilon} \geq n^{2\varepsilon^2/3}\mathrm{Var}\big[|I_{t+1}| \mid I_t\big]^{1/2}.$$

By Chebychev's inequality

$$P\left(\overline{C_t}\mathbf{1}_{|I_t|\leq n'} \mid I_t\right) \leq \; P\left(\Big||I_{t+1}| - \mathbb{E}\big[|I_{t+1}| \mid I_t\big]\Big| > n^{2\varepsilon^2/3}\mathrm{Var}\big[|I_{t+1}| \mid I_t\big]^{1/2} \;\Big|\; I_t\right) \leq n^{-\varepsilon^2}.$$

By combining both cases we get the very crude bound

$$P\big(\overline{C_t} \mid I_t\big) \leq n^{-\varepsilon^2} \quad \text{for all } t \geq t_1. \tag{2.19}$$

The large deviation bounds (1.1) give us that $X_n \geq 2\log_2 n$ has exponentially small probability. Thus

$$P\left(\bigcup_{t\geq 2\log_2 n} \overline{C_t} \mid I_{t_1}\right) = o(1).$$

A union bound and (2.19), applied to $O(\log n)$ many $t_1 \leq t \leq t_2$, then yield the claim. □

Lemma 2.16 shows that $|I_t|$ is closely concentrated around its (conditional) expectation in all rounds. This translates directly to concentration of $|U_{t+1}|$ around $(|U_t|/n)^2 n = \mathbb{E}[|U_{t+1}| \mid U_t]$ for all $t \geq t_1$. Using this, we are now ready to prove Lemma 2.2, that is, $|U_{t_1+t}|$ is close to $(|U_{t_1}|/n)^{2^t} n$ for all $t \in \mathbb{N}$ with high probability.

Proof of Lemma 2.2. We assume that $|U_{t_1}| = \Theta\big(n^{1/3}\big)$, which we know from Corollary 2.13 has high probability. Consequently we can apply Lemma 2.16 with $\varepsilon = 1/10$ and thus we get with high probability for all $t \geq t_1$

$$\big||I_{t+1}| - \mathbb{E}\big[|I_{t+1}| \mid I_t\big]\big| \leq \Big(\min\big\{\mathbb{E}\big[|I_{t+1}| \mid I_t\big], \mathbb{E}\big[|U_{t+1}| \mid I_t\big]\big\}\Big)^{3/5} + n^{1/10}. \tag{2.20}$$

For the rest of this proof we assume in addition (2.20), that is, we assume that $(|I_t|)_{t\geq t_1}$ (and thus also $(|U_t|)_{t\geq t_1}$ and $(\mathbb{E}[|I_{t+1}| \mid I_t])_{t\geq t_1}$) are sequences of numbers with the aforementioned properties. In particular, (2.20) implies for all $\delta > 0$ that $\big|2|I_t| - |I_{t+1}|\big| \leq \delta|I_t|$ for all $t \geq t_1$ and $n > \delta^{-15}$, where $t_1 = \lfloor\log_2(n^{1/3})\rfloor$. Therefore

$$|I_{t_1+s}| \leq (2+\delta)^s|I_{t_1}| \quad \text{for all } s \in \mathbb{N}_0,\ \delta > 0 \text{ and } n > \delta^{-15}. \tag{2.21}$$

Set

$$\beta_{t_1+s} := \big(|U_{t_1}|/n\big)^{2^s}, \quad s \in \mathbb{N}_0.$$

Note that $t_1 + s$ is just a different way to parameterize $t \geq t_1$, which simplifies the notation when t_1 is involved. In particular t_1 is always fixed to the aforementioned value.

We will next argue that for all $t \geq t_1$, abbreviating $\Delta_t := \big||U_t| - \beta_t n\big|$,

$$\Delta_{t+1} \leq \Big(\min\big\{2|I_t|, |U_t|^2/n\big\}\Big)^{3/5} + n^{1/10} + \Big(2|U_t|/n + \Delta_t/n\Big)\Delta_t. \tag{2.22}$$

To see this, note first that by using (2.20) and Lemma 2.14,

$$\big||U_{t+1}| - (|U_t|/n)^2 n\big| = \big||I_{t+1}| - \mathbb{E}\big[|I_{t+1}| \;\big|\; I_t\big]\big| \leq \Big(\min\big\{2|I_t|, |U_t|^2/n\big\}\Big)^{3/5} + n^{1/10}.$$

Secondly, applying the triangle inequality, i.e., $|x+y| \leq 2|x| + |x-y|$ for all $x, y \in \mathbb{R}$, yields

$$\begin{aligned}\big|(|U_t|/n)^2 n - \beta_{t+1} n\big| &= \big||U_t|^2/n - \beta_t^2 n\big| = \big||U_t|/n + \beta_t\big| \cdot \big||U_t| - \beta_t n\big| \\ &\leq \Big(2|U_t|/n + \big||U_t|/n - \beta_t\big|\Big)\big||U_t| - \beta_t n\big|.\end{aligned}$$

The triangle inequality then implies (2.22).

In the remainder of this proof we will look at the bound of Δ_t in (2.22) in three different ways to distinguish in each case a different behaviour. Just to wit, at first Δ_t doubles as long as the number of informed vertices doubles. However, as soon as the doubly exponential shrinking of the uninformed vertices takes over, also the error Δ_t shrinks rapidly, so that Δ_t always remains $o\big(|U_t|\big)$. In end we just make sure that Δ_t stays small and does not increase any more.

We will make this outline more precise by formulating matching claims, which we then use to infer the statement of this lemma. We prove the claims afterwards. Our first claim is that for $\delta = 1/100$ and $d = 400$

$$\Delta_{t_1+s} \leq d(2+\delta)^s |I_{t_1}|^{3/5} \quad \text{for all } 0 \leq s \leq (14/15)\log_2 n - t_1 \text{ and } n > \delta^{-15}. \tag{Claim 1}$$

By using that $|I_{t_1}| = \Theta(n^{1/3}$ and that $14/15 - 1/3 + 1/5 = 4/5 < 0.85$, (Claim 1) implies for sufficiently large n that for $t_2 = \lfloor (14/15)\log_2 n \rfloor$

$$\Delta_{t_1+s} \leq d\,(2+\delta)^{s+1}\,|I_{t_1}|^{3/5} \leq n^{0.85} \quad \text{for all } 0 \leq s \leq t_2 - t_1. \tag{2.23}$$

Moreover, (2.21) yields

$$|I_{t_1+s}| \leq (2+\delta)^{s+1}\,|I_{t_1}| = o(n) \quad \text{for all } 0 \leq s \leq t_2 - t_1,$$

and therefore, as $|U_{t_1+s}| = n - |I_{t_1+s}| = \Theta(n)$,

$$\Delta_{t_1+s} \leq |U_{t_1+s}| \cdot n^{-1/10} \quad \text{for all } 0 \leq s \leq t_2 - t_1. \tag{2.24}$$

From (1.1) we know that whp $X_n \leq \log_2 n + 2\log_2 \ln n$, thus we need to bound Δ_t for at most an additional $(1/15)\log_2 n + 2\log_2 \ln n$ steps. For these steps we will need a different bound, as the bound in (Claim 1) is only useful as long as $t_1 + s < \log_2 n$; otherwise the term $(2+\delta)^s$ blows up. Our next claim is

$$\Delta_{t_2+s} \leq |U_{t_2+s}| \cdot (2+\delta)^s \cdot n^{-1/11} \quad \text{for all } s \in \mathbb{N} \text{ with } |U_{t_2+s}| \geq n^{1/4} \text{ and } n \geq \delta^{-111}. \tag{Claim 2}$$

We saw that it suffices to apply (Claim 2) for at most $s \leq (1/15)\log_2 n + 2\log_2 \ln n$ additional steps. For such s we obtain that $(2+\delta)^{s+1} n^{-1/11} \leq n^{-1/50}$ for sufficiently large n and therefore we conclude from (2.24) for all $t \leq t_2$ and (Claim 2) for all $t > t_2$

$$\Delta_t \leq |U_t| \cdot n^{-1/50} \quad \text{for all} \quad t \geq t_1 \quad \text{as long as} \quad |U_t| \geq n^{1/4}.$$

For $t \geq t_1$ such that $|U_t| \leq n^{1/4}$ the claim of the Lemma follows trivially.

Now we show (Claim 1) and (Claim 2), starting with (Claim 1), which we are going to do by induction. The base case follows directly from (2.22). Let $t \geq t_1$, from (2.22) and using that $|I_t| = \Omega\big(n^{1/3}\big)$ we get that

$$\Delta_{t+1} \leq 2|I_t|^{3/5} + \big(2 + \Delta_t/n\big)\Delta_t.$$

Observe that by using the induction hypothesis, we obtain that

$$\Delta_{t_1+s} \le |U_{t_1+s}| \cdot n^{-1/10}$$

and thus using (2.21), the induction hypothesis, and $n \ge \delta^{-15}$ gives us

$$\begin{aligned}\Delta_{t_1+s+1} &\le 2\big((2+\delta)^s|I_{t_1}|\big)^{3/5} + (2+\delta/2)\left(d(2+\delta)^s|I_{t_1}|^{3/5}\right)\\ &\le \left(\frac{2}{d}(2+\delta)^{-2s/5} + 2 + \frac{\delta}{2}\right) d(2+\delta)^s|I_{t_1}|^{3/5}.\end{aligned}$$

Our choice of d and δ guarantees that $(2/d)(2+\delta)^{(-2/5)s} \le (2/d)(2+\delta)^{(-2/5)} \le \delta/2$ for all $s \in \mathbb{N}$ and (Claim 1) follows.

We continue with (Claim 2) that we we will prove by induction, too. To that end, we observe that (2.20) also implies that as long as $|U_{t+1}| \ge n^{1/4}$ and $n \ge \delta^{-6}$

$$\big||U_t|^2/n - |U_{t+1}|\big| \le \delta|U_{t+1}|/4. \tag{2.25}$$

The base case of (Claim 2) follows directly from (2.24). For the induction step we apply (2.22) together with the induction hypothesis and get

$$\begin{aligned}\Delta_{t_2+s+1} &\le \left(\frac{|U_{t_2+s}|^2}{n}\right)^{3/5} + \left(2\frac{|U_{t_2+s}|}{n} + \Delta_{t_2+s}/n\right)\Delta_{t_2+s}\\ &\le \left(\frac{|U_{t_2+s}|^2}{n}\right)^{3/5} + (2+\delta/2)\frac{|U_{t_2+s}|}{n}\Delta_{t_2+s}\\ &\le \left(\frac{|U_{t_2+s}|^2}{n}\right)^{3/5} + (2+\delta/2)(2+\delta)^s \cdot \frac{|U_{t_2+s}|^2}{n} \cdot n^{-1/11}.\end{aligned}$$

Using the assumption that $|U_{t_2+s+1}| \ge n^{1/4}$ we get, as $(1/4)\cdot(2/5) - 1/11 = 1/110$, that

$$\big(|U_{t_2+s}|^2/n\big)^{3/5} / \big(|U_{t_2+s}|^2/n \cdot n^{-1/11}\big) \le \delta/4 \quad \text{for all} \quad s \in \mathbb{N}_0 \quad \text{and} \quad n > \delta^{-111}.$$

This and (2.25) implies (Claim 2). □

2.5 Proof of Lemma 2.3

We will prove this lemma in two steps, first we start with a simple lemma showing that, if there are much more than $\sqrt{n}$ uninformed vertices remaining, *pull* will not end in the next round. Moreover if there are substantially less than $\sqrt{n}$ uninformed the protocol will end in the next round.

Lemma 2.17. *Let $t, t' \in \mathbb{N}$ such that $|U_t| \le \sqrt{n}/\ln n$ and $|U_{t'}| \ge \sqrt{n}\ln n$. Then*

$$P(|U_{t+1}| = 0 \mid I_t) = o(1) \quad \textit{and} \quad P(|U_{t'+1}| > 0 \mid I_{t'}) = o(1).$$

Proof. Note that

$$\mathbb{E}\big[|U_{t+1}| \bigm| I_t\big] = \frac{|U_t|^2}{n} \le \ln^{-2} n.$$

This yields with Markov's inequality the claim for t. To see the claim for t' we observe first that the probability of one uninformed vertex $v \in U_{t'}$ being informed in the next round is

$$P\big(u \in I_{t'+1} \bigm| u \in U_{t'}\big) = \frac{|I_{t'}|}{n} = \left(1 - \frac{|U_{t'}|}{n}\right).$$

If $|U_{t'+1}| = 0$, then all $|U_{t'}|$ uninformed vertices need to be informed in the next round, and as that happens independently, we get

$$P\big(|U_{t'+1}| = 0 \bigm| I_{t'}\big) = \left(1 - \frac{|U_{t'}|}{n}\right)^{|U_{t'}|} \le e^{-|U_{t'}|^2/n} \le e^{-\ln^2 n}.$$

□

Besides the two cases that we considered in the previous lemma a third case is also possible. Indeed, if there are about $\sqrt{n}$ uninformed vertices, the process ending in the next round may happen with some non-trivial probability. The next lemma shows that, however, that this is very unlikely to happen. This means that once the process crosses the threshold of $\sqrt{n}$ it will terminate in the next round with high probability.

Lemma 2.18. *With high probability for all $t \in \mathbb{N}$,*

$$|U_t| \notin \left[\sqrt{n}/\ln n, \sqrt{n}\ln n\right].$$

Proof. Let $t_1 = \lfloor \log_2(n^{1/3}) \rfloor$ and consider the events

$$\left\{ |I_{t_1}| = \Theta\big(n^{1/3}\big) \right\}, \tag{Event 1}$$

and with $\eta = \eta(n) = \log_2 n + \log_2 \ln n - \lfloor \log_2 n + \log_2 \ln n \rfloor$,

$$\left\{ \log_2\left(2^{-t_1}|I_{t_1}|\right) \notin \bigcup_{k\in\mathbb{Z}} \left[k + \eta - \frac{3\log_2 \ln(n^2)}{\ln n}, k + \eta + \frac{3\log_2 \ln(n^2)}{\ln n} \right] \right\}, \tag{Event 2}$$

as well as

$$\bigcap_{t\geq 0} \left\{ \left| |U_{t_1+t}| - \big(|U_{t_1}|/n\big)^{2^t} n \right| \leq |U_{t_1+t}| \cdot n^{-1/50} + n^{1/4} \right\}. \tag{Event 3}$$

All these events occur with high probability. For (Event 1) this was already established in Corollary 2.13. To see the claim for (Event 2) we observe that $\log_2\left(2^{-t_1}|I_{t_1}|\right)$ converges to the continuous random variable $\log_2 H$, see Lemma 2.1, and that the right side in (Event 2) converges to a $\log_2 H$ null-set. (Event 3) was handled in Lemma 2.2.

In the remainder of this proof we condition on these events, that is, we assume that $(|I_t|)_{t\geq t_1}$ (and thus also $(|U_t|)_{t\geq t_1}$) are sequences of numbers with the aforementioned properties.

Observe that for $n^{1/3} < a < b < n - n^{1/3}$, (Event 1) and (Event 3) imply for $t \geq t_1$ and n large enough

$$(|U_{t_1}|/n)^{2^{t-t_1}} n \notin [a/2, 2b] \quad \Longrightarrow \quad |U_t| \notin [a,b];$$

to see this, note that by assumption $|U_t| = (1+o(1))(|U_{t_1}|/n)^{2^{t-t_1}} n$ and so for large enough n if $|U_t| \in [a,b]$, then with room to spare $(|U_{t_1}|/n)^{2^{t-t_1}} n \in [a/2, 2b]$.

Set $u_t = (1-|I_{t_1}|/n)^{2^{t-t_1}} n$ and define $T_u = \min\{t \in \mathbb{N} : u_t < 2\sqrt{n}\ln n\}$ as well as $T_\ell = \min\{t \in \mathbb{N} : u_t < \sqrt{n}/(2\ln n)\}$. With these definitions we can apply the implication we just derived to the event in the statement of the lemma and obtain that

$$T_u = T_\ell \quad \Longrightarrow \quad |U_t| \notin \left[\sqrt{n}/\ln n, \sqrt{n}\ln n\right]. \tag{2.26}$$

Observe that (Event 1) implies that

$$u_{\lfloor (4/3)\log_2 n \rfloor} n = \big(|U_{t_1}|/n\big)^{2^{\lfloor (4/3)\log_2 n\rfloor - t_1}} n = o(1),$$

thus $T_u, T_\ell < (4/3)\log_2 n$ and consequently we need to study u_t for that range of t only. Therefore, using $1 - x = e^{-x+O(x^2)}$ for small x and $0 \leq t < (4/3)\log_2 n$

$$\begin{aligned} u_t = n\big(1 - |I_{t_1}|/n\big)^{2^{t-t_1}} &= n \cdot \exp\big(- 2^{t-t_1}|I_{t_1}|/n + O(2^{t-t_1}|I_{t_1}|^2/n^2)\big) \\ &= \big(1 + O(n^{-2/3})\big) \cdot n \cdot \exp\left(-2^{t-t_1}|I_{t_1}|/n\right). \end{aligned}$$

To determine T_u and T_ℓ it suffices to solve the equations

$$\exp\left(-2^{t-t_1}|I_{t_1}|/n\right) = \big(1 + O(n^{-2/3})\big) c_n n^{-1/2}, \quad \text{for } c_n \in \big\{2\ln n,\ 1/2\ln n\big\}.$$

Applying logarithms twice and using the first order expansion $\ln(1+x) = x + O(x^2),\ |x| < 1/2$, we readily obtain that

$$T_u = \Big\lfloor \log_2 n - \log_2(2^{-t_1}|I_{t_1}|) + \log_2 \ln n - 1 - \frac{2\log_2 \ln(n^2)}{\ln n} + O(1/\ln n) \Big\rfloor$$

and

$$T_\ell = \Big\lfloor \log_2 n - \log_2(2^{-t_1}|I_{t_1}|) + \log_2 \ln n - 1 + \frac{2\log_2 \ln(n^2)}{\ln n} + O(1/\ln n) \Big\rfloor.$$

Observe that for any $x, y, z \in [0,1]$ with $x \leq y$ it holds that $\lfloor x+z \rfloor \neq \lfloor y+z \rfloor$ if and only if $z \in [1-y, 1-x)$; to see this just note that if $z < 1-y$, then both terms are equal to 0, if $z \geq 1-x$ then both terms are equal to 1 and otherwise just one of them is 0. Thus for $\eta = \log_2 n + \log_2 \ln n - \lfloor \log_2 n + \log_2 \ln n \rfloor$ and n large enough

$$\log_2\left(2^{-t_1}|I_{t_1}|\right) \notin \bigcup_{k\in\mathbb{Z}} \left[k + \eta - \frac{3\log_2 \ln(n^2)}{\ln n}, k + \eta + \frac{3\log_2 \ln(n^2)}{\ln n}\right] \quad \Longrightarrow \quad T_u = T_\ell.$$

Since we have assumed (Event 2) we have just established that $T_u = T_\ell$, and together with (2.26) the proof is completed. $\square$

References

[1] N. Berger, C. Borgs, J. T. Chayes, and A. Saberi. On the spread of viruses on the internet. In *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '05, page 301–310, USA, 2005. Society for Industrial and Applied Mathematics.

[2] F. Chierichetti, G. Giakkoupis, S. Lattanzi, and A. Panconesi. Rumor spreading and conductance. *Journal of the ACM (JACM)*, 65(4):1–21, 2018.

[3] A. Clementi, P. Crescenzi, C. Doerr, P. Fraigniaud, F. Pasquale, and R. Silvestri. Rumor spreading in random evolving graphs. *Random Structures & Algorithms*, 48(2):290–312, 2016.

[4] R. M. Corless, G. H. Gonnet, D. E. Hare, D. J. Jeffrey, and D. E. Knuth. On the LambertW function. *Advances in Computational mathematics*, 5(1):329–359, 1996.

[5] R. Daknama, K. Panagiotou, and S. Reisser. Robustness of randomized rumour spreading. *Combinatorics, Probability and Computing*, page 1–42, 2020.

[6] R. Daknama, K. Panagiotou, and S. Reisser. Asymptotics for push on the complete graph. *Stochastic Processes and their Applications*, 2021.

[7] A. Demers, D. Greene, C. Hauser, W. Irish, J. Larson, S. Shenker, H. Sturgis, D. Swinehart, and D. Terry. Epidemic algorithms for replicated database maintenance. In *Proceedings of the sixth annual ACM Symposium on Principles of distributed computing*, pages 1–12. ACM, 1987.

[8] B. Doerr and A. Kostrygin. Randomized rumor spreading revisited. In *44th International Colloquium on Automata, Languages, and Programming (ICALP 2017)*. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2017.

[9] B. Doerr and M. Künnemann. Tight Analysis of Randomized Rumor Spreading in Complete Graphs. In *Proceedings of the Meeting on Analytic Algorithmics and Combinatorics*, pages 82–91, Philadelphia, PA, USA, 2014. Society for Industrial and Applied Mathematics.

[10] R. Durrett. *Probability: theory and examples*, volume 49. Cambridge university press, 2019.

[11] U. Feige, D. Peleg, P. Raghavan, and E. Upfal. Randomized broadcast in networks. *Random Structures & Algorithms*, 1(4):447–460, 1990.

[12] N. Fountoulakis, A. Huber, and K. Panagiotou. Reliable broadcasting in random networks and the effect of density. In *INFOCOM, 2010 Proceedings IEEE*, pages 1–9. IEEE, 2010.

[13] A. M. Frieze and G. R. Grimmett. The shortest-path problem for graphs with random arc-lengths. *Discrete Applied Mathematics*, 10(1):57–77, 1985.

[14] G. Giakkoupis. Tight Bounds for Rumor Spreading with Vertex Expansion. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2014, Portland, Oregon, USA, January 5-7, 2014*, pages 801–815, 2014.

[15] G. Giakkoupis, Y. Nazari, and P. Woelfel. How asynchrony affects rumor spreading time. In *Proceedings of the 2016 ACM Symposium on Principles of Distributed Computing*, pages 185–194, 2016.

[16] K. Iwanicki and M. van Steen. Gossip-based self-management of a recursive area hierarchy for large wireless sensornets. *IEEE Transactions on Parallel and Distributed Systems*, 21(4):562–576, 2009.

[17] S. Janson. One, two and three times log n/n for paths in a complete graph with random weights. *Combinatorics, Probability and Computing*, 8(4):347–361, 1999.

[18] R. M. Karp, C. Schindelhauer, S. Shenker, and B. Vöcking. Randomized Rumor Spreading. In *41st Annual Symposium on Foundations of Computer Science, FOCS 2000, 12-14 November 2000, Redondo Beach, California, USA*, pages 565–574, 2000.

[19] K. Panagiotou and L. Speidel. Asynchronous Rumor Spreading on Random Graphs. *Algorithmica*, 78(3):968–989, 2017.

[20] C. Patsonakis and M. Roussopoulos. Revisiting Asynchronous Rumor Spreading in the Blockchain Era. In *2019 IEEE 25th International Conference on Parallel and Distributed Systems (ICPADS)*, pages 284–293, Dec 2019.

[21] B. Pittel. On Spreading a Rumor. *SIAM J. Appl. Math.*, 47(1):213–223, Mar. 1987.

[22] A. Pourmiri and B. Mans. Tight analysis of asynchronous rumor spreading in dynamic networks. In *Proceedings of the 39th Symposium on Principles of Distributed Computing*, pages 263–272, 2020.

[23] R. J. Serfling. Some Elementary Results on Poisson Approximation in a Sequence of Bernoulli Trials. *SIAM Review*, 20(3):567–579, 1978.

7 Asymptotics for Push on the Complete Graph

This chapter is the published version of

> Daknama, R., Panagiotou, K., & Reisser, S. (2021). Asymptotics for push on the complete graph. *Stochastic Processes and their Applications*, Volume 137, 35-61.

The published version is online at https://doi.org/10.1016/j.spa.2021.03.008.

My own contribution. This paper is joint work with Rami Daknama and my supervisor Konstantinos Panagiotou. We developed the results in joint discussion and I contributed substantially to all results. Theorem 1.1. and its proof is already included in the thesis of Rami Daknama [17]. Theorem 1.2 and Theorem 1.3, as well as the simulations were done by me. Moreover, the final presentation of all results of this paper is written by me, in particular I substantially reworked and improved all parts that Rami Daknama included in his thesis. This work is based on constant discussions with and continual improvements by Konstantinos Panagiotou.

Available online at www.sciencedirect.com
ScienceDirect
Stochastic Processes and their Applications 137 (2021) 35–61
stochastic processes and their applications
www.elsevier.com/locate/spa

Asymptotics for push on the complete graph☆

Rami Daknama, Konstantinos Panagiotou*, Simon Reisser

Department of Mathematics, LMU Munich, Theresienstraße 39, 80333 Munich, Germany

Received 27 May 2020; received in revised form 8 March 2021; accepted 12 March 2021
Available online 29 March 2021

Abstract

We study the classical randomized rumour spreading protocol *push*. Initially, a node in a graph possesses some information, which is then spread in a round based manner. In each round, each informed node chooses uniformly at random one of its neighbours and passes the information to it. The central quantity of interest is the *runtime*, that is, the number of rounds needed until every node has received the information.

The *push* protocol and variations of it have been studied extensively. Here we study the case where the underlying graph is complete with n nodes. Even in this most basic setting, specifying the limiting distribution and statistics of it have remained open problems since the protocol was introduced. In our main result we describe the limiting distribution of the runtime. We show that it does not converge, and that it becomes, after the appropriate normalization, asymptotically periodic both on the $\log_2 n$ as well as on the $\ln n$ scale. Additionally, on suitable subsequences we determine the expected runtime and higher moments of it.

MSC: 05C85; 68R10

Keywords: Randomized rumour spreading; Complete graph; Asymptotic

1. Introduction

We consider the well-known and well-studied rumour spreading protocol *Push*. It has applications in replicated databases [6], multicast [1] and blockchain technology [22]. *Push*

☆ An extended abstract of this paper was published at the Proceedings of 14th Latin American Theoretical Informatics Symposium (LATIN' 20), 2020.

* Corresponding author.
E-mail addresses: rami.300@hotmail.de (R. Daknama), kpanagio@math.lmu.de (K. Panagiotou), reisser@math.lmu.de (S. Reisser).

https://doi.org/10.1016/j.spa.2021.03.008
0304-4149/

operates on graphs and proceeds in rounds as follows. In the beginning, one node has a piece of information. In subsequent rounds each informed node chooses a neighbour independently and uniformly at random and informs it. For a graph $G = (V, E)$ with $|V| = n$ and a node $v \in V$ we denote by $X(G, v)$ the (random) number of rounds needed to inform all nodes, where at the beginning of the first round only v knows the information. We call $X(G, v)$ the *runtime* (on G with start node v). The most basic case, and the one that we study here, is when G is the complete graph K_n. Since in that case the initially informed node makes no difference, we will abbreviate $X(K_n, v) = X_n$ for any starting node v.

Related work. There are several works studying the runtime of *push* on the complete graph. The first paper considering this protocol is by Frieze and Grimmett [12], who showed that with high probability (whp), that is, with probability $1 - o(1)$ as $n \to \infty$, that

$$X_n = \log_2 n + \ln n + o(\ln n).$$

Moreover, they obtained bounds for (very) large deviations of X_n from its expectation. In [23], Pittel improved upon the results in [12], in particular, he showed that for any $f : \mathbb{N} \to \mathbb{R}^+$ with $f = \omega(1)$, whp,

$$|X_n - \log_2 n - \ln n| \leq f(n).$$

The currently most precise result in this context was obtained by Doerr and Künnemann [7], who considered in great detail the distribution of X_n. They showed that X_n can be stochastically bounded (from both sides) by coupon collector type problems. This gives a lot of control regarding the distribution of X_n, and it allowed them to derive, for example, very sharp bounds for tail probabilities. Apart from that, it enabled them to consider related quantities, as for example the expectation of X_n. Among other results, their bounds on the distribution of X_n imply that

$$\lfloor \log_2 n \rfloor + \ln n - 1.116 \leq \mathbb{E}[X_n] \leq \lceil \log_2 n \rceil + \ln n + 2.765, \tag{1}$$

which pins down the expectation up to a constant additive term. Besides on complete graphs, *push* has been extensively studied on several other graph classes. For example, Erdős–Rényi random graphs [9,10], random regular graphs and expander graphs [5,11,20]. More general bounds that only depend on some graph parameter have also been derived, e.g. the diameter [9], graph conductance [3,4,13,19] and node expansion [4,14,16,26].

A very much related and significantly better understood variation is the so-called *asynchronous rumour spreading*. The key difference is that the communication is no longer synchronized; instead, each node is equipped with an independent rate-1 Poisson process, at whose ticks push operations are performed, see e.g. [15,21,25]. Let us denote by Y_n the runtime in this setting, which is defined as the earliest point in time at which all nodes are informed. By considering the equivalent question about shortest paths on complete graphs with random edge weights (in this case exponentially distributed random variables scaled by n) we can actually derive the limiting distribution of Y_n. In [18] Janson does just exactly that (and much more), and his results imply that as $n \to \infty$, in distribution

$$Y_n - 2\log_2 n \to W_1 + W_2,$$

where W_1, W_2 are independent random variables that follow the *Gumbel distribution*, see also below. Moreover, we obtain that

$$\mathbb{E}[Y_n] = 2\log n + 2\gamma + o(1)$$

where γ denotes the Euler–Mascheroni constant. Our main aim here is to obtain a similar fine grained understanding of the runtime X_n of the (synchronous) push protocol.

Results. In order to state our main result we need some definitions first. Set

$$g = g^{(1)} : [0, 1] \to [0, 1], \quad x \mapsto xe^{x-1}$$

and

$$g^{(i)} : [0, 1] \to [0, 1], \quad g^{(i)} = g \circ g^{(i-1)}, i \geq 2.$$

As we will see later, the function g describes, for a wide range of the parameters, the evolution of the number of uninformed nodes; in particular, if at the beginning of some round there are xn uninformed nodes, then at the end of the same round there will be (roughly) $g(x)n$ uninformed nodes, and after i rounds there will be (roughly) $g^{(i)}(x)n$ uninformed nodes. This fact is not new – at least for bounded i – and has been observed long ago, see for example [23, Lem. 2]. For $x \in \mathbb{R}$ define the function

$$c(x) = -x + \lim_{a\to\infty, a\in\mathbb{N}} \lim_{b\to\infty, b\in\mathbb{N}} -a + b + \ln\left(g^{(b)}(1 - 2^{-a-x})\right), \tag{2}$$

whose actual meaning will become clear later. We will show that the double limit exists, so that this indeed defines a function $c : \mathbb{R} \to \mathbb{R}$. Moreover, we will show that c is continuous and periodic with period 1, that is, if we write $\{x\} = x - \lfloor x \rfloor$ then $c(x) = c(\{x\})$, and that (numerically) $|\sup c - \inf c| \approx 10^{-9}$, cf. Fig. 1. The Gumbel distribution will play a prominent role in our considerations. We say that a real valued random variable G follows a $\mathrm{Gum}(\alpha)$ distribution with parameter $\alpha \in \mathbb{R}$, $G \sim \mathrm{Gum}(\alpha)$, if for all $x \in \mathbb{R}$

$$P[G \leq x] = e^{-e^{-x-\alpha}}, \quad x \in \mathbb{R}.$$

With all these ingredients we can now state our main result, which specifies – see also below – the distribution of the runtime of *push* on the complete graph.

Theorem 1.1. *Let $G \sim \mathrm{Gum}(\gamma)$. Then, as $n \to \infty$*

$$\sup_{k\in\mathbb{N}} \left| P[X_n \geq k] - P\Big[\lceil G + \log_2 n + \ln n + \gamma + c(\{\log_2 n\})\rceil \geq k\Big]\right| = o(1).$$

This theorem does not look completely innocent, and it actually has striking consequences. It readily implies the following result, which establishes that the limiting distribution X_n is periodic both on the $\log_2 n$ and on the $\ln n$ scale. In order to formulate it, we need a version of the Gumbel distribution where we restrict ourselves to integers only. More specifically, we say that a random variable G follows a *discrete Gumbel* distribution, $G \sim \mathrm{dGum}(\alpha)$, if the domain of G is $\mathbb{Z}$ and

$$P[G \leq k] = e^{-e^{-k-\alpha}}, \quad k \in \mathbb{Z}.$$

Theorem 1.2. *Let $x, y \in [0, 1)$ and $(n_i)_{i\in\mathbb{N}}$ be a strictly increasing sequence of natural numbers, such that $\log_2 n_i - \lfloor \log_2 n_i \rfloor \to x$ and $\ln n_i - \lfloor \ln n_i \rfloor \to y$ as $i \to \infty$. Then in distribution, as $i \to \infty$*

$$X_{n_i} - \left(\lfloor \log_2 n_i \rfloor + \lfloor \ln n_i \rfloor\right) \to \mathrm{dGum}(-x - y - c(x)).$$

Some remarks are in place. First, it is a priori not obvious (at least it was not to us) that subsequences as required in the theorem indeed exist. They do, and the fundamental reason for this is that real numbers can be approximated arbitrarily well by rational numbers; we include a short proof of the existence in the Appendix. Second, it is a priori not clear that $x + c(x)$ is not constant for $x \in [0, 1)$. If it was constant, Theorem 1.2 would imply that the limiting

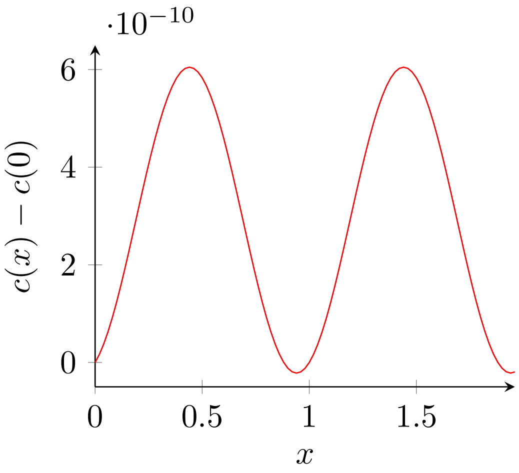

Fig. 1. The function $c(x) - c(0)$, $c(0) \approx 0.105$, plotted for values of x between 0 and 2. The periodic nature of the function and its small amplitude are evident.

distribution of X_n is periodic on the $\ln n$ scale only. Although we did not manage to *prove* that $x + c(x)$ is not constant, we have strong numerical evidence that it indeed is not so. In particular, as we shall also see later, the double limit in the definition of c converges exponentially fast and thus it is not difficult to obtain accurate estimates for it and explicit error bounds. We leave it as an open problem to study the behaviour of c more accurately.

Our next result addresses moments of X_n. Bounds given in previous works, for example in [7], guarantee that $X_n - \log_2 n - \ln n$ and all integer powers of it are uniformly integrable. This allows us to conclude that the expectation and all of its moments also converge.

Theorem 1.3. *Let $x, y \in [0, 1)$ and $(n_i)_{i \in \mathbb{N}}$ be a strictly increasing sequence of natural numbers, such that $\log_2 n_i - \lfloor \log_2 n_i \rfloor \to x$ and $\ln n_i - \lfloor \ln n_i \rfloor \to y$ as $i \to \infty$. Then for all $k \in \mathbb{N}$, as $i \to \infty$*

$$\mathbb{E}\Big[\big(X_{n_i} - (\lfloor \log_2 n_i \rfloor + \lfloor \ln n_i \rfloor)\big)^k\Big] \to \mathbb{E}\Big[\big(\mathrm{dGum}(-x - y - c(x))\big)^k\Big].$$

For $x, y \in [0, 1)$ and a strictly increasing sequence of natural numbers $(n_i)_{i \in \mathbb{N}}$ such that $\{\log_2 n_i\} \to x$ and $\{\ln n_i\} \to y$ Theorem 1.3 immediately implies that, as $i \to \infty$,

$$\mathbb{E}\big[X_{n_i}\big] = \log_2 n_i + \ln n_i + h(x, y) + o(1),$$

where we abbreviated $h(x, y) = \mathbb{E}\big[\mathrm{dGum}(-x - y - c(x))\big] - x - y$, cf. Fig. 2 for a visualization of h. Similarly, to obtain an expression for the variance of the runtime, see that

$$\begin{aligned}\mathrm{Var}[X_{n_i}] &= \mathrm{Var}\big[X_{n_i} - (\lfloor \log_2 n_i \rfloor + \lfloor \ln n_i \rfloor)\big] \\ &= \mathbb{E}\Big[\big(X_{n_i} - (\lfloor \log_2 n_i \rfloor + \lfloor \ln n_i \rfloor)\big)^2\Big] - \mathbb{E}\big[X_{n_i} - (\lfloor \log_2 n_i \rfloor + \lfloor \ln n_i \rfloor)\big]^2\end{aligned}$$

and using Theorem 1.3, consequently

$$\mathrm{Var}[X_{n_i}] = \mathbb{E}\big[\mathrm{dGum}(-x - y - c(x))^2\big] - \mathbb{E}\big[\mathrm{dGum}(-x - y - c(x))\big]^2 + o(1).$$

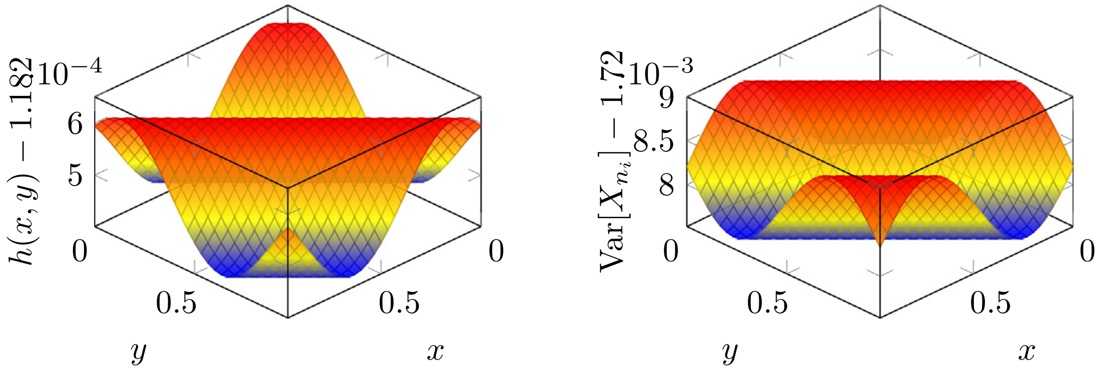

Fig. 2. Let $(n_i)_{i\in\mathbb{N}}$ be a sequence of natural numbers such that $\{\log_2 n_i\} \to x$ and $\{\ln n_i\} \to y$ for $x, y \in [0, 1)$. The left figure shows the function $h(x, y)$ (appearing in the expectation of X_{n_i}) for values of x and y between 0 and 1. The right figure shows $\mathrm{Var}[X_{n_i}]$ as a function of x, y.

To determine the expectation and variance of the runtime we need to consider various moments of the discrete Gumbel distribution. To this end, let X be an integer valued random variable with finite kth moment, then

$$\mathbb{E}\left[X^k\right] = \sum_{\ell\in\mathbb{Z}} \ell^k P[X = \ell] = \sum_{\ell\in\mathbb{Z}} \ell^k \big(P[X \le \ell] - P[X \le \ell - 1]\big),$$

and therefore, for all $\alpha \in \mathbb{R}$ and $k \in \mathbb{N}$,

$$\mathbb{E}\left[\mathrm{dGum}(\alpha)^k\right] = \sum_{\ell\in\mathbb{Z}} \ell^k \left(e^{-e^{-\ell-\alpha}} - e^{-e^{-\ell-\alpha+1}}\right).$$

This sum converges exponentially fast, both for $\ell \to \infty$ and $\ell \to -\infty$, and thus allows for effective numerical treatment. In summary, improving (1), we get for all $n \in \mathbb{N}$ the numerical bounds

$$\log_2 n + \ln n + 1.18242 \le \mathbb{E}[X_n] \le \log_2 n + \ln n + 1.18263,$$

as $\inf_{0\le x,y\le 1} h(x, y) = 1.18242\ldots$, $\sup_{0\le x,y\le 1} h(x, y) = 1.18262\ldots$ and

$$1.7277 \le \mathrm{Var}[X_n] \le 1.7289.$$

These numerical bounds are (essentially) best possible, see also Fig. 2. Higher moments can be estimated similarly. Let us close this section with a final remark on the function c defined in (2), as this might be helpful in future works. This function is defined as the limit of a sequence in two parameters a, b; the main reason for this is its combinatorial origin, which will become apparent in the proofs. However, all that is actually important is that b is large enough, in the sense that the difference $b - a \to \infty$ as $a \to \infty$. So, if we write h for an integer function that diverges to infinity, then we could define

$$d(x) = -x + \lim_{a\to\infty, a\in\mathbb{N}} h(a) + \ln\left(g^{(a+h(a))}(1 - 2^{-a-x})\right).$$

Then $c(x) = d(x)$ (which we state without proof, as we do not need it here), and c can be represented as a limit of an (one-dimensional) sequence.

Outline. In the next section we give an outline of the proof of our main results. At the beginning of the rumour spreading process *push* is dominated by an exponential growth of the informed nodes (Lemma 2.2). For the main part, where most nodes get informed, it closely follows a deterministic recursion (Lemma 2.1) and at the end it is described by a coupon collector type problem (Lemma 2.3). Based on these lemmas we give the rigorous proof of our claims in Section 3. The proof to these three important lemmas can also be found there, in Sections 3.3–3.5. Sections 3.6 and 3.7 contain all other proofs.

Further notation. Unless stated otherwise, all asymptotic behaviour in this paper is for $n \to \infty$. Consider a graph $G = (V, E)$. For $t \in \mathbb{N}_0$ $(= \mathbb{N} \cup \{0\})$ we denote by $I_t \subseteq V$ the set of informed nodes at the end of round t; in particular $|I_0| = 1$. Analogously we write $U_t = V \backslash I_t$ for the set of uninformed nodes. For an event A, we sometimes write $P_A[\cdot]$ instead of $P[\cdot \mid A]$ to denote the conditional probability and we write $\mathbb{E}_A[\cdot] = \mathbb{E}[\cdot \mid A]$. If we condition on I_t, then we also abbreviate $P[\cdot \mid I_t] = P_t[\cdot]$ and $\mathbb{E}[\cdot \mid I_t] = \mathbb{E}_t[\cdot]$.

2. Proof overview

Let us start the proof of Theorem 1.1 about the distribution of the runtime of *push* on K_n with a simple observation, that is more or less explicit also in previous works. Note that as long as the *total number* of pushes performed is $o(\sqrt{n})$, then whp no node will be informed twice – this is a simple consequence of the famous birthday paradox. That is, whp as long as $|I_t| = o(\sqrt{n})$, every node in I_t will inform a currently uninformed node and thus $|I_{t+1}| = 2|I_t|$. In particular, whp

$$|I_{t_0}| = 2^{t_0}, \quad \text{where} \quad t_0 := \lfloor 0.49 \cdot \log_2 n \rfloor. \tag{3}$$

Soon after round t_0 things get more complicated. We continue with a definition. Apart from the functions $g^{(i)}$ defined in the previous section, we will also need the following functions. Set

$$f = f^{(1)} : [0, 1] \to [0, 1], \quad x \mapsto 1 - e^{-x}(1 - x)$$

and

$$f^{(i)} : [0, 1] \to [0, 1], \quad f^{(i)} = f \circ f^{(i-1)}, i \geq 2.$$

Some elementary properties of f are: f is strictly increasing and concave, and $f^{(b)}(x) \to 1$ as $b \to \infty$ for all $x \in (0, 1]$. Moreover, $f^{(i)}(x) = 1 - g^{(i)}(1 - x)$ for all $x \in [0, 1]$ and $i \in \mathbb{N}$. It is also not difficult to establish, see also [23] and Lemma 3.5, that f captures the behaviour of the expected number of informed nodes after one round of the protocol. Moreover, $|I_{t+1}|$ is typically close to $f(|I_t|/n)n$. Here we will need a more explicit qualitative control of how $|I_t|$ behaves, since our aim is to specify the limiting distribution. We show the following statement, which implies that if we start in round t_0 (set $T = t_0$ in that lemma) then whp for *all* succeeding rounds $t_0 + t$ the number of informed nodes is close to $f^{(t)}(|I_{t_0}/n|)n$.

Lemma 2.1. *Let* $0 < c < 0.49$ *and* $T \geq c \log_2 n$. *Then*

$$P_T\left[\bigcap_{t \in \mathbb{N}_0} \left\{\left||I_{T+t}| - f^{(t)}\left(|I_T|/n\right) n\right| \leq n^{1-c/4}\right\}\right] = 1 - O(n^{-c^2/10}).$$

Thus, the key to understanding $|I_t|$ is to understand how f behaves when iterated very many times. Note that when the number of informed nodes is xn for some very small x, then the e^{-x} term in the definition of f can be approximated by $1-x$ and therefore $f(x) \approx 1-(1-x)^2 \approx 2x$.

This crude estimate suggests that the number of informed nodes doubles every round as long as there are only few informed nodes, and we know already that the doubling is perfect if $xn = o(\sqrt{n})$. Our next lemma actually shows that the doubling continues to be *almost* perfect, as long as the total number of nodes is not close to n.

Lemma 2.2. *Let $a, T \in \mathbb{N}$ be such that $2^{-a} < 0.1$ and $T \le \lfloor 0.49 \cdot \log_2 n \rfloor$. Set $t_1 := \lfloor \log_2 n \rfloor - a$. Then*

$$\left|2^{t_1} - f^{(t_1-T)}\left(2^T/n\right)n\right| \le 2^{-2a+1}n.$$

Combining the previous lemmas we have thus established that for any $a \in \mathbb{N}$ with $2^{-a} < 0.1$ whp

$$(1 - 2^{-a+2}) \cdot 2^{t_1} \le |I_{t_1}| \le 2^{t_1}, \quad t_1 := \lfloor \log_2 n \rfloor - a. \tag{4}$$

Here we can think of a being very large (but fixed) and then the two bounds are very close to each other; in particular, $|I_{t_1}| \approx 2^{\lfloor \log_2 n \rfloor - a}$ and thus I_t contains a linear number of nodes. Up to that point we have studied the behaviour of the process up to time t_1. Next we perform another b steps, where again b is fixed. Applying Lemma 2.1 once more and using that $f^{(b)}(x)$ is increasing and is less than 1 for $x < 1$ yields with room to spare that for $t_2 = t_1 + b$ whp

$$\left(1 - n^{-1/6}\right) f^{(b)}\left((1 - 2^{-a+2})2^{t_1}/n\right) \le n^{-1}\left|I_{t_2}\right| \le \left(1 + n^{-1/6}\right) f^{(b)}\left(2^{t_1}/n\right). \tag{5}$$

In essence, this says that if we write $x = \log_2 n - \lfloor \log_2 n \rfloor = \{\log_2 n\}$, then (we begin getting informal and obtain that)

$$|I_{t_2}| \approx f^{(b)}\left(2^{t_1}/n\right)n = f^{(b)}\left(2^{-a-x}\right)n, \quad \text{where} \quad t_2 = \lfloor \log_2 n \rfloor - a + b.$$

In particular, choosing a priori b large enough makes the fraction $|I_{t_2}|/n$ arbitrarily close to 1, that is, almost all nodes except for a tiny fraction are informed. All in all, up to time t_2 we have very fine control of the number of informed nodes, and we also see how the quantity $\{\log_2 n\}$ slowly sneaks in.

After time t_2 the behaviour changes once more. In this regime there is an interesting connection to the well-known Coupon Collector Problem (CCP), which was also exploited in [7]. In order to formulate the connection, note that the number of pushes that are needed to inform one uninformed node, having N informed nodes, is (in distribution) equal to the number of coupons needed to draw the $(N + 1)$st distinct coupon out of n. The CCP is very well understood, and it is a classic result that, appropriately normalized, the total number of coupons drawn (until all are collected) tends to a Gumbel distribution. However, translating the number of required pushes to the number of rounds – the quantity we are interested in – is not straightforward. In particular, the number of pushes in one round depends on the current number of informed nodes. On the other hand, after round t_2 there are $n - o(n)$ informed nodes, so that we may hope to approximate the remaining number of rounds with n^{-1} times the number of coupons in the CCP. The next lemma establishes the precise bridge between the two problems. There, for two sequences of random variables $(X_n)_{n\in\mathbb{N}}$ and $(Y_n)_{n\in\mathbb{N}}$ we write $X_n \precsim Y_n$ if there is a function $h : \mathbb{N} \to \mathbb{R}^+$ with $h = o(1)$ such that $P[X_n \ge x] \le P[Y_n \ge x] + h(n)$ for all $n \in \mathbb{N}, x \in \mathbb{R}$; $X_n \succsim Y_n$ is defined with "$\ge$" instead of "$\le$".

Lemma 2.3. *Let $G \sim \mathrm{Gum}(\gamma)$, $b > 2a \in \mathbb{N}$ and assume that $\ell \cdot n \le |I_{\lfloor \log_2 n \rfloor - a + b}| \le u \cdot n$ for some $\ell, u \in [0, 1)$. Then*

$$X_n - \lfloor \log_2 n \rfloor + a - b \succsim \left\lceil \ln n + \ln\left(\frac{1}{u} - 1\right) + \gamma \right\rceil$$

and

$$X_n - \lfloor \log_2 n \rfloor + a - b \precsim \left\lceil \ln n + \ln\left(\frac{1}{\ell} - 1\right) + \ln\left(\frac{\ell}{e\ell - e + 1}\right) + \gamma + G \right\rceil.$$

Note that the previous discussion guarantees that ℓ, u in Lemma 2.3 are very close to 1 and very close to each other. So, the term $\ln(\ell/(e\ell - e + 1))$ is very close to 0. We obtain that in distribution

$$X_n - \lfloor \log_2 n \rfloor + a - b \approx \left\lceil \ln n + \ln\left(\frac{1}{u} - 1\right) + \gamma + G \right\rceil, \quad \text{where} \quad u = f^{(b)}(2^{-a-x}).$$

and equivalently with $x = \log_2 n - \lfloor \log_2 n \rfloor$

$$X_n \approx \left\lceil \log_2 n + \ln n - a + b + \ln\left(g^{(b)}(2^{-a-x})\right) - x + \gamma + G \right\rceil. \tag{6}$$

Here we now encounter the mysterious function c from (2). The next lemma collects some important properties of it that will turn out to be very helpful.

Lemma 2.4. *The function*

$$c(x) = \lim_{a\to\infty, a\in\mathbb{N}} \lim_{b\to\infty, b\in\mathbb{N}} -a + b + \ln\left(g^{(b)}(1 - 2^{-a-x})\right) - x$$

is well-defined, continuous and periodic with period 1.

With all these facts at hand, the proof of Theorem 1.1 is completed by considering the random variable on the right-hand side of (6); in particular, the dependence on $y = \ln n - \lfloor \ln n \rfloor$ arises naturally. The complete details of the proof, which is based on Lemmas 2.1–2.3 and follows the strategy outlined here can be found in Section 3 (together with the proofs of the lemmas).

As described in the introduction, apart from the limiting distribution we are interested in the asymptotic expectation of the runtime. A key ingredient towards the proof of Theorem 1.3 is uniform integrability, which can be shown by using the distributional bounds from [7]. Uniform integrability is a sufficient condition that convergence in distribution also implies convergence of the means.

Lemma 2.5 (*Uniform Integrability*)**.** *Let $k \in \mathbb{N}$ and set $Y_n := X_n - \lfloor \log_2 n \rfloor - \lfloor \ln n \rfloor$. Then Y_n^k is uniformly integrable, that is*

$$\lim_{N\to\infty} \sup_{n\in\mathbb{N}} \mathbb{E}\left[|Y_n|^k \;\middle|\; \mathbb{1}\left[|Y_n|^k > N\right]\right] = 0.$$

3. Proof of the main result

In this section we complete the proof of Theorem 1.1 outlined in Section 2. Afterwards we give the (short) proofs for Theorems 1.2 and 1.3.

3.1. Proof of Theorem 1.1

As the outline was indeed rigorous until (5) we take the proof up from there. Choose the quantities $a, b \in \mathbb{N}$ such that $2a < b$ and recall that $t_1 = \lfloor \log_2 n \rfloor - a$. Set furthermore for brevity

$$\ell = \left(1 - n^{-1/6}\right) f^{(b)}\left((1 - 2^{-a+2})2^{t_1}/n\right) \quad \text{and} \quad u = \left(1 + n^{-1/6}\right) f^{(b)}\left(2^{t_1}/n\right).$$

Then (5) states that, for $t_2 = \lfloor \log_2 n \rfloor - a + b$,

$$\ell \le n^{-1} \left| I_{t_2} \right| \le u,$$

and Lemma 2.3 yields, for $Y_n = X_n - \lfloor \log_2 n \rfloor + a - b$, that

$$Y_n \precsim \left\lceil \ln n + \ln\left(\frac{1}{\ell} - 1\right) + \ln\left(\frac{\ell}{e\ell - e + 1}\right) + \gamma + G \right\rceil$$

and

$$Y_n \succsim \left\lceil \ln n + \ln\left(\frac{1}{u} - 1\right) + \gamma + G \right\rceil.$$

The next lemma establishes that both ℓ, u tend to 1 as a gets large, and moreover that the difference $\ln(1/\ell - 1) - \ln(1/u - 1)$ can be made arbitrarily small. Its proof can be found in Section 3.7.

Lemma 3.1. *For ℓ, u defined as above, where $b > 2a$*

$$\lim_{a\to\infty} \sup_{n\in\mathbb{N}} |\ln \ell| = \lim_{a\to\infty} \sup_{n\in\mathbb{N}} |\ln u| = \lim_{a\to\infty} \sup_{n\in\mathbb{N}} \left| \ln\left(\frac{\ell}{e\ell - e + 1}\right) \right| = 0.$$

Furthermore,

$$\lim_{a\to\infty} \sup_{n\in\mathbb{N}} |\ln(1-\ell) - \ln(1-u)| = 0.$$

Thus, as $n \to \infty$,

$$\ln(1-u) = \ln\left(1 - f^{(b)}\left(2^{t_1}/n\right)\right) + o(1) = \ln\left(g^{(b)}\left(1 - 2^{-a-\{\log_2 n\}}\right)\right) + o(1).$$

Let $\varepsilon > 0$. Lemma 3.1 readily implies that there are $a_0, n_0 \in \mathbb{N}$ such that for all $a > a_0$ and $n > n_0$,

$$Y_n \succsim \lceil \ln n + \ln\left(g^{(b)}\left(1 - 2^{-a-\{\log_2 n\}}\right)\right) + \gamma + G - \varepsilon \rceil$$

and similarly also

$$Y_n \precsim \lceil \ln n + \ln\left(g^{(b)}\left(1 - 2^{-a-\{\log_2 n\}}\right)\right) + \gamma + G + \varepsilon \rceil.$$

Lemma 2.4 guarantees that there is an $a_1 \ge a_0$ such that for all $a \ge a_1$

$$\left| \ln\left(g^{(b)}\left(1 - 2^{-a-\{\log_2 n\}}\right)\right) - a + b - (c(\{\log_2 n\}) + \{\log_2 n\}) \right| \le \varepsilon.$$

Thus for all $a > a_1$ and $n > n_0$

$$X_n \succsim \lceil \log_2 n + \ln n + c(\{\log_2 n\}) + \gamma + G - 2\varepsilon \rceil,$$

as well as

$$X_n \precsim \lceil \log_2 n + \ln n + c(\{\log_2 n\}) + \gamma + G + 2\varepsilon \rceil.$$

Thus we are left with getting rid of the ε terms in the previous equations. The following lemma accomplishes exactly that and therefore implies the claim of Theorem 1.1. Its proof can be found in Section 3.7.

Lemma 3.2. *Let $h : \mathbb{N} \to \mathbb{R}^+$ and $G \sim \mathrm{Gum}(\gamma)$. Then*

$$\forall \varepsilon > 0 : X_n \precsim \lceil h(n) + G + \varepsilon \rceil \implies X_n \precsim \lceil h(n) + G \rceil.$$

The respective statement also holds for "$\succsim$".

3.2. Proof of Theorems 1.2 and 1.3

Proof of Theorem 1.2. Recall that $\{z\} = z - \lfloor z \rfloor$, $z \in \mathbb{R}$. Let $(n_i)_{i\in\mathbb{N}}$ be a strictly increasing subsequence of $\mathbb{N}$ such that $\{\log_2 n_i\} \to x$ and $\{\ln n_i\} \to y$. Substituting $k = \lfloor \log_2 n_i \rfloor + \lfloor \ln n_i \rfloor + 1 + t$ for any $t \in \mathbb{Z}$ we get that

$$\begin{aligned} & P\Big[\lceil G + \log_2 n_i + \ln n_i + \gamma + c(\{\log_2 n_i\})\rceil \geq k\Big] \\ &= P\Big[\lceil G + \log_2 n_i + \ln n_i + \gamma + c(\{\log_2 n_i\})\rceil \geq \lfloor \log_2 n_i \rfloor + \lfloor \ln n_i \rfloor + 1 + t\Big] \\ &= P\Big[\lceil G + \{\log_2 n_i\} + \{\ln n_i\} + \gamma + c(\{\log_2 n_i\})\rceil > t\Big] \\ &= P\Big[G + \{\log_2 n_i\} + \{\ln n_i\} + \gamma + c(\{\log_2 n_i\}) > t\Big]. \end{aligned}$$

Thus using Theorem 1.1, Lemma 2.4 and Lemma 3.2 we get that, as $i \to \infty$,

$$\sup_{t\in\mathbb{Z}} \Big| P\big[X_{n_i} \geq \lfloor \log_2 n_i \rfloor + \lfloor \ln n_i \rfloor + 1 + t\big] - P\big[G + x + y + \gamma + c(x) > t\big]\Big| = o(1).$$

Using the distribution function of $G \sim \mathrm{Gum}(\gamma)$ we get

$$P\big[X_{n_i} \geq \lfloor \log_2 n_i \rfloor + \lfloor \ln n_i \rfloor + 1 + t\big] \stackrel{i\to\infty}{\longrightarrow} 1 - \exp\big(-\exp(-t + x + y + c(x))\big),$$

that is,

$$P\big[X_{n_i} \leq \lfloor \log_2 n_i \rfloor + \lfloor \ln n_i \rfloor + t\big] \stackrel{i\to\infty}{\longrightarrow} P(\mathrm{dGum}(-x - y - c(x)) \leq t). \quad \square$$

Next we prove Theorem 1.3.

Proof of Theorem 1.3. Lemma 2.5 states that $\big(X_n - \lfloor \log_2 n \rfloor - \lfloor \ln n \rfloor\big)^k$ is uniformly integrable and Theorem 1.2 established its convergence in distribution to $\big(\mathrm{dGum}(-x - y - c(x))\big)^k$. Together this implies

$$\mathbb{E}\Big[\big(X_n - \lfloor \log_2 n \rfloor - \lfloor \ln n \rfloor\big)^k\Big] \to \mathbb{E}\Big[\big(\mathrm{dGum}(-x - y - c(x))\big)^k\Big]. \quad \square$$

3.3. Proof of Lemma 2.1

The number of informed nodes, $|I_t|$, fulfils a so-called self-bounding property, for reference see [2]. One striking consequence thereof is the following bound.

Lemma 3.3 *([5]). For any $t \in \mathbb{N}$,*

$$\mathrm{Var}\big[|I_{t+1}| \mid I_t\big] \leq \mathbb{E}\big[|I_{t+1}| \mid I_t\big].$$

This bound on the variance and Chebychev's inequality ensure that the number of informed nodes is highly concentrated around its expectation as soon as enough nodes are informed. Moreover, even stronger concentration results are possible, as self-bounding functions admit exponential concentration inequalities, see e.g. [2]. Here, Chebyshev is sufficient for our application.

Lemma 3.4. *Let $0 < c \leq 1$, let $t_0 \in \mathbb{N}$ and assume that $|I_{t_0}| \geq n^c$. For $t \in \mathbb{N}$ and $\varepsilon > 0$ let C_t denote the event that*

$$\big||I_{t+1}| - \mathbb{E}_t[|I_{t+1}|]\big| \leq (\mathbb{E}_t[|I_{t+1}|])^{1/2+\varepsilon}.$$

Then

$$P_{t_0}\left[\bigcap_{t\geq t_0} C_t\right] = 1 - O\left(n^{-c\varepsilon}\right).$$

Proof. From [7, Corollary 3.2] it is known that for any $r > 0$

$$P\big[X_n \geq \lceil \log_2 n\rceil + \ln n + 2.188 + r\big] \leq 2e^{-r}.$$

Thus it suffices (with lots of room to spare) to show

$$P_{t_0}\left[\bigcup_{t_0\leq t\leq \log^2 n} \overline{C_t}\right] = O\left(n^{-3c\varepsilon/2}\right). \tag{7}$$

By using Chebychev's inequality and Lemma 3.3,

$$P_t\big[\overline{C_t}\big] = P_t\left[\big||I_{t+1}| - \mathbb{E}_t[|I_{t+1}|]\big| > \mathbb{E}_t[|I_{t+1}|]^{1/2+\varepsilon}\right] \leq \frac{\mathrm{Var}[|I_{t+1}|]}{\mathbb{E}_t[|I_{t+1}|]^{1+2\varepsilon}}$$
$$\leq \mathbb{E}_t[|I_{t+1}|]^{-2\varepsilon}.$$

Since $\mathbb{E}_t[|I_{t+1}|] \geq |I_{t+1}| \geq |I_{t_0}|$ the claim follows from (7) and the union bound. $\square$

Lemma 3.5 establishes a connection between the expected value of $|I_{t+1}|$ and our previously defined function f, see below Eq. (3). This has also been observed (though not so precise) in [23] and we include a quick proof for completeness.

Lemma 3.5. *Let $t \in \mathbb{N}$ and $n \geq 3$. Then*

$$f(|I_t|/n)n \leq \mathbb{E}_t|I_{t+1}| \leq f(|I_t|/n)n + 5.$$

Proof. Each uninformed node $u \in U_t$ remains uninformed if all $|I_t|$ informed nodes do not push to u. Since all these events are independent, we obtain that the probability that u remains uninformed is $(1 - 1/(n-1))^{|I_t|}$. Thus by linearity of expectation

$$\mathbb{E}_t[|I_{t+1}|] = |I_t| + \big(n - |I_t|\big)\left(1 - \left(1 - \frac{1}{n-1}\right)^{|I_t|}\right)$$
$$= n - \big(n - |I_t|\big)\left(1 - \frac{1}{n-1}\right)^{|I_t|}.$$

For a lower bound we use $1 - x \leq e^{-x}$ and get

$$\mathbb{E}_t[|I_{t+1}|] \geq n - \big(n - |I_t|\big)e^{-|I_t|/(n-1)} \geq n - \big(n - |I_t|\big)e^{-|I_t|/n} = f(|I_t|/n)n.$$

For an upper bound we use $1 - x \geq e^{-x-x^2}$ for all $x \leq 1/2$

$$\mathbb{E}_t[|I_{t+1}|] \leq n - \big(n - |I_t|\big)e^{-|I_t|/(n-1)-|I_t|/(n-1)^2}$$
$$\leq n - \big(n - |I_t|\big)e^{-|I_t|/n}\exp\left(-\frac{2|I_t|}{(n-1)^2}\right)$$

and again using $1 - x \leq e^{-x}$

$$\mathbb{E}_t[|I_{t+1}|] \leq n - \big(n - |I_t|\big)e^{-|I_t|/n}\left(1 - \frac{2|I_t|}{(n-1)^2}\right) \leq f(|I_t|/n)n + 5. \quad \square$$

Lemma 3.6 is an auxiliary result that we use in the proof of Lemma 2.1. It shows that f is concave and has decreasing derivative on the interval $[0, 1]$, the stated property is a direct consequence.

Lemma 3.6. *Let* $0 < x_1 \le x_2 < 1$. *Then* $|f(x_2) - f(x_1)| \le (2 - x_1)e^{-x_1}(x_2 - x_1)$.

Proof. It is $f'(x) = (2 - x)e^{-x}$ and $f''(x) = (x - 3)e^{-x}$; in particular, f' is monotonically decreasing and takes only positive values on $[x_1, x_2]$. Furthermore

$$\max_{x\in[x_1,x_2]} f'(x) = (2 - x_1)e^{-x_1}$$

and therefore, as a direct consequence of the mean value theorem, we have

$$|f(x_2) - f(x_1)| \le (x_2 - x_1) \max_{x\in[x_1,x_2]} f'(x) = (2 - x_1)e^{-x_1}(x_2 - x_1). \qquad \square$$

We state a simple corollary for later reference.

Corollary 3.1. *Let* $i \in \mathbb{N}$ *and* $r, s \in [0, 1/2]$. *Then* $f^{(i)}(r + s) \le f^{(i)}(r) + 2^i s$.

Having these lemmas as ingredients we can prove the main result of this subsection. Lemma 3.5 shows that the expectation of $|I_{t+1}|$ is given by $f(|I_t|/n)n$ and Lemma 3.4 shows that $|I_{t+1}|$ is closely concentrated around its expectation in (nearly) all rounds. To then prove that $|I_{t+\tau}|$ is close to $f^{(\tau)}(|I_t|/)n$ for any $\tau \in \mathbb{N}$ we need to make sure that the errors in the concentration and the approximation of the expectation are not blown up by repeated applications of f. We will show that f can indeed increase the error in each step by a factor that can be as large as $\sqrt{2}$, but luckily this only happens when $|I_{t+\tau}| = o(n)$ and thus the accumulated error will remain small (as $|I_t|$ nearly doubles in this regime).

Proof of Lemma 2.1. Let $0 < \varepsilon < c/10$, and assume, with foresight, that $n \ge n_0$, where n_0 satisfies the inequalities

$$\sqrt{2} + 10n_0^{-\varepsilon} < \sqrt{2 + \varepsilon} \quad \text{and} \quad n_0^c \ge 25.$$

As $T \ge c \log_2 n$ and because of (3), that is, $|I_t| = 2^t$ for all $t \le \lfloor 0.49 \log_2 n \rfloor$, we have $|I_T| \ge n^c$. Consequently we can apply Lemma 3.4 and thus get with probability $1 - O(n^{-c\varepsilon})$

$$\big||I_{t+1}| - \mathbb{E}_t[|I_{t+1}|]\big| \le \mathbb{E}_t[|I_{t+1}|]^{1/2+\varepsilon}, \qquad \text{for all } t \ge T. \tag{8}$$

For the rest of this proof we assume that (8) holds. Set

$$\alpha_{T+t} = f^{(t)}(|I_T|/n), \quad t \in \mathbb{N}_0.$$

We will first argue that

$$\big||I_t| - \alpha_t n\big| \le \alpha_t^{1/2+\varepsilon} n^{1/2+2\varepsilon} \sqrt{2 + \varepsilon}^{\,t-T} =: d_t. \tag{9}$$

for all $t \ge T$ such that $d_t \le n^{1-\varepsilon}$. Note that this is obviously true for $t = T$. For the induction step we argue that

$$\big||I_{t+1}| - \alpha_{t+1} n\big| \le \alpha_{t+1}^{1/2+\varepsilon} n^{1/2+2\varepsilon} \sqrt{2 + \varepsilon}^{\,t+1-T} = d_{t+1}. \tag{10}$$

To see this, we use Lemma 3.5, (8) and the fact that $|I_{t+1}| \le 2|I_t|$ (in this order) to obtain the bound

$$\big||I_{t+1}| - f(|I_t|/n)n\big| \le \big||I_{t+1}| - \mathbb{E}_t[|I_{t+1}|]\big| + 5 \le (2|I_t|)^{1/2+\varepsilon} + 5.$$

Then we apply Lemma 3.6 to estimate the difference of $f(|I_t|/n)$ and $\alpha_{t+1} = f(\alpha_t)$, and infer from (9), using $e^x \le 1 + 2x$ for all $0 \le x \le 1$, that

$$\begin{aligned}\left|f(|I_t|/n)n - \alpha_{t+1}n\right| &\le \left||I_t| - \alpha_t n\right|\big(2 - \min\{\alpha_t, |I_t|/n\}\big)e^{-\min\{\alpha_t,|I_t|/n\}} \\ &\le d_t\big(2 - \alpha_t + d_t/n\big)e^{-\min\{\alpha_t,|I_t|/n\}} \\ &\le d_t\big(2 - \alpha_t\big)e^{-\alpha_t + d_t/n} + d_t^2/n \le d_t(2-\alpha_t)e^{-\alpha_t} + 5d_t^2/n.\end{aligned}$$

All in all we have argued that for all t such that $d_t \le n^{1-\varepsilon}$

$$\begin{aligned}\left||I_{t+1}| - \alpha_{t+1}n\right| &\le \left||I_{t+1}| - f(|I_t|/n)n\right| + \left|f(|I_t|/n)n - \alpha_{t+1}n\right| \\ &\le (2|I_t|)^{1/2+\varepsilon} + 5 \ + \ d_t(2-\alpha_t)e^{-\alpha_t} + 5d_t^2/n \\ &\le 2(\alpha_t n + d_t)^{1/2+\varepsilon} + 5 + d_t(2-\alpha_t)e^{-\alpha_t} + 5d_t n^{-\varepsilon}.\end{aligned}$$

Our assumptions on ε and n imply that $d_t^{1/2+\varepsilon} \le d_t n^{-\varepsilon}$. Moreover, $\alpha_T n \ge n^c \ge 25$, and thus

$$\begin{aligned}\left||I_{t+1}| - \alpha_{t+1}n\right| &\le 3(\alpha_t n)^{1/2+\varepsilon} + d_t(2-\alpha_t)e^{-\alpha_t} + 7d_t n^{-\varepsilon} \\ &\le d_t(2-\alpha_t)e^{-\alpha_t} + 10d_t n^{-\varepsilon}.\end{aligned} \tag{11}$$

To understand (11) consider the auxiliary function

$$H(x) = \sqrt{\frac{f(x)}{x}} - \frac{f'(x)}{\sqrt{2}} = \sqrt{\frac{1-(1-x)e^{-x}}{x}} - \frac{(2-x)e^{-x}}{\sqrt{2}}.$$

As $(1-x)e^{-x} = 1 - 2x + O(x^2)$ as $x \to 0$ we have that $\lim_{x\to 0}(1-(1-x)e^{-x})/x = 2$ and thus $\lim_{x\to 0} H(x) = 0$. Furthermore is H an increasing function on the interval $[0, 1]$ as,

$$H' = \frac{1}{2}\left(\frac{2(1-x)e^{-x}}{x^2}\right)^{-1/2} + \frac{(3-x)e^{-x}}{\sqrt{2}} \ge 0 \quad \text{for} \quad x \le 1.$$

Therefore $H(x) \ge 0$ for all $0 \le x \le 1$ and consequently, using $\alpha_{t+1} > \alpha_t$,

$$\left(\frac{\alpha_{t+1}}{\alpha_t}\right)^{1/2+\varepsilon} \ge \left(\frac{\alpha_{t+1}}{\alpha_t}\right)^{1/2} \ge \frac{(2-\alpha_t)e^{-\alpha_t}}{\sqrt{2}}.$$

Since $d_t = \alpha_t^{1/2+\varepsilon} n^{1/2+2\varepsilon}\sqrt{2+\varepsilon}^{\,t-T}$, applying the previous bound to (11) implies (10) for all $n \ge n_0$, that is, all n such that $\sqrt{2} + 10n^{-\varepsilon} < \sqrt{2+\varepsilon}$. This completes the induction step and the proof of (9) is completed.

Actually our arguments yield also the following statement, which is stronger than (9) when there are "many" informed nodes. In particular, for all $t' \in \mathbb{N}$ such that $(2-\alpha_{t'})e^{-\alpha_{t'}} < 1-\varepsilon$ Eq. (11) also yields for all $n \ge n_0$

$$\left||I_{T+t'}| - \alpha_{T+t'}n\right| \le d_{t'} \quad \Rightarrow \quad \left||I_{T+t'+1}| - \alpha_{T+t'+1}n\right| \le d_{t'},$$

meaning that the absolute error does not increase any more after round t'. (Actually the error decreases by a factor of at least ε after that round, but we do not need this.) To complete the proof we show that we can choose t' such that $d_{t'} \le n^{1-c/4}$ and $(2-\alpha_{t'})e^{-\alpha_{t'}} < 1-\varepsilon$. To this end, consider

$$T' = \lfloor \log_2 n \rfloor - 4 - T$$

and applying Lemma 2.2 to $\alpha_{T'}$ yields

$$\alpha_{T+T'} = f^{(T')}(|I_T|/n) \ge f^{(\lfloor \log_2 n\rfloor - 4 - T)}\big(2^T/n\big) \ge 2^{-4}\big(1 - 2^{-8+1}\big)$$

and furthermore, a simple computation yields that $\alpha_{T+T'+5} \geq 3/4$. Thus

$$(2-\alpha_{T+T'+5})e^{-\alpha_{T+T'+5}} \leq (2-3/4)e^{-3/4} < 1-\varepsilon$$

and we set $t' := T'+5$. Moreover,

$$d_{t'} \leq n^{1/2+2\varepsilon}\sqrt{2+\varepsilon}^{t'} \leq (2+\varepsilon)n^{1/2+2\varepsilon+(1-c)\log_2(2+\varepsilon)/2}.$$

Note that $\log_2(2+\varepsilon) \leq 1+\varepsilon$. Plugging this into the exponent yields that if $\varepsilon < c/10$ and n is large enough then $d_{t'} \leq n^{1-c/4}(\leq n^{1-\varepsilon})$, as claimed. □

3.4. *Proof of* Lemma 2.2

We begin with showing the basic inequality

$$2x(1-x) \leq f(x) \leq 2x. \tag{12}$$

To see this, note that $e^{-x} \leq 1-x+x^2/2$ for $x \in [0,1]$ and so

$$f(x) = 1-e^{-x}(1-x) \geq 1-\left(1-x+\frac{x^2}{2}\right)(1-x) \geq x\left(2-\frac{3}{2}x\right) \geq 2x-2x^2,$$

which establishes the first inequality in (12). The other inequality follows directly from the simple bound $e^{-x} \geq 1-x$.

Let us write $z_0 = 2^{t_0}/n$ and $z_i = f(z_{i-1}) = f^{(i)}(z_0)$; we want to bound $z_{t_1-t_0}$, where $t_1 = \lfloor \log_2 n \rfloor - a$ and $t_0 \leq \lfloor 0.49 \log_2 n \rfloor$. Clearly $z_i \leq 2^i z_0$, which shows the upper bound in Lemma 2.2. Using (12) we obtain by induction

$$z_i \geq 2^i z_0 \cdot \prod_{j=0}^{i-1}(1-2^j z_0), \quad i \in \mathbb{N}.$$

Further, using the bound $1-x \geq e^{-x-x^2/2(1-x)}$, valid for any $x \in [0,1)$ we obtain

$$z_i \geq 2^i z_0 \cdot \exp\left\{-z_0 \sum_{0\leq j<i} 2^j - z_0^2 \sum_{0\leq j<i} \frac{4^j}{2(1-2^j z_0)}\right\}$$

Note that our assumptions guarantee that $2^{t_1-t_0}z_0 = 2^{-a} < 0.1$, and so for any $1 \leq i \leq t_1-t_0$

$$z_i \geq 2^i z_0 \cdot \exp\left\{-2^i z_0 - (2^i z_0)^2\right\} \geq 2^i z_0 \cdot (1-2^{-a}-2^{-2a}).$$

Finally note that $1-y-y^2 \geq 1-2y$ for any $y \in [0,1]$, and so the last term is bounded by $2^i z_0 \cdot (1-2^{-2a+1})$, which coincides with the lower bound claimed in Lemma 2.2.

Corollary 3.2. *For all $x \in [0,1]$ and $i \in \mathbb{N}$*

$$2^i x\left(1-2^i x-2^{2i}x^2\right) \leq f^{(i)}(x) \leq 2^i x.$$

3.5. *Proof of* Lemma 2.3

A main tool in the forthcoming proof is the following result, which states that a sum of normalized independent geometric random variables converges to a Gumbel distributed random variable.

Theorem 3.7 *([8]). Let $T_1, \dots, T_{n-1}$ be independent random variables such that $T_i \sim \mathrm{Geo}((n-i)/(n-1))$ for $1 \le i < n$. Then, in distribution*

$$n^{-1} \sum_{1\le i<n} \big(T_i - \mathbb{E}[T_i]\big) \to \mathrm{Gum}(\gamma).$$

Unfortunately we cannot apply directly Theorem 3.7 to our setting, as we will have to deal with a sum of independent geometric random variables that are not normalized with the 'correct' factor n^{-1}. However, the next well-known statement assures that if the error is small enough we will still converge to the same limiting distribution.

Theorem 3.8 (*Slutsky's Theorem, See, e.g., [27, p. 19]*)**.** *Let $(X_n)_{n\in\mathbb{N}}$, $(Y_n)_{n\in\mathbb{N}}$ and $(Z_n)_{n\in\mathbb{N}}$ be sequences of real-valued random variables. Suppose that $X_n \to X$ in distribution and that there are constants $a, b \in \mathbb{R}$ such that $Y_n \to a$ and $Z_n \to b$ in probability. Then $Y_n X_n + Z_n \to aX + b$ in distribution.*

We now show a more general version of Theorem 3.7 that is applicable to our setting.

Lemma 3.9. *Let $T_1, \dots, T_{n-1}$ be independent random variables such that $T_i \sim \mathrm{Geo}((n-i)/(n-1))$ for $1 \le i < n$. Let furthermore $\varepsilon > 0$ and $s : \mathbb{N} \to [1, n]$ be a function such that $s(n-i) \ge \big(1 - o(1)\big)(n - c\cdot i)$ for any positive integer $i < \varepsilon n$. Then, in distribution*

$$\sum_{(1-\varepsilon)n\le i<n} \frac{T_i - \mathbb{E}[T_i]}{s(i)} \to \mathrm{Gum}(\gamma).$$

Proof. Let $D_i = T_i - \mathbb{E}[T_i]$ be the centralized version of T_i. Then

$$\sum_{(1-\varepsilon)n\le i<n} \frac{D_i}{s(i)} = \sum_{1\le i<n} \frac{D_i}{n} - \sum_{1\le i<(1-\varepsilon)n} \frac{D_i}{n} + \sum_{(1-\varepsilon)n\le i<n} \left(\frac{D_i}{s(i)} - \frac{D_i}{n}\right).$$

A direct application of Theorem 3.7 guarantees that the first sum converges to $\mathrm{Gum}(\gamma)$ in distribution. To complete the proof it is sufficient to argue that in probability

$$\sum_{1\le i<(1-\varepsilon)n} \frac{D_i}{n} \to 0 \quad \text{and} \quad \sum_{(1-\varepsilon)n\le i<n} \left(\frac{D_i}{s(i)} - \frac{D_i}{n}\right) \to 0, \tag{13}$$

from which the claim in the lemma follows immediately from Theorem 3.8. Since the D_i's are centralized

$$\mathbb{E}\left[\sum_{1\le i<(1-\varepsilon)n} \frac{D_i}{n}\right] = 0,$$

and using that $\mathrm{Var}[T_i] = \big((n-1)(i-1)\big)/(n-i)^2$ for all $i < n$

$$\begin{aligned}\mathrm{Var}\left[\sum_{1\le i<(1-\varepsilon)n} \frac{D_i}{n}\right] &= \sum_{1\le i<(1-\varepsilon)n} \frac{\mathrm{Var}[T_i]}{n^2} = \sum_{1\le i<(1-\varepsilon)n} \frac{1}{n^2}\frac{(n-1)(i-1)}{(n-i)^2} \\ &\le \sum_{1\le i<(1-\varepsilon)n} \frac{1}{(\varepsilon n)^2} = o(1).\end{aligned}$$

Thus Chebychev's inequality directly implies that

$$\sum_{1\le i<(1-\varepsilon)n} \frac{D_i}{n} \to 0 \quad \text{in probability.}$$

It remains to treat the second term in (13). We compute the variance as before

$$\begin{aligned}\mathrm{Var}\left[\sum_{(1-\varepsilon)n\le i<n} \frac{D_i}{s(i)} - \frac{D_i}{n}\right] &= \sum_{(1-\varepsilon)n\le i<n} \left(\frac{1}{s(i)} - \frac{1}{n}\right)^2 \frac{(n-1)(i-1)}{(n-i)^2} \\ &\le \sum_{1\le i\le \varepsilon n} \left(\frac{1}{s(n-i)} - \frac{1}{n}\right)^2 \frac{n^2}{i^2}.\end{aligned}$$

However, this is also $o(1)$, as $s(n-i) \ge \big(1-o(1)\big)(n-c\cdot i)$ for all integers $i \le \varepsilon n$ by assumption, and therefore

$$0 \le \frac{1}{s(i)} - \frac{1}{n} \le \frac{1}{(1+o(1))(n-c\cdot i)} - \frac{1}{n} = (1+o(1))\frac{c\cdot i + o(n)}{n^2}, \quad i \le \varepsilon n.$$

In summary we have shown that

$$\mathrm{Var}\left[\sum_{(1-\varepsilon)n\le i<n} \left(\frac{D_i}{s(i)} - \frac{D_i}{n}\right)\right] = o(1)$$

and clearly

$$\mathbb{E}\left[\sum_{(1-\varepsilon)n\le i<n} \left(\frac{D_i}{s(i)} - \frac{D_i}{n}\right)\right] = 0.$$

Thus Chebychev's inequality implies also the second statement in (13) and the proof is complete. □

A further ingredient that we shall exploit is the following fact. If a sequence of random variables $X_n \to X$ in distribution with distribution functions $F_n \to F$ and if F is continuous everywhere, then the convergence of F_n to F is even uniform.

Theorem 3.10 (*Polya's Theorem, [24, Theorem 1]*)**.** *For each $n \in \mathbb{N}$ let X_n be a real-valued random variable with distribution function F_n. Assume that $X_n \to X$ in distribution. If X has continuous distribution function F, then*

$$\lim_{n\to\infty} \sup_{x\in\mathbb{R}} |F_n(x) - F(x)| = 0.$$

We need one more auxiliary lemma that gives an upper bound on the informed nodes when going one round backwards in order to later convert the number of Coupons into the number of rounds that are needed to finish the protocol. Appropriately, Lemma 3.4 assures that in all rounds the number of informed nodes is tightly concentrated around its expectation, which in turn is described by f, thus applying f^{-1} will give a good bound.

Lemma 3.11. *Let $t_0 \in \mathbb{N}$ and $0 < \varepsilon < 1/6$. Let C_t be the event that $\big||I_{t+1}| - E_t[|I_{t+1}|]\big| \le (E_t[|I_{t+1}|])^{1/2+\varepsilon}$, as given in Lemma 3.4. Then for n large enough the event $\bigcap_{t\ge t_0} C_t$ implies* *for all* $t \geq t_0$

$$|I_t| \geq \left(1 - n^{-1/3}\right) \cdot e \cdot \left(|I_{t+1}| - (1 - 1/e)n\right).$$

Proof. Lemma 3.5 and C_t together give that

$$|I_{t+1}| \leq E_t[|I_{t+1}|]| + (E_t[|I_{t+1}|])^{1/2+\varepsilon} = f(|I_t|/n)n + o\left(n^{2/3}\right).$$

Using the definition of $f(x) = 1 - (1-x)e^{-x}$ and that $|I_t| \leq n$ for all t we get that

$$|I_{t+1}| \leq n - e^{-|I_t|/n}(n - |I_t|) + o\left(n^{2/3}\right) \leq (1 - 1/e)n + |I_t|/e + o\left(n^{2/3}\right).$$

Rearranging yields the claimed statement. □

Let us briefly outline the proof of Lemma 2.3. We have already shown bounds for the number of informed nodes after $\lfloor \log_2 n \rfloor - a + b$ rounds in (5). Starting from these bounds we will use the Coupon Collector Problem to compute the number of *pushes* that are needed to inform all remaining uninformed nodes. This will yield sums of independent geometric random variables (one summand for each uninformed node). Using Lemma 3.11 we will translate these numbers of pushes into numbers of rounds, which results in an almost correctly normalized sum of geometric random variables that Lemma 3.9 assures to converge to a Gumbel distribution. We will end up with upper and lower bounds to the distribution function of *push*.

Proof of Lemma 2.3. In this proof we will establish a connection between the Coupon Collector Problem and the behaviour of *push*. Let $v \in V$ be the node that was initially informed. Instead of every informed node choosing one of its neighbours uniformly at random, we now assume that it samples one node in $V \setminus \{v\}$ uniformly at random. This defines an equivalent model, as for all $u \in V$ the probability to choose any specific node in $V \setminus \{u, v\}$ does not change (it equals $1/(n-1)$ in both models) and choosing u or v makes no difference for the distribution of the set of informed nodes. Thus *push* is the same as drawing coupons out of a pool of $n-1$ different coupons, but doing so in batches with size being the number of distinct coupons already collected plus one, the 'plus one' representing the initially informed node v. It is widely known and easy to see that assuming $1 \leq i \leq n-1$ coupons (including v) have already been collected, then

$$T_i \sim \text{Geo}\left(\frac{n-i}{n-1}\right), \quad 1 \leq i \leq n-1. \tag{14}$$

describes the number of coupons one needs to draw in order to draw the next, $(i+1)$st new, distinct coupon. Thus in order to collect all n coupons one needs to draw $\sum_{i=1}^{n-1} T_i$ coupons, where the summands are independent random variables. However, we are not particularly interested in the total number of coupons drawn, but in the number of batches needed. If a batch has size $s \leq n-1$, then this batch is worth s coupons, or vice versa, each coupon drawn in this batch is worth $1/s$ batches. Thus we need to estimate the size of the batch that contained all coupons that were needed to draw the $(i+1)$st distinct coupon, or if these coupons were contained in multiple batches, then we bound all those involved — we call these batches the batches that are *linked to* $i+1$. Let L_i be the smallest and U_i the largest size of a batch linked to the $(i+1)$st coupon. Then certainly $U_i \leq i$, as at the time that the $(i+1)$st distinct coupon gets collected there are obviously at most i distinct collected coupons. Using our assumption $\ell \cdot n \leq |I_{\lfloor \log_2 n \rfloor - a + b}| \leq u \cdot n$ we thus obtain

$$\left\lceil \sum_{i=\lfloor un \rfloor}^{n-1} \frac{T_i}{U_i} \right\rceil \leq X_n - \left(\lfloor \log_2 n \rfloor - a + b\right) \leq \left\lceil \sum_{i=\lceil \ell n \rceil}^{n-1} \frac{T_i}{L_i} \right\rceil. \tag{15}$$

Abbreviating $Y_n = X_n - (\lfloor \log_2 n \rfloor - a + b)$ and recalling that $U_i \le i$ yields

$$Y_n \ge \left\lceil \sum_{i=\lfloor un \rfloor}^{n-1} \frac{T_i}{U_i} \right\rceil = \left\lceil \sum_{i=\lfloor un \rfloor}^{n-1} \frac{T_i}{i} \right\rceil.$$

As the T_i are independent and geometrically distributed, we can apply Lemma 3.9 and for $G \sim \mathrm{Gum}(\gamma)$ we obtain with Theorem 3.10

$$\sup_{k\in\mathbb{Z}} \left| P\left[\sum_{i=\lfloor un \rfloor}^{n-1} \frac{T_i - \mathbb{E}[T_i]}{i} \ge k \right] - P\left[G \ge k\right] \right| = o(1)$$

and therefore

$$\left\lceil \sum_{i=\lfloor un \rfloor}^{n-1} \frac{T_i}{i} \right\rceil = \left\lceil \sum_{i=\lfloor un \rfloor}^{n-1} \frac{\mathbb{E}[T_i]}{i} + \sum_{i=\lfloor un \rfloor}^{n-1} \frac{T_i - \mathbb{E}[T_i]}{i} \right\rceil \gtrsim \left\lceil \sum_{i=\lfloor un \rfloor}^{n-1} \frac{1}{i(1-i/n)} + G \right\rceil.$$

The partial fraction decomposition $\big(i(1-i/n)\big)^{-1} = (n-i)^{-1} + i^{-1}$ allows us to simplify

$$\begin{aligned}\left\lceil \sum_{i=\lfloor un \rfloor}^{n-1} \frac{1}{i(1-i/n)} + G \right\rceil &= \left\lceil \sum_{i=\lfloor un \rfloor}^{n-1} \frac{1}{n-i} + \sum_{i=\lfloor un \rfloor}^{n-1} \frac{1}{i} + G \right\rceil \\ &= \left\lceil \sum_{i=1}^{n-\lfloor un \rfloor} \frac{1}{i} + \sum_{i=\lfloor un \rfloor}^{n-1} \frac{1}{i} + G \right\rceil.\end{aligned}$$

Expressing these partial harmonic sums using the asymptotic expansion for the nth harmonic number

$$\sum_{1\le k\le n} k^{-1} = H_n = \ln n + \gamma + O(1/n) \tag{16}$$

we get, using Lemma 3.2,

$$\begin{aligned} Y_n &\gtrsim \lceil \ln(n-un) + \gamma + \ln n + \gamma - \ln(un) - \gamma + G + O(1/n) \rceil \\ &= \left\lceil \ln n + \ln\left(\frac{n(1-u)}{un}\right) + \gamma + G + O(1/n) \right\rceil \\ &\gtrsim \left\lceil \ln n + \ln(1/u - 1) + \gamma + G \right\rceil. \end{aligned}$$

We now look at the upper bound in (15). For all $\lfloor \ell n \rfloor \le i \le n-1$ we specify an appropriate bound for L_i. To obtain it, assume that t is the round in which the ith vertex was informed. Then all batches that are linked to the $(i+1)$st coupon have size at least $|I_t|$, i.e. $L_i \ge |I_t|$, as the $(i+1)$st distinct coupon is drawn after the ith distinct coupon, i.e., it cannot be drawn in any round $t' < t$. However, we do not know $|I_t|$, but we certainly can say that $|I_{t+1}| \ge i$. So, Lemma 3.11, guarantees that whp

$$|I_t| \ge \left(1 - n^{-1/3}\right) \cdot e \cdot \big(i - (1-1/e)n\big) \quad \text{for all } i \in \{\lfloor \ell n \rfloor, \ldots, n-1\}.$$

(Note that $t = t(i)$ in that statement.) In particular, whp

$$L_i \ge |I_t| \ge \left(1 - n^{-1/3}\right) \cdot (n - e\cdot(n-i)) \quad \text{for all } i \in \{\lfloor \ell n \rfloor, \ldots, n-1\}.$$

Let C be the event that Lemma 3.11 conditions on, that is that $|I_t|$ (for all $t \in \mathbb{N}$) is closely concentrated around its expectation. Let $k \in \mathbb{N}$ and

$$B = \left\{ \left\lceil \sum_{i=\lceil \ell n \rceil}^{n-1} \frac{T_i}{L_i} \right\rceil \geq k \right\}.$$

Then $P(B) = P(C \cap B) + o(1)$ and as

$$\{C \cap B \geq k\} \Rightarrow \left\{ \left\lceil \sum_{i=\lceil \ell n \rceil}^{n-1} \frac{T_i}{(1-n^{-1/3})(n - e \cdot (n-i))} \right\rceil \geq k \right\}$$

we get, recalling $Y_n = X_n - (\lfloor \log_2 n \rfloor - a + b)$, that

$$Y_n \leq \left\lceil \sum_{i=\lceil \ell n \rceil}^{n-1} \frac{T_i}{L_i} \right\rceil \precsim \left\lceil \sum_{i=\lceil \ell n \rceil}^{n-1} \frac{T_i}{(1-n^{-1/3})(n - e \cdot (n-i))} \right\rceil.$$

Again applying Lemma 3.9 and Theorem 3.10, for $G \sim \mathrm{Gum}(\gamma)$ and $c = e$, we obtain

$$\begin{aligned} Y_n &\precsim \left\lceil \sum_{i=\lceil \ell n \rceil}^{n-1} \frac{\mathbb{E}[T_i]}{(1-n^{-1/3})(n - e \cdot (n-i))} + \sum_{i=\lceil \ell n \rceil}^{n-1} \frac{T_i - \mathbb{E}[T_i]}{(1-n^{-1/3})(n - e \cdot (n-i))} \right\rceil \\ &\precsim \left\lceil \left(1 + O(n^{-1/3})\right) \sum_{i=\lceil \ell n \rceil}^{n-1} \frac{1}{(n - e \cdot (n-i))(1 - i/n)} + G \right\rceil. \end{aligned}$$

Let $c = 1 - 1/e$. Using that $\left((n - e \cdot (n-i))(1-i/n)\right)^{-1} = (n-i)^{-1} + (i - cn)^{-1}$ gives

$$Y_n \precsim \left\lceil \left(1 + O(n^{-1/3})\right) \left(\sum_{i=\lceil \ell n \rceil}^{n-1} \frac{1}{n-i} + \sum_{i=\lceil \ell n \rceil}^{n-1} \frac{1}{i - cn} \right) + G \right\rceil.$$

Using index shifts, the asymptotic expansion for the harmonic number (16) and Lemma 3.2 yields

$$\begin{aligned} Y_n &\precsim \left\lceil \left(1 + O(n^{-1/3})\right) \left(\sum_{i=1}^{n - \lceil \ell n \rceil} \frac{1}{i} + \sum_{i=\lceil \ell n \rceil - \lfloor cn \rfloor}^{n-1-\lfloor cn \rfloor} \frac{1}{i} \right) + G + o(1) \right\rceil \\ &\precsim \left\lceil \ln n + \ln(1/\ell - 1) - \ln(1/\ell) + \gamma + \ln\left(\frac{1-c}{\ell - c}\right) + G \right\rceil. \end{aligned}$$ □

3.6. *Proof of Lemma 2.4*

In this subsection we investigate the double limit

$$\lim_{a \to \infty, a \in \mathbb{N}} \lim_{b \to \infty, b \in \mathbb{N}} -a + b + \ln\left(g^{(b)}(1 - 2^{-a-x})\right) - x$$

where $g(x) = xe^{x-1}$. We will show that this limit exists and defines a continuous function $c(x)$. It being periodic with period 1 is an immediate consequence of substituting $a \to a+1$ in the limit. A similar proof would also yield that c is continuously differentiable, but we only need continuity in the proof of our main theorem.

Before we actually prove Lemma 2.4 let us state two auxiliary statements first. In Definition 3.12, we quantify "exponentially fast convergence" and in Lemma 3.13 we state some simple properties.

Definition 3.12 (*Exponentially Fast Convergence*). Let $(a_n)_{n\in\mathbb{N}}$ be a real-valued sequence and let $c \in (0, 1)$. If there is an $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ we have $|a_{n+1}| < c|a_n|$, then we say that a_n *converges exponentially fast to zero at rate c with start number* n_0.

Lemma 3.13.

(a) Let $c \in (0, 1)$ and let $(a_n)_{n\in\mathbb{N}}$ be a real-valued sequence that converges exponentially fast to zero at rate c. Then $\sum_{n\geq 1} a_n$ converges absolutely.

(b) Let $c \in (0, 1)$, $n_0 \in \mathbb{N}$ and let $(h_n)_{n\in\mathbb{N}}$ denote a sequence of functions with $h_n : [0, 1] \to \mathbb{R}$ such that for any $x \in [0, 1]$ the sequence $(h_n(x))_{n\in\mathbb{N}}$ converges exponentially fast to zero at rate c with start number at most n_0. Define $h : [0, 1] \to \mathbb{R}$ by $h(x) := \sum_{n\geq 1} h_n(x)$. Then the sequence of functions $(\sum_{j=1}^{n} h_j)_{n\in\mathbb{N}}$ converges uniformly to h, i.e.

$$\lim_{n\to\infty} \sup_{x\in[0,1]} \left| h(x) - \sum_{j=1}^{n} h_j(x) \right| = 0.$$

Proof. (a) is elementary. We prove (b). Let $\varepsilon > 0$. We show that there is an $n_1 \in \mathbb{N}$ such that for all $n \geq n_1$ and for all $x \in [0, 1]$ it holds $\left|\sum_{j=1}^{n} h_j(x) - h(x)\right| < \varepsilon$. For $n \geq n_0$ it is

$$\left|\sum_{j=1}^{n} h_j(x) - h(x)\right| = \left|\sum_{j=n+1}^{\infty} h_n(x)\right| \leq \sum_{j=n+1}^{\infty} |h_n(x)| \leq |a_{n_0}| \sum_{j=n+1}^{\infty} c^j = |a_{n_0}| \frac{c^{n+1}}{1-c}$$

which implies that an n_1 as required exists. $\square$

Proof of Lemma 2.4. We show first, that for a fixed and any $x \in [0, 1]$ the limit

$$\lim_{b\to\infty, b\in\mathbb{N}} b + \ln\left(g^{(b)}\left(1 - 2^{-a-x}\right)\right)$$

exists and the convergence is uniform. Inductively we get

$$\begin{aligned} b + \ln\left(g^{(b)}\left(1 - 2^{-a-x}\right)\right) &= b + \ln\left(g^{(b-1)}\left(1 - 2^{-a-x}\right)\right) + g^{(b-1)}\left(1 - 2^{-a-x}\right) - 1 \\ &= 1 + \ln\left(1 - 2^{-a-x}\right) - 2^{-a-x} + \sum_{j=1}^{b-1} g^{(j)}\left(1 - 2^{-a-x}\right) \end{aligned} \tag{17}$$

which, according to Lemma 3.13(a), converges for $b \to \infty$ because $g^{(j)}\left(1 - 2^{-a-x}\right)$ converges exponentially fast to zero at rate at most $\exp(-2^{-a-1}) < 1$ and start number 1 for $j \to \infty$ in the sense of Definition 3.12. For $x \in [0, 1]$, according to Lemma 3.13(b), the convergence is even uniform with respect to x. By the Uniform Limit Theorem we thus showed that

$$\gamma_a(x) = -a + \sum_{j\geq 1} g^{(j)}(1-2^{-a-x}) \quad \text{is continuous for } a \in \mathbb{N}. \tag{18}$$

To complete the proof we show that the sequence of continuous functions $(\gamma_a)_{a\in\mathbb{N}}$ converges uniformly. But first we make an observation. Let $a' > a \in \mathbb{N}$ and $x \in [0,1]$, then, using $g(x) = 1 - f(1-x)$,

$$\begin{aligned}\gamma_{a'}(x) &= -a' + \sum_{j\geq 1} g^{(j)}(1-2^{-a'-x})\\ &= -a' + \sum_{j=1}^{a'-a} g^{(j)}\big(1-2^{-a'-x}\big) + \sum_{j\geq 1}^{\infty} g^{(j)}\left(g^{(a'-a)}\big(1-2^{-a'-x}\big)\right)\\ &= -a - \sum_{j=1}^{a'-a} f^{(j)}\big(2^{-a'-x}\big) + \sum_{j\geq 1} g^{(j)}\left(1 - f^{(a'-a)}\big(2^{-a'-x}\big)\right).\end{aligned}$$

Furthermore, we can bound the repeated application of f using Corollary 3.2 and therefore

$$0 \leq \sum_{j=1}^{a'-a} f^{(j)}\big(2^{-a'-x}\big) \leq 2^{-a-x+1}$$

and

$$2^{-a-x}\big(1-2^{-a-x}-2^{-2a-2x}\big) \leq f^{(a'-a)}\big(2^{-a'-x}\big) \leq 2^{-a-x}.$$

Thus there is $x' \in [0,1]$ such that $|x-x'| \leq 2^{-a}$ and $\gamma_{a'}(x) = \gamma_a(x') + O\big(2^{-a}\big)$.

With this at hand we show uniform convergence of $(\gamma_a)_{a\in\mathbb{N}}$. In particular, for any $0 < \varepsilon < 1/8$ we will show that there is some $N \in \mathbb{N}$ such that $\sup_{x\in[0,1]} |\gamma_a(x) - \gamma_{a'}(x)| \leq \varepsilon$ for all $a' > a > N$. To achieve this we use our previous observation and obtain that

$$\begin{aligned}\sup_{x\in[0,1]} |\gamma_a(x)-\gamma_{a'}(x)| &\leq \sup_{x\in[0,1],|x-x'|\leq 2^{-a}} |\gamma_a(x)-\gamma_a(x')| + O\big(2^{-a}\big)\\ = \sup_{x\in[0,1],|x-x'|\leq 2^{-a}} &\left|\sum_{j\geq 1}\Big(g^{(j)}\big(1-2^{-a-x'}\big) - g^{(j)}\big(1-2^{-a-x}\big)\Big)\right| + O\big(2^{-a}\big).\end{aligned}$$

We bound this sum by splitting it into three parts. There is $M_1 \in \mathbb{N}$ such that for any $a > M_1$ there is $N_1 \in \mathbb{N}$ (N_1 depending on a and ε) such that

$$\varepsilon \leq f^{(N_1)}\big(2^{-a-1}\big) \leq f^{(N_1+1)}\big(2^{-a}\big) \leq 8\varepsilon. \tag{19}$$

That is, N_1 is the number of iterations such that $f^{(N_1)}\big(2^{-a}\big) \approx \varepsilon$, in particular $N_1 \leq a$, as $f^{(a)}\big(2^{-a-1}\big) \geq 1/8$ by Corollary 3.2 and the fact that f is increasing. Furthermore, using again that $g^{(j)}\left(1-2^{-a-x}\right)$ converges exponentially fast to zero with rate at most $\exp(-2^{-a-1}) < 1$ for $j \to \infty$, there is $c \in \mathbb{N}$ depending only on ε such that for $N_2 := N_1 + c$

$$0 \leq \sup_{x\in[0,1]} \sum_{j\geq N_2} g^{(j)}\big(1-2^{-a-x}\big) \leq \varepsilon \quad \text{for all } a > M_1. \tag{20}$$

Then, abbreviating $h^{(j)} = g^{(j)}\big(1 - 2^{-a-x'}\big) - g^{(j)}\big(1 - 2^{-a-x}\big)$, we can write

$$\sum_{j=1}^{\infty}\Big(g^{(j)}\big(1 - 2^{-a-x'}\big) - g^{(j)}\big(1 - 2^{-a-x}\big)\Big) = \sum_{j=1}^{N_1} h^{(j)} + \sum_{j=N_1+1}^{N_2} h^{(j)} + \sum_{j>N_2} h^{(j)}. \tag{21}$$

In the rest of the proof estimate these sums individually, starting with the first one. Again using (17) and $f(x) = 1 - g(1-x)$ we have as $a \to \infty$

$$\begin{aligned}
&\sum_{j=1}^{N_1} g^{(j)}\big(1 - 2^{-a-x'}\big) - g^{(j)}\big(1 - 2^{-a-x}\big)\\
&= \ln g^{(N_1)}\big(1 - 2^{-a-x'}\big) - \ln g^{(N_1)}\big(1 - 2^{-a-x}\big) + O(2^{-a})\\
&= \ln\Big(1 - f^{(N_1)}\big(2^{-a-x'}\big)\Big) - \ln\big(1 - f^{(N_1)}\big(2^{-a-x}\big)\big) + O(2^{-a}).
\end{aligned}$$

By our choice of N_1, see (19), and the elementary inequalities $z/(1+z) \le \ln(1+z) \le z$ for all $z > -1$ this yields the upper bound

$$\sup_{x\in[0,1]}\left|\sum_{j=1}^{N_1} h^{(j)}\right| \le \varepsilon + \frac{8\varepsilon}{1+8\varepsilon} + O(2^{-a}) \quad \text{for all } a > M_1. \tag{22}$$

We continue with the second sum in (21). Corollary 3.1 yields

$$\begin{aligned}
\left|\sum_{j=N_1+1}^{N_2} h^{(j)}\right| &= \left|\sum_{j=N_1+1}^{N_2} \big(f^{(j)}\big(2^{-a-x}\big) - f^{(j)}\big(2^{-a-x'}\big)\big)\right|\\
&\le \Big(2^{-a-x} - 2^{-a-x'}\Big)\sum_{j=N_1+1}^{N_2} 2^j.
\end{aligned}$$

Thus, as $N_1 \le a$ and $N_2 = N_1 + c$, where c depends on ε only, and our assumption $|x - x'| \le 2^{-a}$ there is $M_2 \ge M_1$ such that for all $a > M_2$

$$\left|\sum_{j=N_1+1}^{N_2} h^{(j)}\right| \le \Big(2^{2^{-a}} - 1\Big)\cdot 2^{-a}\cdot \sum_{j=N_1+1}^{N_2} 2^j \le \Big(2^{2^{-a}} - 1\Big)\cdot 2^{c+1} \le \varepsilon. \tag{23}$$

In summary, (21) gives

$$\begin{aligned}
&\sup_{x\in[0,1]}\big|\gamma_a(x) - \gamma_{a'}(x)\big|\\
&\le \sup_{x\in[0,1],|x-x'|\le 2^{-a}}\left|\sum_{j=1}^{N_1} h^{(j)} + \sum_{j=N_1+1}^{N_2} h^{(j)} + \sum_{j>N_2} h^{(j)}\right| + O\big(2^{-a}\big).
\end{aligned}$$

and for $a > M_2 > M_1$, applying (22), (23) and (20) yields the uniform convergence of $(\gamma_a)_{a\in\mathbb{N}}$. □

3.7. Other proofs

In this subsection we complete the rigorous treatment of our main theorems and give the last two remaining proofs. First we prove Lemma 2.5, which states that $X_n - \lfloor \log_2 n\rfloor - \lfloor \ln n\rfloor$ is uniformly integrable.

Proof of Lemma 2.5. Doerr and Künnemann show in [7, Cor. 3.2 and Thm. 4.1] that for all $r \in \mathbb{N}$

$$P\Big[X_n \geq \lfloor \log_2 n \rfloor + \ln n + 2.188 + r\Big] \leq 2e^{-r} \quad \text{and}$$

$$P\Big[X_n \leq r\Big] \leq P\left[\lfloor \log_2 n \rfloor - 1 + \frac{C_n(\left\lceil n/2 \right\rceil)}{n} \leq r\right],$$

where $C_n(\lceil n/2 \rceil)$ is the number of rounds a coupon collector needs to draw the last $n/2$ out of n coupons. These two bounds together with common deviation bounds for the coupon collector problem imply, see e.g. [8], that

$$P[Y_n \notin \lfloor \log_2 n \rfloor + \lfloor \ln n \rfloor \pm (r+5)] \leq 4e^{-r}.$$

Using this inequality we get that for any $N \in \mathbb{N}$

$$\mathbb{E}\Big[|Y_n|^k \;\Big|\; \mathbb{1}\big[|Y_n|^k > N\big]\Big] \leq \sum_{t \geq \sqrt[k]{N}} (t+5)^k 4e^{-t},$$

which implies the claim. □

We close the section with the proof of Lemmas 3.1 and 3.2.

Proof of Lemma 3.1. First we observe that the $(1-n^{-1/6})$ error term in the definition of ℓ, u is negligible as is factors out as a small additional term. Thus it suffices to consider $\ell = f^{(b)}(L)$ and $u = f^{(b)}(U)$ where $L = \left(1 - 2^{-a+2}\right) 2^{-a-x}$ and $U = 2^{-a-x}$ for some $x \in (0, 1]$. We assume that $a \geq 3$.

We start by showing an analogue to Corollary 3.1 but concerning g. For all $r \geq s \in [0, 1]$, using $1 - x \leq e^{-x}$,

$$g(r-s) = (r-s)e^{r-s-1} \geq re^{r-s-1} - se^{r-1} \geq g(r) - s(1+r)e^{r-1}$$

and consequently

$$g^{(i)}(r-s) \geq g^{(i)}(r) - s\big((1+r)e^{r-1}\big)^i \quad \text{for all } r \geq s \in [0,1) \text{ and } i \in \mathbb{N}. \tag{24}$$

This completes our preparations. In order to show that $(1-\ell)/(1-u) \to 1$ as $a \to \infty$ we argue that ℓ and u are very close together and approach 1 as a (and $b > 2a$) gets big. We start by bounding the distance between ℓ and u. Applying Corollary 3.1 to $U = L + 2^{-2a-x+2}$ we get that

$$f^{(a)}(U) = f^{(a)}\left(L + 2^{-2a-x+2}\right) \leq f^{(a)}(L) + 2^{-a-x+2} \tag{25}$$

and Corollary 3.2 bounds $f^{(a-1)}(U)$ from below with $2^{-x-1}\left(1 - 2^{-x-1} - 2^{-2x-2}\right) \geq 1/8$, thus $f^{(a+2)}(U) \geq 1/2$, and therefore we get using the monotonicity of f

$$\frac{1}{2} \leq f^{(a+2)}(U) \leq f^{(a+3)}(L) \leq f^{(a+3)}(U) \overset{(25)}{\leq} f^{(a+3)}(L) + 2^{-a-x+5}. \tag{26}$$

We switch our focus to g. Observe that $z := 3e^{-1/2}/2 < 1$ and, using (24),

$$\begin{aligned} & g^{(b-a-3)}\left(1 - f^{(a+3)}(L) - 2^{-a-x+5}\right) \\ \geq\; & g^{(b-a-3)}\left(1 - f^{(a+3)}(L)\right) - 2^{-a-x+5} \cdot z^{b-a-3}. \end{aligned}$$

This implies, using (26) and the previous equation, that

$$g^{(b-a-3)}\left(1-f^{(a+3)}(U)\right) \geq g^{(b-a-3)}\left(1-f^{(a+3)}(L)\right) - 2^{-a-x+5}\cdot z^{b-a-3},$$

and therefore, as $1-f^{(b)}(L) \geq 1-f^{(b)}(U) = g^{(b-a-3)}\left(1-f^{(a+3)}(U)\right)$,

$$|u-\ell| = |f^{(b)}(U) - f^{(b)}(L)| \leq 2^{-a-x+5}z^{b-a-3} \to 0 \quad \text{as } a\to\infty, b-a\to\infty. \tag{27}$$

Next we show that u, ℓ approach 1. Using $g(x) = 1-f(1-x)$, (26), g being increasing and $g(x) \leq xe^{-1/2}$ for all $x \leq 1/2$ (in that order), we get for all $b > a+3$

$$g^{(b)}(1-L) = g^{(b-a-3)}\left(1-f^{(a+3)}(L)\right) \leq g^{(b-a-3)}\left(\frac{1}{2}\right) \leq \frac{1}{2}e^{-(b-a-3)/2}.$$

Moreover, using that $f(x) \leq 2x$ and $g(x) \geq x/e$,

$$g^{(b)}(1-U) = g^{(b-a)}\left(1-f^{(a)}(U)\right) \geq \left(1-2^{-x}\right)e^{-(b-a)}.$$

Thus, these two bounds together give for all $b > a+3$

$$1-\frac{1}{2}e^{-(b-a-3)/2} \leq f^{(b)}(L) \leq f^{(b)}(U) \leq 1-\left(1-2^{-x}\right)e^{-(b-a)}. \tag{28}$$

We just showed that $u, \ell \to 1$ as a (and b) tends to infinity. This yields that $\ln u, \ln \ell$ and $\ln\left(\ell/(e\ell - e + 1)\right)$ tend to 0, leaving us with the term $\ln\left((1-\ell)/(1-u)\right)$. The fact $U \leq f(L)$ (and so $f^{(b-2)}(U) \leq f^{(b-1)}(L)$) implies that

$$\begin{aligned}
\frac{1-\ell}{1-u} &= \frac{g^{(b)}\left(1-L\right)}{g^{(b)}\left(1-U\right)} \\
&= \frac{\exp\left(g^{(b-1)}(1-L)-1\right)\cdot\exp\left(g^{(b-2)}(1-L)-1\right)}{\exp\left(g^{(b-1)}(1-U)-1\right)\cdot\exp\left(g^{(b-2)}(1-U)-1\right)}\cdot\frac{g^{(b-2)}\left(1-L\right)}{g^{(b-2)}\left(1-U\right)} \\
&\leq \frac{\exp\left(g^{(b-2)}(1-L)-1\right)}{\exp\left(g^{(b-1)}(1-U)-1\right)}\cdot\frac{g^{(b-2)}\left(1-L\right)}{g^{(b-2)}\left(1-U\right)}.
\end{aligned}$$

Applying the same estimate to the latter fraction inductively we get for any $c \in \mathbb{N}$

$$\frac{1-\ell}{1-u} \leq \frac{\exp\left(g^{(c)}(1-L)-1\right)}{\exp\left(g^{(b-1)}(1-U)-1\right)}\cdot\frac{g^{(c)}\left(1-L\right)}{g^{(c)}\left(1-U\right)} \leq \exp\left(g^{(c)}(1-L)\right)\cdot\frac{g^{(c)}\left(1-L\right)}{g^{(c)}\left(1-U\right)}.$$

Set $c = \lceil a(1+\ln 2)\rceil$. Using (27) and (28), where we set $b = c$, we obtain (for large enough a) that $\left|g^{(c)}(1-L) - g^{(c)}(1-U)\right| \leq 2^{-a-x+5}z^{c-a-3}$ as well as $f^{(c)}\left(U\right) \leq 1-\left(1-2^{-x}\right)e^{-(c-a)}$ and $f^{(c)}(L) \geq 1-e^{-(c-a-3)/2}/2$. Thus

$$\begin{aligned}
\frac{1-\ell}{1-u} &\leq \exp\left(1-f^{(c)}(L)\right)\left(1+\frac{2^{-a-x+5}z^{c-a-3}}{1-f^{(c)}\left(U\right)}\right) \\
&\leq \exp\left(\frac{1}{2}e^{-(c-a-3)/2}\right)\left(1+\frac{2^{-a-x+5}z^{c-a-3}}{(1-2^{-x})e^{-c+a}}\right).
\end{aligned}$$

Using $e^x \leq 1+2x, x \in [0,1]$ this yields the bounds

$$1 \leq \frac{1-\ell}{1-u} \leq \left(1+\sqrt{2}^{-a+4}\right)\left(1+\frac{2^{-x+5}e}{1-2^{-x}}\cdot z^{a\ln 2-2}\right) \quad \text{for all } a \in \mathbb{N}.$$

Therefore, as $0 < z < 1$ we obtain $\frac{1-\ell}{1-u} \to 1$, and consequently $\ln\left((1-\ell)/(1-u)\right) \to 0$, as $a \to \infty$. □

Lemma 3.2 states that disturbing a Gumbel distributed random variable by a small amount does not significantly alter its distribution.

Proof of Lemma 3.2. Observe that $\lceil h(n) + G \pm \varepsilon \rceil \neq \lceil h(n) + G \rceil$ is equivalent to

$$G \in \left[j - h(n) - \varepsilon, \ \ j - h(n) + \varepsilon \right] \quad \text{for some } j \in \mathbb{Z}.$$

But as G is absolutely continuous, for any $\delta > 0$ we can choose ε small enough such that

$$P\left[G \in \bigcup_{j \in \mathbb{Z}} [j - h(n) - \varepsilon, j - h(n) + \varepsilon] \right] \leq \delta. \quad \square$$

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix. Existence of subsequence

Let $x, y \in [0, 1]$. In this section we show that there is an unbounded sequence of natural numbers $(n_i)_{i \in \mathbb{N}}$ such that $\log_2 n_i - \lfloor \log_2 n_i \rfloor \to x$ and $\ln n_i - \lfloor \ln n_i \rfloor \to y$ as $i \to \infty$. To this end, set $z = y - x \ln 2$. According to a Theorem of Kronecker, see e.g. [17, Thm. 440], for all $i \in \mathbb{N}$, there are $p_i, q_i \in \mathbb{N}$ such that

$$\left| q_i \ln 2 - p_i - z \right| \leq i^{-1}. \tag{A.1}$$

Actually even more is true: there are infinitely many $p_i, q_i \in \mathbb{N}$ that solve (A.1). To see this, assume that there are only finitely many, then there is $k, \ell \in \mathbb{N}$ such that $k \ln 2 = \ell + z$, otherwise there would be some $i \in \mathbb{N}$ where (A.1) has no solution. However, according to a Theorem of Hurwitz, see e.g. [17, Thm. 193], there are infinitely many $r_j, s_j \in \mathbb{N}$ such that

$$\left| r_j \ln 2 - s_j \right| \leq r_j^{-2}.$$

But then

$$\left| r_j \ln 2 - s_j \right| = \left| (r_j + k) \ln 2 - (s_j + \ell) - z \right| \leq r_j^{-2},$$

a contradiction, thus there are infinitely many solutions to (A.1). We continue with that equation, which we can restate, as $i \to \infty$,

$$q_i \ln 2 + x \ln 2 = p_i + y + O\left(i^{-1}\right).$$

Taking the exponential on both sides thus yields, as $i \to \infty$,

$$2^{q_i + x} = e^{p_i + y + O(i^{-1})}.$$

Set $n_i = \lfloor 2^{q_i + x} \rfloor$ for all $i \in \mathbb{N}$, where we choose q_i such that $q_i \geq i$ from the infinitely many solutions to (A.1). Then $n_i \in \mathbb{N}$ for all $i \in \mathbb{N}$ and

$$\log_2 n_i - \lfloor \log_2 n_i \rfloor = x + O\left(2^{-i}\right) \quad \text{as well as} \quad \ln n_i - \lfloor \ln n_i \rfloor = y + O\left(i^{-1}\right).$$

Thus the subsequence of natural numbers that is induced by $\log_2 n_i - \lfloor \log_2 n_i \rfloor \to x$ and $\ln n_i - \lfloor \ln n_i \rfloor \to y$ is non-empty and unbounded.

References

[1] K.P. Birman, M. Hayden, O. Ozkasap, Z. Xiao, M. Budiu, Y. Minsky, Bimodal multicast, ACM Trans. Comput. Syst. 17 (2) (1999) 41–88.
[2] S. Boucheron, G. Lugosi, P. Massart, Concentration Inequalities: A Nonasymptotic Theory of Independence, Oxford University Press, 2013.
[3] F. Chierichetti, S. Lattanzi, A. Panconesi, Almost tight bounds for rumour spreading with conductance, in: Proceedings of the Forty-Second ACM Symposium on Theory of Computing, ACM, 2010, pp. 399–408.
[4] F. Chierichetti, S. Lattanzi, A. Panconesi, Rumour spreading and graph conductance, in: Proceedings of the Twenty-First Annual ACM–SIAM Symposium on Discrete Algorithms, SIAM, 2010, pp. 1657–1663.
[5] R. Daknama, K. Panagiotou, S. Reisser, Robustness of randomized rumour spreading, in: 27th Annual European Symposium on Algorithms, ESA 2019, in: Leibniz International Proceedings in Informatics (LIPIcs), vol. 144, Dagstuhl, Germany, 2019, pp. 36:1–36:15, Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.
[6] A. Demers, D. Greene, C. Hauser, W. Irish, J. Larson, S. Shenker, H. Sturgis, D. Swinehart, D. Terry, Epidemic algorithms for replicated database maintenance, in: Proceedings of the Sixth Annual ACM Symposium on Principles of Distributed Computing, ACM, 1987, pp. 1–12.
[7] B. Doerr, M. Künnemann, Tight analysis of randomized rumor spreading in complete graphs, in: 2014 Proceedings of the Eleventh Workshop on Analytic Algorithmics and Combinatorics, ANALCO, SIAM, 2014, pp. 82–91.
[8] P. Erdős, A. Rényi, On a classical problem of probability theory, in: Magyar Tud. Akad. Mat. Kutató Int. Közl. 6, 1961, pp. 215–220.
[9] U. Feige, D. Peleg, P. Raghavan, E. Upfal, Randomized broadcast in networks, Random Struct. Algorithms 1 (4) (1990) 447–460.
[10] N. Fountoulakis, A. Huber, K. Panagiotou, Reliable broadcasting in random networks and the effect of density, in: INFOCOM 2010. 29th IEEE International Conference on Computer Communications, Joint Conference of the IEEE Computer and Communications Societies, 15–19 March, 2010, San Diego, CA, USA, 2010, pp. 2552–2560.
[11] N. Fountoulakis, K. Panagiotou, Rumor spreading on random regular graphs and expanders, in: Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, 2010, pp. 560–573.
[12] A.M. Frieze, G.R. Grimmett, The shortest-path problem for graphs with random arc-lengths, Discrete Appl. Math. 10 (1) (1985) 57–77.
[13] G. Giakkoupis, Tight bounds for rumor spreading in graphs of a given conductance, in: 28th International Symposium on Theoretical Aspects of Computer Science, STACS 2011, March 10–12, 2011, Dortmund, Germany, 2011, pp. 57–68.
[14] G. Giakkoupis, Tight bounds for rumor spreading with vertex expansion, in: Proceedings of the Twenty-Fifth Annual ACM–SIAM Symposium on Discrete Algorithms, SODA 2014, Portland, Oregon, USA, January 5–7, 2014, 2014, pp. 801–815.
[15] G. Giakkoupis, Y. Nazari, P. Woelfel, How asynchrony affects rumor spreading time, in: Proceedings of the 2016 ACM Symposium on Principles of Distributed Computing, pp. 185–194.
[16] G. Giakkoupis, T. Sauerwald, Rumor spreading and vertex expansion, in: Proceedings of the Twenty-Third Annual ACM–SIAM Symposium on Discrete Algorithms, SODA 2012, Kyoto, Japan, January 17–19, 2012, 2012, pp. 1623–1641.
[17] G.H. Hardy, E.M. Wright, et al., An Introduction to the Theory of Numbers, Oxford University Press, 1979.
[18] S. Janson, One, two and three times log n/n for paths in a complete graph with random weights, 8 (4) 347–361.
[19] D. Mosk-Aoyama, D. Shah, Fast distributed algorithms for computing separable functions, IEEE Trans. Inform. Theory 54 (7) (2008) 2997–3007.
[20] K. Panagiotou, X. Pérez-Giménez, T. Sauerwald, H. Sun, Randomized rumour spreading: The effect of the network topology, Comb. Probab. Comput. 24 (2) (2015) 457–479.
[21] K. Panagiotou, L. Speidel, Asynchronous rumor spreading on random graphs, 78 (3) 968–989.
[22] C. Patsonakis, M. Roussopoulos, Revisiting asynchronous rumor spreading in the blockchain era, in: 2019 IEEE 25th International Conference on Parallel and Distributed Systems, ICPADS, 2019, pp. 284–293.
[23] B. Pittel, On spreading a rumor, SIAM J. Appl. Math. 47 (1) (1987) 213–223.
[24] G. Pólya, Über den zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung und das Momentenproblem, Math. Z. 8 (3) (1920) 171–181.
[25] A. Pourmiri, B. Mans, Tight analysis of asynchronous rumor spreading in dynamic networks, in: Proceedings of the 39th Symposium on Principles of Distributed Computing, pp. 263–272.

[26] T. Sauerwald, A. Stauffer, Rumor spreading and vertex expansion on regular graphs, in: Proceedings of the Twenty-Second Annual ACM–SIAM Symposium on Discrete Algorithms, SODA 2011, San Francisco, California, USA, January 23–25, 2011, 2011, pp. 462–475.
[27] R. Serfling, Approximation Theorems of Mathematical Statistics, in: Wiley Series in Probability and Statistics, Wiley, 2009.

8 Solving a conjecture by Alon et al.

This chapter is a reprint of

> Panagiotou, K., & Reisser, S. (2021). The Effect of Iterativity on Adversarial Opinion Forming.

A preprint of this version is online at `https://arxiv.org/abs/2111.15445`.

My own contribution. This paper is joint work with my supervisor Konstantinos Panagiotou. I contributed substantially to all results. In particular, I worked out the main idea and its proof. The final version was edited under guidance of my supervisor Konstantinos Panagiotou.

The Effect of Iterativity on Adversarial Opinion Forming

Konstantinos Panagiotou   Simon Reisser

Tuesday 30$^{\text{th}}$ November, 2021

Abstract

Consider the following model to study adversarial effects on opinion forming. A set of initially selected experts form their binary opinion while being influenced by an adversary, who may convince some of them of the falsehood. All other participants in the network then take the opinion of the majority of their neighbouring experts. Can the adversary influence the experts in such a way that the majority of the network believes the falsehood? Alon et al. [2] conjectured that in this context an iterative dissemination process will always be beneficial to the adversary. This work provides a counterexample to that conjecture.

1 Introduction

Understanding how opinions are formed is as important as ever, as the spread of misinformation becomes more prevalent every day. Assume there is some new innovation being either good or bad that is introduced to a group of people who want to form their (binary) opinion about it. Following a key insight by Rogers [22], the opining forming process can be modelled as follows. At first, a small set of so-called early adopters, or experts, forms their opinion about the newly introduced innovation. Afterwards, they disseminate their opinion to all other non-experts in the network.

When looking at that network from the outside an observer wants to infer the quality of the new innovation by observing the opinion of all individuals, but without taking the actual structure of the network into consideration (maybe by doing a poll). One popular method to achieve this is using the *wisdom of the crowd*. In this case that corresponds to a simple majority rule, that is, the observer takes the majority of opinions as an estimate. Wisdom of the crowd has been shown to have a plethora of useful applications in decision making, see e.g. [9, 19, 21, 5, 8, 18].

Assume furthermore that there is an adversary who can influence the opinion of some early adopters so as to falsely convince the observer of the new innovation's quality. Let us look at some examples. Consider the so-called Black-Hat ASIN Piggybacking on Amazons Marketplace [17]. This is the method of hijacking the listing of an Amazon vendor to sell counterfeit products under the (dis-)guise of a genuine listing. Some customers then buy the real product and some buy the fake one. This results in the vendor to lose profit as well as him getting negative reviews that do not correspond to the actual product. The second example is a newly opened restaurant, that in its opening phase invites food critics to try and rate the restaurant. However, when those critics dine at the restaurant, the restaurant puts in more effort than it would when catering to a regular customer, e.g., by providing better quality food and service. Lastly, consider the common practice of online vendors to buy positive reviews for their products by either giving directly monetary incentives to reviewers or providing them with free products. In particular, on Amazon in certain product categories, like Bluetooth speakers and headphones, ReviewMeta [20] finds more than half of reviews to be fake [11].

A Model for Opinion Forming In the previous examples we saw three different sorts of adversaries: the hijacking seller influenced negatively the opinions of some customers; the restaurant owner could actively choose which critics to influence; finally, the seller that bought his reviews could select the reviewers as well as guarantee their opinion. Alon et al. [2] introduced a model that implements the ideas outlined above. Given a graph $G = (V, E)$ on n vertices and parameters $0 \leq \mu < 1/2, 0 < \delta \leq 1/2$ we define the set of experts as a set $\mathcal{E} \subseteq V$ with the property that $|\mathcal{E}| = \mu|V| = \mu n$. Let $\mathcal{E}$ be furthermore divided into two subsets: the experts that know the truth $\mathcal{E}_1 \subseteq \mathcal{E}$ and the experts that are convinced of the falsehood $\mathcal{E}_0 = \mathcal{E} \setminus \mathcal{E}_1$. The sets $\mathcal{E}_1, \mathcal{E}_0$ are chosen in three different ways that correspond to the various adversaries described in the previous paragraphs.

The *random* adversary has actually no choice. He chooses the expert set $\mathcal{E}$ uniformly at random among all sets of size $\mu|V|$. Then $\mathcal{E}$ is in turn partitioned into $\mathcal{E}_1$ and $\mathcal{E}_0$ by adding each vertex in $\mathcal{E}$ to $\mathcal{E}_1$ independently with probability $1/2 + \delta$ and to $\mathcal{E}_0$ otherwise. The *weak* adversary is allowed to choose the expert set with the restriction that $|\mathcal{E}| = \mu|V|$; the selected set is then partitioned into $\mathcal{E}_1$ and $\mathcal{E}_0$ like in the random adversary. Finally, the *strong* adversary chooses $\mathcal{E}, \mathcal{E}_1$ and $\mathcal{E}_0 = \mathcal{E} \setminus \mathcal{E}_1$ arbitrarily such that $|\mathcal{E}| = \mu|V|, |\mathcal{E}_1| = (1/2 + \delta)|\mathcal{E}|$ and consequently $|\mathcal{E}_0| = (1/2 - \delta)|\mathcal{E}|$. We will ignore rounding issues througout to facilitate the presentation.

All vertices that know the truth in a graph are assigned the label '1', including all vertices in $\mathcal{E}_1$, and all vertices that believe a falsehood are labeled '0'. Vertices without an opinion bear no label. The experts disseminate their opinions to the non-experts $V \setminus \mathcal{E}$ by a majority rule, that is, every vertex in $V \setminus \mathcal{E}$ takes the opinion of the majority of its neighbouring experts. To be completely explicit, a non-expert is labeled '1'/'0' if more that half of its neighbouring experts are labeled '1'/'0'. Vertices at which there is no majority – because of a tie of '1's and '0's or because they have no expert neighbours – decide upon their opinion uniformly at random, i.e., each of these vertices is independently labeled '1' with probability 1/2 and '0' otherwise.

We say that a graph is *robust* against the random/ weak/ strong adversary if with high probability, for any choice of the expert set, after the dissemination process more than half of the vertices are labeled '1'. 'With high probability' means with probability approaching 1 as n approaches infinity, which we sometimes abbreviate with whp. In [2] the authors studied which properties of a graph make it robust. They discovered that all graphs with maximal degree being sub-linear in n are robust against the weak adversary. Furthermore, they showed that certain well-connected networks are robust against the strong adversary. In particular, such networks are either Erdős-Rényi random graphs having edge probability p greater than c/n for a suitable constant $c > 0$, or expander graphs, with d, λ_2 being the largest and second largest eigenvalue of its adjacency matrix, satisfying $d \geq \lambda_2/(\delta\sqrt{\mu(1 - \mu + 2\delta\mu)})$.

Iterative Dissemination In [2] the authors also introduced an iterative version of the model with a more dynamic dissemination process. The *iterative model* also starts with labeled experts and all non-experts are labeled according to the majority of their neighbouring experts. Ties that involve at least one expert are broken uniformly at random. All non-experts without any expert neighbours, however, do not form their opinion right away, but remain unlabeled. This process is then iterated by considering all vertices with label '1' and all vertices with label '0', until all vertices are labeled.

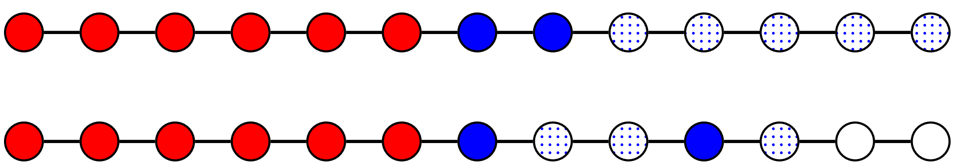

Figure 1: This figure shows an example from [2]. The colors red/blue correspond to the experts labeled '1'/'0'. The dotted vertices indicate their label after the dissemination process, the unmarked vertices are decided randomly. In the first line graph we consider the iterative strong adversary, where only the rightmost blue vertex determines the label of all remaining vertices. In the second line we consider the non-iterative setting, each blue expert can at most convince two non-experts. If $1/2 + \delta > 3(1/2 - \delta)$ and n is large, the adversary can not hope to convince more that half of all vertices.

A natural question is whether iterativity helps or hinders the adversary. Intuitively, iterativity ought to be beneficial for the adversary. If a graph is not robust against a non-iterative adversary, then there is a choice of expert sets such that after one round of dissemination there are more vertices that are labeled '0' than '1'. The remaining vertices without expert neighbours are then either decided randomly (non-iterative) or there are subsequent rounds of dissemination (iterative). As there now are more '0' labeled vertices that '1' labeled vertices, deciding the label of the remaining vertices by dissemination should be beneficial for the adversary. Indeed, the authors of [2] provided examples where this is the case. For example, they showed that for suitable values of μ and δ, a line graph is robust against the non-iterative strong/weak adversaries, but not against their iterative versions, see Fig. 1.

However, in [2] an additional example, where for the weak adversary the opposite is true, was constructed. Consider a graph that is a disjoint union of a star and a d-regular expander graph. Place one expert in the center of the star and distribute the other experts as evenly as possible on the expander. In the non-iterative setting, each expert in the expander will spread its label to d many non-experts. If the expert in the center of the star is labeled '0', all vertices in the star are labeled '0' as well, outweighing the difference between '1's and '0's in the expander. In the iterative setting however, each expert does not only spread its label to d many other vertices, but all vertices in the expander will be labeled at the end of the dissemination, roughly in the same ratio as that of the experts in the beginning. Now the difference in '1's and '0's is so large that even if all vertices in the star were labeled with '0' can sway the majority.

Guided by the intuition described previously, it seems that no such construction can work for a strong adversary. In the previous example of the graph consisting of an expander and a star the adversary can place all '1'-labeled experts on the star and all others in the expander. Then, all vertices in the expander will be labeled '0' resulting in a clear majority. Consequently, in [2] the following conjecture concerning the effect of iterativity in that case was made.

Conjecture 1.1 ([2]). *In the case of a strong adversary an iterative propagation can never harm the adversary.*

Equivalently, the conjecture states that there is no graph that is robust against the iterative strong adversary and simulaneously not robust against the non-iterative version – in this precise sense iterativity does not harm/can only help the adversary.

Related Results Besides of [2], where this model for opinion formation was introduced, there is one more work that studies questions in this precise framework. In his doctoral thesis [10], Daknama studied resilience properties of random graphs. 'Local resilience' in this context refers to the largest number of edges, which are adjacent to any vertex, that can be removed so that the graph still is robust against the strong adversary. In [10] it was shown that one can delete up to a fraction of $2(1 - \mu + 2\delta\mu)\delta/(1 + 2\delta)$ of all edges at each vertex without affecting robustness.

There are also other directly related studies in opinion forming, which, however, do not use the exact model presented here. These papers include studies on word of mouth [24], group recommendation [3, 15, 16, 12] and informational cascades [6, 7, 23, 1, 13]. For further references see also [2].

Result The contribution of this paper is to refute Conjecture 1.1. The idea is to consider a graph that has non-robustness against the non-iterative strong adversary in a very weak way. More concretely, the majority for '0' labeled vertices is only achieved if a majority of vertices without expert neighbours is labeled '0'. If we consider iterativity, then the adversary has no clear advantage in a subsequent round of the dissemination, as there are roughly equally many '1'- and '0'-labeled vertices. Additionally, we can construct the graph in a way such that the vertices without expert neighbours are connected to vertices that are labeled '1' in the first round of the dissemination, so that the adversary *gets harmed*.

Consider the following graph that implements these ideas. Let $0 < \mu, \delta < 1/2$ and $0 < \varepsilon_1 < 2\delta/(1/2 + \delta)$, $0 < \varepsilon_2 < (1/2 - \delta)/(1/2 + \delta)$ as well as $0 < d < (1 - \mu - 2\delta\mu)/3$. Then the graph

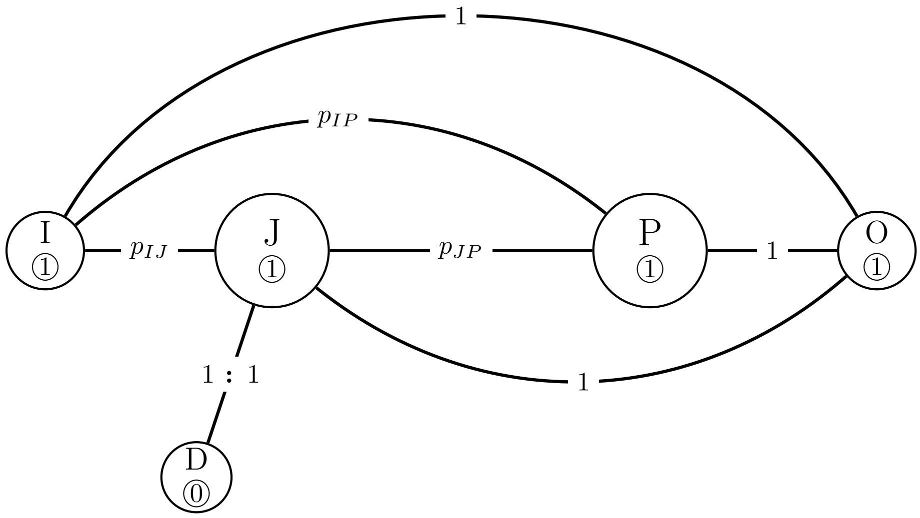

Figure 2: This figure shows the graph G. The numbers on the edges and in the vertices give the probability that an edge is present between/in the components. For example, any edge with one vertex in I and one in J exists independently with probability p_{IJ}. Every vertex in D has exactly one distinct neighbour in J and no other neighbours, i.e., every vertex in D has degree 1 and no two vertices in D have a common neighbour.

$G = (V, E),\ |V| = n$ is given by $V = I \mathbin{\dot\cup} J \mathbin{\dot\cup} O \mathbin{\dot\cup} P \mathbin{\dot\cup} D$ such that

$$|I \cup J| = |O \cup P| = (1-d)\frac{n}{2}, \quad |D| = dn, \quad |I| = \mu\left(\frac{1}{2}+\delta\right)n \quad \text{and} \quad |O| = \mu\left(\frac{1}{2}-\delta\right)n.$$

The subset D forms an independent set. In contrast, I, J, O and P each form a clique. Every vertex in O is connected to all other vertices except to those in D. Between I and J, I and P and J and P are random bipartite graphs with edge-probabilities p_{IJ}, p_{IP} and p_{JP} respectively. Every vertex in D has degree one, with the unique neighbor being in J; moreover, no two vertices in D have the same neighbour. There are no more edges. Set

$$p_{IJ} = p_{JP} = \frac{1/2-\delta}{1/2+\delta} + \varepsilon_1 \qquad \text{and} \qquad p_{IP} = \frac{1/2-\delta}{1/2+\delta} - \varepsilon_2.$$

See Fig. 2 for a depiction of G.

Assume for now that the adversary chooses $\mathcal{E}_1 = I$ and $\mathcal{E}_0 = O$. Then whp all vertices in P will have $\approx \varepsilon_2|I|$ more neighbours in O than in I by choice of p_{IP}, thus they will be labeled '0' independently of iterativity. In contrast, vertices in J have $\approx \varepsilon_1|I|$ more neighbours in I than in O and will consequently be labeled '1'. Summarizing, we have that all vertices in $I \cup J$ are labeled '1' and all vertices in $O \cup P$ are labeled '0'. As both unions have by construction the same size, the labels of vertices in D decide whether the adversary succeeds or not. This is where (non-)iterativity comes into play. In the non-iterative setting, vertices in D will choose uniformly at random, as they have no neighbours in $I \cup O$. So, with positive probability there will be more vertices labeled '0' than '1' in D, consequently granting a majority of '0'-labeled vertices. In the iterative setting however, all vertices in D will be labeled '1', as they are exclusively connected to vertices in J. Thus, when choosing $\mathcal{E}_1 = I$ and $\mathcal{E}_0 = O$ the iterative adversary fails, while the non-iterative adversary succeeds. Choosing the proportions of I, O, J and P and the edges between them suitably, we can make sure that choosing $\mathcal{E}_1$ and $\mathcal{E}_0$ differently is not advantageous for the adversary and therefore G is indeed robust against the iterative strong adversary. The main result of this paper is to show that the graph G has indeed the properties outlined above.

Theorem 1.2. *For all $0 < \mu < 1/2$ and $1/6 < \delta < 1/2$ there are $\varepsilon_1, \varepsilon_2, d > 0$ such that G is whp robust against the iterative strong adversary, but not against the non-iterative strong adversary.*

Note that $\delta > 1/6$ is a necessary constraint for our construction, but we are certain that there is an example for smaller δ as well. Permissible values in Theorem 1.2 are, e.g. $\mu = \delta = 1/5$, $\varepsilon_1 = 10^{-2}$, $d = 10^{-4}$ and $\varepsilon_2 = 10^{-6}$. The remainder of this paper will consist of the proof of Theorem 1.2. We first state and prove a well known description of the edge distribution of random graphs and then show the claimed (non-) robustness.

2 Proof

For a graph $G = (V, E)$ let $N(v) = \{w \in V \mid (v, w) \in E\}$ be the set of neighbours of v. We begin with a statement about the distribution of edges in random graphs.

Lemma 2.1. *Let $\varepsilon > 0$. The Erdős-Rényi random graph $G(n, p)$ with vertex set V and $p \geq \varepsilon$ has whp the following property. For any set $S \subseteq V$ of size $|S| \geq \varepsilon n$ there is a set $X_S \subset V \setminus S$ of size at most $4\varepsilon^{-3}(\ln \varepsilon^{-1} + 2)$ such that*

$$\forall v \in (V \setminus S) \setminus X_S : \big||N(v) \cap S| - p|S|\big| \leq \varepsilon p|S|.$$

Similar versions of Lemma 2.1 with (somehow) different bounds exist in the literature, see for example [14, Lem. IV.1 and IV.3]. However, as we did not find the exact statement we will need in the literature we include a proof. We will utilize the following Chernoff bound.

Theorem 2.2 ([4], Cor 7.11)**.** *Let X be a binomially distributed random variable. Then*

$$P\Big(|X - \mathbb{E}[X]| > \delta\mathbb{E}[X]\Big) \leq 2\exp\left(-\min\{\delta^2, \delta\}\mathbb{E}[X]/4\right), \qquad \delta > 0.$$

Proof of Lemma 2.1. Let $S \subseteq V, |S| \geq \varepsilon n$ and let

$$X_S = \left\{v \in V \setminus S \;\middle|\; \big||N(v) \cap S| - p|S|\big| > \varepsilon p|S|\right\}$$

be the set of vertices not satisfying the claim of the lemma. The number of neighbours of any vertex $v \in V \setminus S$ is a binomially distributed random variable, $|N(v) \cap S| = \text{Bin}(|S|, p)$, and the expected number of neighbours of v in S is $p|S|$. Thus the probability of $v \in X_S$ can be bounded with Theorem 2.2 by

$$P\left(\big||N(v) \cap S| - p|S|\big| > \varepsilon p|S|\right) \leq \exp\left(-\varepsilon^2 p|S|/4\right).$$

Let furthermore $t \in \mathbb{N}$; the probability that t distinct vertices are in X_S is at most $\exp(-\varepsilon^2 p|S|/4 \cdot t)$ as the events of vertices being elements of X_S are independent. There are $\binom{n}{k} \leq \left(\frac{en}{k}\right)^k$ possibilities to choose S, a set of size $k \geq \varepsilon n$. Hence the probability that for fixed $k \geq \varepsilon n$ there is a set $S, |S| = k$ such that $|X_S| = t$ is by union bound at most

$$\exp\left(k\ln(en/k) - \frac{\varepsilon^2 pk}{4} \cdot t\right) \leq \exp\left(k\left(-\ln \varepsilon + 1 - \frac{\varepsilon^3}{4} \cdot t\right)\right),$$

where we used the assumption that $p \geq \varepsilon$. Thus, if $t \geq 4\varepsilon^{-3}(\ln \varepsilon^{-1} + 2)$ this expression is $\leq e^{-k}$ and summing over $k \geq \varepsilon n$ yields the claim. $\square$

This concludes the preparations. Next we prove the main theorem, by proving the two claims separately. We show the robustness of G against the iterative strong adversary first.

Lemma 2.3. *For all $0 < \mu < 1/2$ and $1/6 < \delta < 1/2$ there are values ε_1, ε_2, $d > 0$ such that G is whp robust against the iterative strong adversary.*

Proof. Let $0 < \mu < 1/2,\ 1/6 < \delta < 1/2$ and $\varepsilon_1, \varepsilon_2,\ d > 0$ such that

$$\varepsilon_1 < \min\left\{\frac{\delta\mu}{2}, \frac{4\delta}{1/2+\delta} - 1\right\} \tag{2.1}$$

and furthermore

$$d < \min\left\{\frac{\varepsilon_1\delta}{1/2+\delta}, \frac{\varepsilon_1\delta\mu}{4}, \frac{1-\mu-2\delta\mu}{3}\right\} \tag{2.2}$$

as well as

$$\varepsilon_2 < \min\left\{\frac{d}{6}\left(\frac{4\delta}{1/2+\delta} - 1 - \varepsilon_1\right), \frac{1/2-\delta}{1+2\delta}\right\}. \tag{2.3}$$

We will show that for any choice of experts, at the end of the dissemination the majority will be labeled '1' thus proving robustness. Let therefore $\mathcal{E} = \mathcal{E}_1 \cup \mathcal{E}_0$ be any set of experts as chosen by the iterative strong adversary and define

$$i_1 := |I \cap \mathcal{E}_1|, \quad j_1 := |J \cap \mathcal{E}_1|, \quad o_1 := |O \cap \mathcal{E}_1|, \quad p_1 := |P \cap \mathcal{E}_1|, \quad d_1 := |D \cap \mathcal{E}_1|$$

as well as

$$i_0 := |I \cap \mathcal{E}_0|, \quad j_0 := |J \cap \mathcal{E}_0|, \quad o_0 := |O \cap \mathcal{E}_0|, \quad p_0 := |P \cap \mathcal{E}_0|, \quad d_0 := |D \cap \mathcal{E}_0|.$$

By definition of the model we have that $|\mathcal{E}_1| = (1/2+\delta)\,\mu n$ as well as $|\mathcal{E}_0| = (1/2-\delta)\,\mu n$ and therefore

$$i_1 + j_1 + o_1 + p_1 + d_1 = \left(\frac{1}{2} + \delta\right)\mu n \quad \text{and} \quad i_0 + j_0 + o_0 + p_0 + d_0 = \left(\frac{1}{2} - \delta\right)\mu n, \tag{2.4}$$

which readily implies that

$$2\delta\mu n = \frac{2\delta}{1/2+\delta}(i_1 + j_1 + p_1 + o_1 + d_1). \tag{2.5}$$

We will see that the iterative dissemination will be finished after two rounds only. We start by determining the label of each vertex in the different components after the first round of dissemination. This is decided by the difference in '0'/'1' labeled expert neighbours. Consider the difference

$$\Delta(v) := |N(v) \cap \mathcal{E}_1| - |N(v) \cap \mathcal{E}_0|, \quad v \in V.$$

In particular, $\Delta(v) > 0$ means that $v \in V \setminus (\mathcal{E}_1 \cup \mathcal{E}_0)$ will be labeled '1' and $\Delta(v) < 0$ means it will be labeled '0'. Note that vertices v with $\Delta(v) = 0$ could be either labeled randomly (if they have the same positive number of '0'/'1' labeled neighbours) or not at all in this round.

We begin with a vertex $v \in O$. Using the construction of G and (2.4) we get

$$|N(v) \cap \mathcal{E}_1| = i_1 + j_1 + p_1 + o_1 = \left(\frac{1}{2} + \delta\right)\mu n - d_1$$

and similarly

$$|N(v) \cap \mathcal{E}_0| = i_0 + j_0 + p_0 + o_0 = \left(\frac{1}{2} - \delta\right)\mu n - d_0.$$

Combining these two equations, $d < \varepsilon_1 < \delta\mu/2$ given by (2.2) and (2.1) implies

$$\Delta(v) = 2\delta\mu n - (d_1 - d_0) > 0 \qquad \forall\ v \in O.$$

We continue with $v \in J$. Using Lemma 2.1, we get that for all $\varepsilon > 0$ whp there is $J_P \subset J$, $|J_P| \le 4\varepsilon^{-3}(\ln \varepsilon^{-1} + 2)$ such that

$$\big||N(v) \cap P \cap \mathcal{E}_1| - p_{JP} \cdot p_1\big| \le \varepsilon \cdot p_{JP} \cdot p_1 + \varepsilon n \quad \text{for all } v \in J \setminus J_P.$$

As $\varepsilon > 0$ is arbitrary we infer that

$$|N(v) \cap P \cap \mathcal{E}_1| = p_{JP} \cdot p_1 + o(n) \quad \text{for all } v \in J \setminus J_P.$$

Completely analogous calculations for I and $\mathcal{E}_0$ yield that whp there is $J' \subset J$, $|J'| = o(n)$ such that for all $v \in J \setminus J'$

$$\begin{aligned} |N(v) \cap \mathcal{E}_1| &= p_{IJ} \cdot i_1 + j_1 + p_{JP} \cdot p_1 + o_1 + o(n) \\ &= \left(\frac{1}{2} + \delta\right) \mu n - (1 - p_{IJ}) i_1 - (1 - p_{JP}) p_1 - d_1 + o(n) \end{aligned}$$

and

$$\begin{aligned} |N(v) \cap \mathcal{E}_0| &= p_{IJ} \cdot i_0 + j_0 + p_{JP} \cdot p_0 + o_0 + o(n) \\ &= \left(\frac{1}{2} - \delta\right) \mu n - (1 - p_{IJ}) i_0 - (1 - p_{JP}) p_0 - d_0 + o(n). \end{aligned}$$

Computing the difference of the above expressions we get for all $v \in J \setminus J'$

$$\begin{aligned} \Delta(v) &= 2\delta\mu n - (1 - p_{IJ})(i_1 - i_0) - (1 - p_{JP})(p_1 - p_0) - (d_1 - d_0) + o(n) \\ &= 2\delta\mu n - \left(\frac{2\delta}{1/2 + \delta} - \varepsilon_1\right) \Big((i_1 - i_0) + (p_1 - p_0)\Big) - (d_1 - d_0) + o(n) \\ &= 2\delta\mu n - \frac{2\delta}{1/2 + \delta}\Big((i_1 - i_0) + (p_1 - p_0)\Big) + \varepsilon_1\Big((i_1 - i_0) + (p_1 - p_0)\Big) - (d_1 - d_0) + o(n). \end{aligned}$$

Applying (2.5) and (2.4) we can obtain a lower bound for $\Delta(v), v \in J \setminus J'$

$$\begin{aligned} \Delta(v) &\ge \frac{2\delta}{1/2 + \delta}(j_1 + o_1 + d_1 + i_0 + p_0) + \varepsilon_1\Big((i_1 - i_0) + (p_1 - p_0)\Big) - d_1 + o(n) \\ &\ge \varepsilon_1(i_1 + j_1 + p_1 + o_1 + d_1) + \left(\frac{2\delta}{1/2 + \delta} - \varepsilon_1\right)(i_0 + p_0) - d_1. \end{aligned}$$

According to (2.1) and (2.2) we have $\varepsilon_1 < 2\delta/(1/2 + \delta)$ as well as $d < \varepsilon_1 \delta\mu/4$ and therefore

$$\Delta(v) \ge \varepsilon_1 \cdot 2\delta\mu n - dn > 0 \qquad \forall\, v \in J \setminus J'.$$

Next we look at $v \in I$. Using again Lemma 2.1 and (2.4) we infer that whp there is $I' \subset I$, $|I'| = o(n)$ such that for all $v \in I \setminus I'$

$$\begin{aligned} |N(v) \cap \mathcal{E}_1| &= i_1 + p_{IJ} \cdot j_1 + p_{IP} \cdot p_1 + o_1 + o(n) \\ &= \left(\frac{1}{2} + \delta\right) \mu n - (1 - p_{IJ})\, j_1 - (1 - p_{IP})\, p_1 - d_1 + o(n) \end{aligned}$$

and

$$\begin{aligned} |N(v) \cap \mathcal{E}_0| &= i_0 + p_{IJ} \cdot j_0 + p_{IP} \cdot p_0 + o_0 + o(n) \\ &= \left(\frac{1}{2} - \delta\right) \mu n - (1 - p_{IJ})\, j_0 - (1 - p_{IP})\, p_0 - d_0 + o(n). \end{aligned}$$

By combining those bounds we obtain for all $v \in I \setminus I'$

$$\begin{aligned}\Delta(v) &= 2\delta\mu n - (1-p_{IJ})\,(j_1 - j_0) - (1-p_{IP})\,(p_1 - p_0) - (d_1 - d_0) + o(n)\\ &= 2\delta\mu n - \frac{2\delta}{1/2+\delta}\Big((j_1 - j_0) + (p_1 - p_0)\Big) - (d_1 - d_0) + \varepsilon_1(j_1 - j_0) - \varepsilon_2(p_1 - p_0) + o(n).\end{aligned} \tag{2.6}$$

Before we conclusively determine $\Delta(v)$ for $v \in I \setminus I'$ we look at vertices $v \in P$. Using once more Lemma 2.1 and (2.4) we infer that whp there is $P' \subset P$, $|P'| = o(n)$ such that for all $v \in P \setminus P'$

$$\begin{aligned}|N(v) \cap \mathcal{E}_1| &= p_{IP} \cdot i_1 + p_{JP} \cdot j_1 + p_1 + o_1 + o(n)\\ &= \left(\frac{1}{2} + \delta\right)\mu n - (1-p_{IP})\,i_1 - (1-p_{JP})\,j_1 - d_1 + o(n)\end{aligned}$$

and

$$\begin{aligned}|N(v) \cap \mathcal{E}_0| &= p_{IP} \cdot i_0 + p_{JP} \cdot j_0 + p_0 + o_0 + o(n)\\ &= \left(\frac{1}{2} - \delta\right)\mu n - (1-p_{IP})\,i_0 - (1-p_{JP})\,j_0 - d_0 + o(n).\end{aligned}$$

Together these two expressions yield for all $v \in P \setminus P'$

$$\Delta(v) = 2\delta\mu n - \frac{2\delta}{1/2+\delta}\Big((i_1 - i_0) + (j_1 - j_0)\Big) - (d_1 - d_0) + \varepsilon_1(j_1 - j_0) - \varepsilon_2(i_1 - i_0) + o(n). \tag{2.7}$$

We argue next, that either "$\Delta(v) < 0$ for some $v \in I \setminus I'$" or "$\Delta(v) < 0$ for some $v \in P \setminus P'$" but never both. To see this, observe that

$$\text{“}\Delta(v) < 0 \text{ for some } v \in I \setminus I' \quad \text{and} \quad \Delta(v) < 0 \text{ for some } v \in P \setminus P\text{”}$$

implies that

$$(i_1 - i_0) + (j_1 - j_0) \quad \text{and} \quad (j_1 - j_0) + (p_1 - p_0) \quad \text{are both} \quad \geq ((1/2+\delta)\mu - \varepsilon_1)n. \tag{2.8}$$

Otherwise assumptions (2.2) and (2.3) give that $\varepsilon_2 < d/6$ as well as $d < \varepsilon_1\delta/(1/2+\delta)$ and therefore either by (2.6)

$$\Delta(v) \geq \frac{2\delta}{1/2+\delta}\varepsilon_1 n - dn - \varepsilon_2 n > 0, \qquad \text{for all } v \in I \setminus I'$$

or by (2.7)

$$\Delta(v) \geq \frac{2\delta}{1/2+\delta}\varepsilon_1 n - dn - \varepsilon_2 n > 0, \qquad \text{for all } v \in P \setminus P'.$$

However, as $(i_1 - i_0) + (p_1 - p_0) \leq (1/2+\delta)\mu n$ equation (2.8) implies that $j_1 - j_0 \geq (\delta\mu - \varepsilon_1)n$. Again (2.1),(2.2) and (2.3) give that $\varepsilon_1 < \delta\mu/2$, $d < \varepsilon_1\delta\mu/$ and $\varepsilon_2 < d/6$ and thus (2.6) yields

$$\Delta(v) \geq \varepsilon_1 \cdot (\delta\mu - \varepsilon_1)n - dn - \varepsilon_2 n > 0, \qquad \text{for all } v \in I \setminus I'.$$

Summarizing, we have shown that $\Delta(v) > 0$ for all $v \in (J \setminus J') \cup O$ and either $\Delta(v) > 0$ for all $v \in I \setminus I'$ or $\Delta(v) > 0$ for all $v \in P \setminus P'$.

In the rest of the proof we consider the second round of the iterative dissemination process. We will distinguish two cases. Assume first that $j_0 + d_0 < (d/2 - \varepsilon_2)n$. As $\Delta(v) > 0$ for all $v \in J \setminus J'$ we infer that at most $(d/2 - \varepsilon_2)n + o(n)$ vertices in D will be labeled '0' after the second round of the dissemination process, all other vertices in D will be labeled '1'. Thus counting the total number of vertices labeled '1' after the process, we get in this case for n large enough

$$\begin{aligned}\#(\text{vertices labeled '1'}) &> |I \setminus I'| \;+\; |J \setminus J'| \;+\; |O| \;+\left(\frac{d}{2} + \varepsilon_2 - o(1)\right) n - |\mathcal{E}_0| \\ &= (1-d)\frac{n}{2} + \frac{dn}{2} + \varepsilon_2 n - o(n) > \frac{n}{2}.\end{aligned}$$

We are left with the case $j_0 + d_0 \geq (d/2 - \varepsilon_2)n$. Observe that $d_1 < (d/2 + \varepsilon_2)n$ as otherwise the conclusion of the previous case applies. We revisit $\Delta(v),\ v \in P \setminus P'$ using (2.7) and (2.5)

$$\begin{aligned}\Delta(v) &= \frac{2\delta}{1/2+\delta}(p_1 + o_1 + d_1 + i_0 + j_0) - (d_1 - d_0) + \varepsilon_1(j_1 - j_0) - \varepsilon_2(i_1 - i_0) + o(n) \\ &\geq \left(\frac{2\delta}{1/2+\delta} - \varepsilon_1\right) j_0 + d_0 - \left(1 - \frac{2\delta}{1/2+\delta}\right) d_1 - \varepsilon_2 i_1 + o(n).\end{aligned}$$

Using the assumptions $j_0 + d_0 \geq (d/2 - \varepsilon_2)n$ and $d_1 < (d/2 + \varepsilon_2)n$, this simplifies to

$$\Delta(v) > \left(\frac{4\delta}{1/2+\delta} - 1 - \varepsilon_1\right) dn/2 - 3\varepsilon_2 n + o(n).$$

Assumption (2.3) guarantees that $\Delta(v) > 0,\ v \in P \setminus P'$ and thus in this case for n large enough

$$\begin{aligned}\#(\text{vertices labeled '1'}) &> |I \setminus I'| \;+\; |J \setminus J'| \;+\; |O| \;+\; |P \setminus P'| \;-\; |\mathcal{E}_0| \\ &= (1-d)n - \left(\frac{1}{2} - \delta\right)\mu n - o(n) > \frac{n}{2},\end{aligned}$$

and the proof is completed. ☐

The next lemma together with Lemma 2.3 implies Theorem 1.2.

Lemma 2.4. *For all $0 < \mu < 1/2$ and $1/6 < \delta < 1/2$ there are values $\varepsilon_1, \varepsilon_2, d > 0$ such that whp G is not robust against the non-iterative strong adversary.*

Proof. Let $0 < \mu < 1/2$ and $1/6 < \delta < 1/2$ and ε_1, ε_2, $d > 0$ as given in (2.1) to (2.3). We show that G is indeed not robust by giving a suitable choice of the expert set. Set $\mathcal{E} = \mathcal{E}_1 \cup \mathcal{E}_0$ with $\mathcal{E}_1 = I$ and $\mathcal{E}_0 = O$. By definition, these sets have matching cardinalities. We compute the quantity

$$\Delta(v) = |N(v) \cap \mathcal{E}_1| - |N(v) \cap \mathcal{E}_2|$$

for vertices $v \in P$ to find their labels. Using $\varepsilon_2 < (1/2 - \delta)/(1 + 2\delta)$ by (2.3) and Theorem 2.2 we readily obtain that whp for all $v \in P$

$$\Delta(v) \leq (p_{IP} + o(1))|\mathcal{E}_1| - |\mathcal{E}_0| = \left(\frac{1/2-\delta}{1/2+\delta} - \varepsilon_2 + o(1)\right)|\mathcal{E}_1| - |\mathcal{E}_0| = -(\varepsilon_2 + o(1))|\mathcal{E}_1| < 0.$$

Therefore the set of vertices labeled '0' contains $O \cup P$ which has cardinality $(1-d)\frac{n}{2}$. However, vertices in D do not have any expert neighbours and as we are in the non-iterative setting, those vertices will be decided uniformly at random. Hence with probability $1/2$ there will be at least $dn/2 + 1$ vertices labeled '0' in D and therefore G is not robust against the non-iterative strong adversary. ☐

References

[1] N. Alon, M. Babaioff, R. Karidi, R. Lavi, and M. Tennenholtz. Sequential voting with externalities: herding in social networks. In *Proceedings of the 13th ACM Conference on Electronic Commerce (EC)*. ACM, 2012.

[2] N. Alon, M. Feldman, O. Lev, and M. Tennenholtz. How Robust Is the Wisdom of the Crowds? In *Proceedings of the 24th International Joint Conference on Artificial Intelligence (IJCAI 2015)*, pages 2055–2061, 2015.

[3] R. Andersen, C. Borgs, J. Chayes, U. Feige, A. Flaxman, A. Kalai, V. Mirrokni, and M. Tennenholtz. Trust-based recommendation systems: an axiomatic approach. In *Proceedings of the 17th International Conference on World Wide Web*, pages 199–208, 2008.

[4] S. Arora and B. Barak. *Computational complexity: a modern approach.* Cambridge University Press, 2009.

[5] W. Aspinall. A route to more tractable expert advice. *Nature*, 463(7279):294–295, 2010.

[6] S. Bikhchandani, D. Hirshleifer, and I. Welch. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy*, 100(5):992–1026, 1992.

[7] S. Bikhchandani, D. Hirshleifer, and I. Welch. Learning from the behavior of others: Conformity, fads, and informational cascades. *Journal of economic perspectives*, 12(3):151–170, 1998.

[8] D. V. Budescu and E. Chen. Identifying expertise to extract the wisdom of crowds. *Management Science*, 61(2):267–280, 2015.

[9] R. M. Cooke and L. L. Goossens. Tu delft expert judgment data base. *Reliability Engineering & System Safety*, 93(5):657–674, 2008.

[10] R. Daknama. *Theoretical runtime bounds for information spreading and a new vehicle routing algorithm.* PhD thesis, lmu, 2018.

[11] E. Dwoskin and C. Timberg. How merchants use Facebook to flood Amazon with fake reviews, 2018. `https://www.washingtonpost.com/business/economy/how-merchants-secretly-use-facebook-to-flood-amazon-with-fake-reviews/2018/04/23/5dad1e30-4392-11e8-8569-26fda6b404c7_story.html?`, visited 02.11.21.

[12] P. Faliszewski, J. Rothe, and H. Moulin. *Control and Bribery in Voting*, page 146–168. Cambridge University Press, 2016.

[13] M. Feldman, N. Immorlica, B. Lucier, and S. M. Weinberg. Reaching consensus via non-bayesian asynchronous learning in social networks. *Approximation, Randomization, and Combinatorial Optimization*, pages 192–208, 2014.

[14] N. Fountoulakis, A. Huber, and K. Panagiotou. Reliable broadcasting in random networks and the effect of density. In *2010 Proceedings IEEE INFOCOM*, pages 1–9. IEEE, 2010.

[15] U. Grandi and P. Turrini. A network-based rating system and its resistance to bribery. In *Proceedings of the 25th International Joint Conference on Artificial Intelligence*, pages 301–307, 2016.

[16] O. Lev and M. Tennenholtz. Group recommendations: Axioms, impossibilities, and random walks. *arXiv preprint arXiv:1707.08755*, 2017.

[17] K. Masters. The New Black Hat Tactics Amazon Sellers Are Using To Take Out Their Competition, 2019. `https://www.forbes.com/sites/kirimasters/2019/02/19/the-new-black-hat-tactics-amazon-sellers-are-using-to-take-out-their-competition/#5440a8c53f58`, visited 02.11.21.

[18] B. Mellers, L. Ungar, J. Baron, J. Ramos, B. Gurcay, K. Fincher, S. E. Scott, D. Moore, P. Atanasov, S. A. Swift, et al. Psychological strategies for winning a geopolitical forecasting tournament. *Psychological Science*, 25(5):1106–1115, 2014.

[19] M. G. Morgan. Use (and abuse) of expert elicitation in support of decision making for public policy. *Proceedings of the National academy of Sciences*, 111(20):7176–7184, 2014.

[20] T. Noonan. ReviewMeta, 2016. `https://reviewmeta.com`, visited 02.11.21.

[21] T. I. Oprea, C. G. Bologa, S. Boyer, R. F. Curpan, R. C. Glen, A. L. Hopkins, C. A. Lipinski, G. R. Marshall, Y. C. Martin, L. Ostopovici-Halip, et al. A crowdsourcing evaluation of the nih chemical probes. *Nature chemical biology*, 5(7):441–447, 2009.

[22] E. M. Rogers. *Diffusion of innovations (5th ed.).* Free Press, 2003.

[23] D. Watts. A simple model of global cascades on random networks. *Proceedings of the National Academy of Sciences*, 99(9):5766–5771, 2002.

[24] H. P. Young. Innovation diffusion in heterogeneous populations: Contagion, social influence, and social learning. *American economic review*, 99(5):1899–1924, 2009.