Adaptive Rumor Spreading

José Correa[1], Marcos Kiwi[2(✉)], Neil Olver[3], and Alberto Vera[1]

[1] Department of Industrial Engineering, Universidad de Chile, Santiago, Chile
[2] Department of Mathematical Engineering and Center for Mathematical Modelling, Universidad de Chile, Santiago, Chile
mkiwi@dim.uchile.cl
[3] Department of Econometrics and Operations Research, Vrije Universiteit Amsterdam; and CWI, Amsterdam, The Netherlands

Abstract. Motivated by the recent emergence of the so-called opportunistic communication networks, we consider the issue of adaptivity in the most basic continuous time (asynchronous) rumor spreading process. In our setting a rumor has to be spread to a population; the service provider can push it at any time to any node in the network and has unit cost for doing this. On the other hand, as usual in rumor spreading, nodes share the rumor upon meeting and this imposes no cost on the service provider. Rather than fixing a budget on the number of pushes, we consider the cost version of the problem with a fixed deadline and ask for a minimum cost strategy that spreads the rumor to every node. A non-adaptive strategy can only intervene at the beginning and at the end, while an adaptive strategy has full knowledge and intervention capabilities. Our main result is that in the homogeneous case (where every pair of nodes randomly meet at the same rate) the benefit of adaptivity is bounded by a constant. This requires a subtle analysis of the underlying random process that is of interest in its own right.

1 Introduction

A basic question in the study of social networks concerns the diffusion of a rumor, which may refer to adopting a new technology, updating content on a cell phone, or buying a new product or service. In this setting we are given a network in which vertices represent agents and edges represent social links. Initially, a single agent knows the rumor and we would like to estimate the time by which the full network is informed. The flow of information is governed by a certain stochastic process which may evolve in discrete or continuous time. The most widely studied discrete time models are the push model and the pull model; in the former at each time step a vertex knowing the rumor pushes it to a random neighbor, while in the latter a vertex not knowing pulls the rumor from

Supported by the Millennium Nucleus Information and Coordination in Networks ICM/FIC RC130003, CONICYT via Basal in Applied Mathematics, and an NWO Veni grant.

E. Markakis and G. Schäfer (Eds.): WINE 2015, LNCS 9470, pp. 272–285, 2015.
DOI: 10.1007/978-3-662-48995-6_20

a random neighbor. On the other hand, in the continuous time (asynchronous) model, every pair of connected vertices meet at random times following a Poisson process. The latter model was first formulated by Boyd et al. [8] as a way around the unrealistic time-synchronization issue implicit in discrete time models.

The diffusion of information through a social network has also posed a number of fundamental algorithmic questions, particularly in so-called *viral marketing campaigns.* The study of how the initial selection of vertices (who adopt a new product or gets it for free) influences further adoption through a cascading effect was pioneered by Domingos and Richardson [14], and rigorously addressed by Kempe et al. [20], who designed approximation algorithms for the influence maximization problem subject to a budget constraint on the number of initial nodes to which the rumor is pushed. Interestingly, these viral marketing ideas have permeated not only the technological industry, but also more traditional markets like automotive ones [4]. Unlike the rumor spreading process described above, Kempe et al. [20] consider "static" diffusion models, particularly the *independent cascade model*, in which the spread of the rumor is probabilistic, but time plays no role. An alternative approach, which we take in this paper (see also [17]), is to keep the standard rumor spreading process, but rather than fixing a budget on the initial set of selected vertices, fix a time horizon and only account for the vertices that receive the rumor within this time.

When working with a dynamic diffusion model a new problem pops up, namely that of *adaptivity.* Already Domingos and Richardson [14] identify this issue and state that: *A more sophisticated alternative would be to plan a marketing strategy by explicitly simulating the sequential adoption of a product by customers given different interventions at different times, and adapting the strategy as new data on customer response arrives.* Along these lines, Seeman and Singer [24] consider a two stage extension of the Kempe et al. model.

The central concern of this paper is that of adaptivity when speeding up rumor spreading on a social network. More precisely, we are given a network and a fixed deadline. As stated above, the diffusion model is the standard for asynchronous rumor spreading. Thus, every pair of connected vertices meet at random times following a Poisson process, and the rumor is spread whenever, upon meeting, one vertex knows the rumor and the other one does not. Along the way we are able to *push* the rumor to any vertex in order to speed up the diffusion process. We consider both the profit maximization and the cost minimization versions of the problem, which are equivalent from an optimization viewpoint. When maximizing profit we get zero profit for every vertex to which the rumor is pushed (say because we are giving the product for free) and get unit profit whenever a vertex gets the rumor through the diffusion process. The objective is thus to maximize the number of vertices that got the rumor through the diffusion process within the time horizon. The cost minimization version is exactly the opposite; every time we push the rumor we make a unit payment and if a vertex gets the rumor through the diffusion process this cost is not realized. The goal is to minimize the total payment made by the time horizon subject to the constraint that everyone should be informed by then.

A non-adaptive policy does not track the evolution of the process and therefore can only push the rumor at the starting time (and also at the deadline in the cost minimization version). In contrast, an adaptive policy may monitor the evolution of the diffusion process and intervene by pushing the rumor to additional vertices. The main contribution of this paper is to show that the advantage of adaptivity is small (in terms of cost or profit) in the setting of *homogeneous* networks, where interactions occur at the same rate between any pair of nodes. While the homogeneous case seems unrealistic, it is already highly nontrivial and we believe it will be a useful first step towards tackling more general situations.

The seemingly less natural cost minimization version of our problem actually constitutes our main motivation and finds its roots in opportunistic communication networks. The widespread adoption of networked mobile devices and the deployment of new technologies (3G, 4G), through which ever increasing data intensive services can be delivered, has generated an explosion of mobile data traffic. This trend is likely to continue, thus exacerbating current cellular network data overload [12]. Therefore, it is critical for operators and service providers to design networks and communication mechanisms that can not only handle the current traffic overload, but also allow rapid data dissemination that will be required by next-generation mobile-enabled devices and applications. A promising, less than a decade old, proposal to address the cellular network data overload consists in offloading traffic through so-called opportunistic communications. The key idea is for service providers to push mobile application content to a small subset of interested users through the cellular network and let them opportunistically spread the content to other interested users upon meeting them. Opportunistic communication can occur when mobile device users are (temporarily) in each others proximity, making it possible for their devices to establish local peer-to-peer connections (e.g., via Wifi or Bluetooth). Opportunistic communication based services have been proposed across several domains. For instance, studies have been done in the dissemination of dynamic content such as news using real world data sets, as well as that of traffic update information using dataset of the municipality of Bologna [19,25]. In several of the aforementioned application scenarios, the usefulness and/or relevance of the disseminated content crucially depends on it being opportunely delivered. Moreover, quality of service contractual obligations in subscription based contexts might entail strict deadlines for the delivery of data. Another key issue that arises in this context is the presumed feedback capabilities of network nodes and the service providers' knowledge of how data has propagated through the network up to a given moment of time.

Whitbeck et al. [25] were the first to study a fixed deadline scenario. They propose a *Push-and-Track* framework, where a subset of users receive the content from the infrastructure and start disseminating it epidemically. The main feature of *Push-and-Track* is the closed control loop, this supervises the injection of copies of the content via the infrastructure whenever it estimates that the ad hoc mode alone will fail to achieve full dissemination within the target horizon. Upon reaching the deadline, the system enters into a "panic zone" and pushes the content to all nodes that have not yet received it. Sciancalepore et al. [23]

initiate a more rigorous analysis of *Push-and-Track* type proposals. In particular, they derive formulas (although not explicit algebraic expressions) for the optimal number of nodes to initially push data in order to minimize the overall number of pushes. Furthermore, they propose a control theory based adaptive heuristic.

Our work is thus motivated by the natural question left open by Sciancalepore et al. [23]: whether or not an adaptive strategy, that harnesses the accrued information of how data has propagated through the network up to any given instant, can actually outperform an optimal non-adaptive strategy, and to what extent.

Model and Main Contributions. Consider a network of n nodes labeled by the elements of $[n] := \{1, \ldots, n\}$. Nodes are presumed mobile and such that the encounters of any two nodes i and j, $i \neq j$, are governed by a Poisson process of rate $\lambda_{i,j}$. Thus the time elapsed between two consecutive encounters of i and j is distributed as an exponential random variable of rate $\lambda_{i,j}$, henceforth denoted by $\text{Exp}(\lambda_{i,j})$. All these random variables are independent, including those associated to distinct inter-encounter intervals for the same pair of nodes. As usual, if upon encountering each other one node is informed (i.e., is active) and the other is not, then the information is spread. We refer to the case where all the rates are identical as the *homogeneous* case; our main positive result will be for this setting.

We assume that there is a service provider who wishes to cost efficiently disseminate one unit of information to all nodes within a *deadline* of time $\tau > 0$. The set of nodes that posses the unit of information at time t will be denoted $S(t) \subseteq [n]$ and referred to as the set of *active notes at time* t. Initially (at time $t = 0$), the service provider selects a set of nodes $S(0)$ and *activates* them by pushing to them the unit of data of interest. Subsequently, nodes become active by either one of the following two mechanisms:

- Opportunistic communication: If nodes $i \neq j$ encounter each other at time t and either i or j belong to $S(t^-)$, say j, then i becomes informed at time t, i.e., i belongs to $S(t')$ for all $t' \geq t$. Here we used the convention $S(t^-) := \cup_{0 \leq t' < t} S(t')$. When a node becomes informed via opportunistic communication, it signals the service provider that his state has changed.
- Pushes: Because of the network's feedback capabilities, at any time $0 \leq t \leq \tau$, the service provider has full knowledge of the evolution of the set of active nodes, i.e., of $(S(t'))_{0 \leq t' < t}$, and based on this knowledge she decides whether or not, and which nodes to activate. Formally, at time t a set of nodes $C(t) \subseteq [n] \setminus S(t^-)$ is chosen and added to $S(t)$, in which case we say that at time t the nodes in $C(t)$ are activated and $|C(t)|$ pushes performed.

For most of the paper we deal with the cost minimization version of the problem, in which the service provider incurs a unit cost for activating a single node, independent of the nodes label and the time it happens. When the deadline is reached, all nodes not in $S(\tau)$ must be activated at a total cost of $n - |S(\tau)|$. Note that this is equivalent, from an optimization perspective, to the maximization problem where the service provider gets unit profit for each node informed via opportunistic communication and zero profit for the pushes she makes. We say that the service provider's strategy is *non-adaptive* if it can only activate nodes at time $t = 0$ and $t = \tau$. Otherwise, we say that its strategy is *adaptive*.

Of course, the cost of an optimal adaptive strategy is at most that of a non-adaptive strategy that initially activates an optimal number of nodes. A natural question is thus to determine the *adaptivity gap*, defined as the ratio between the cost of an optimal non-adaptive strategy and that of an optimal adaptive strategy. This question is certainly of practical significance – if the adaptivity gap turns out to be close to 1 for realistic ranges of the relevant parameters (n, τ, and $\lambda_{i,j}$'s), then at least from a purely cost effective point of view there is no justification for incurring the overhead of relying in the network's feedback capabilities, nor the extra cost required to implement a more computational demanding adaptive on-line strategy. Our main result is the following.

Theorem 1. *In the homogeneous case, i.e., $\lambda_{i,j} = \lambda$ for all $i, j \in [n]$, the adaptivity gap is bounded by a small constant, irrespective of n and the deadline.*

We also show that the adaptivity gap with respect to the profit objective, defined in the obvious way, is at most $(1 + o(1))$.

From a technical viewpoint, our analysis turns out to be significantly different for small, intermediate and large values of τ.

- For sufficiently large values of τ (say $\tau \geq \frac{1}{\lambda n}(2 + o(1)) \log n$), activating a single node initially will cause, with high probability, the entire network to be active by the deadline. This follows from classical work on stochastic epidemic models (for an overview, see [16] and [1]). So the optimal nonadaptive policy pays essentially 1, and the advantage of adaptivity is negligible.
- For the case of small τ we use a coupling argument to formalize the intuition that the process is "too deterministic" for adaptivity to win much. This already implies a $(1 + o(1))$ bound on the adaptivity gap of the profit maximization version of the problem.
- The case of intermediate values of τ is by far the most challenging. Unlike in the case of small τ, the number of nodes initially activated by an optimal nonadaptive policy is relatively small. The implication of this is that the behaviour of the process is initially *not* very concentrated, moreover, fluctuations in the rumor spreading behaviour in this initial phase can have a large impact on the cost. Since the optimal adaptive policy can be rather complicated, we consider a relaxation of an adaptive strategy which may push for free when certain conditions (which are always satisfied by an optimal adaptive strategy) are met; understanding the optimal behaviour of this relaxation turns out to be more tractable. The analysis then involves understanding an underlying martingale accounting for the expected final cost given the current situation.

The result holds also in the synchronous setting with the push protocol, in fact the argument is substantially easier than for our asynchronous setting. The reason for this is that the spreading process is much more predictable, even in the initial phase where few nodes are active. For example, the time required to activate all nodes starting from a single active node is very tightly concentrated [13], whereas, in the asynchronous case, even the time needed to go from one to two active nodes has substantial variance. As such, the result follows along the same lines as the small τ case in the asynchronous model. We defer further discussion on the synchronous model to the full version of this paper.

More General Networks. The adaptivity gap cannot be bounded by a constant in the general inhomogeneous setting. It can be shown that taking a 2-level, k-regular tree, with unit rates on the tree and all other rates 0, and a deadline of $k \log(k \log k)$, yields an adaptivity gap of $\Omega(\log k / \log\log k)$. A slight variant of this construction, with higher rates on the edges adjacent to the leaves, yields an adaptivity gap of $\Omega(\sqrt{n})$. Despite this, there remains a large scope for better understanding what network features affect the adaptivity gap. In particular, we leave bounds on the adaptivity gaps in the following settings as open questions.

- **Good expansion:** communicating pairs are described by a graph with good expansion, and all communicating pairs interact at the same rate. The lower bound constructions crucially exploit very poor connectivity.
- **Metric constraints:** the inverse rate $\lambda_{i,j}^{-1}$ describing the expected time between interactions between i and j satisfy the triangle inequality. This captures the natural idea that if i and j are frequently in the same vicinity, and likewise for j and k, then i and k are likely also frequently nearby.

Further Related Work. The existing literature on rumor spreading is vast, particularly in the discrete time (synchronous) model. The natural problem here is to estimate the time at which every node in the network has the rumor. This question is quite well understood and extremely precise estimates are known when the network is a complete graph [13]. These estimates state that the time is highly concentrated around a logarithmic function of n, depending on the specific protocol. The arguably more realistic continuous time (asynchronous) model is not as well understood [8]. This is possibly due to inherent additional randomness of this process, particularly in the beginning, although logarithmic estimates for the expected time to activate the whole network, starting with one node, have been obtained for various classes of graphs [6].

Viral marketing is also an area of much interest, where models for the diffusion of information have received a lot of attention [22]. Closest to our work is the influence maximization problem, in which the goal is to find a subset, of at most k nodes, maximizing the total final number of informed nodes. The most studied underlying diffusion model is that of *independent cascades*: when node v becomes informed it has a single chance of informing each currently uninformed neighbor w and succeeds with probability p_{vw}. This problem was studied by Kempe et al. [20], who showed that the underlying optimization problem is a monotone submodular maximization problem, and therefore can be approximated efficiently within a factor of $1 - 1/e$. A long list of follow-up papers have studied the problem (see e.g. [7,9,10,21]) as well as several variations (see e.g. [2,11,15,18,24]).

Note. Due to lack of space, proofs are omitted from this extended abstract.

2 Preliminaries

In this section we further specify the model and the notation we will work with. While introducing the model we try to build some intuition and elicit how it behaves. We also establish some basic facts, which both capture some of the aforementioned intuition and will be needed in subsequent sections.

Recall that our study concerns the homogeneous case, i.e., when the rate $\lambda_{i,j}$ is a fixed value independent of the pair of nodes $i \neq j$. Moreover, everything is invariant if the rates and the deadline τ are both scaled by the same amount, so we assume $\lambda_{i,j} = 1/n$ for all $i \neq j$.

Because of symmetry considerations, the specific labels of active nodes is irrelevant and only their number at any given time matters. We henceforth denote by $K(t)$ the number of active nodes at time t for a non-adaptive scheme. Observe that $K(\cdot)$ is right continuous (i.e., $K(t^-) \leq K(t) = K(t^+)$). Also, define

$$u_k(t) := \mathbb{E}(n - K(\tau)|K(t) = k),$$

i.e., the expected number of pushes to be made at the end of the process given $K(t) = k$. We will need some information about the optimal non-adaptive choice $k_{\mathrm{N}} := k_{\mathrm{N}}(\tau)$ for the number of initially active nodes, i.e., the value of $k \in [n-1]$ that minimizes $k + u_k(0)$. For small values of n, one can compute k_{N} and $u_{k_{\mathrm{N}}}(0)$ explicitly. To do so, it is convenient to consider the elapsed time between the i-th and $(i+1)$-th node activation, henceforth denoted X_i. Since X_i is the minimum of $i(n-i)$ random variables distributed according to $\mathrm{Exp}(1/n)$, well known facts imply that X_i is distributed as $\mathrm{Exp}(\lambda_i)$ for $\lambda_i := i(n-i)/n$.

Analogously, let $K^*(t)$ be the number of active nodes at time t for an *optimal* adaptive scheme, still assuming an explicit deadline τ. Since exponentially distributed random variables are memoryless, the optimal adaptive scheme is completely determined by a sequence $0 \leq t_0^* \leq \ldots \leq t_{n-1}^* \leq \tau$ (depending on τ), so if at time $t \in \{t_0^*, \ldots, t_{n-1}^*\}$ it holds that $K^*(t^-) \leq k$, where k is the largest index for which $t_k^* = t$, then the optimal scheme makes $k + 1 - K^*(t^-)$ pushes at time t. We interpret t_k^* as the first time when it is optimal to push more than k rumors. Let $P(t)$ denote the number of pushes performed by the optimal scheme up to (and including) time t, but excluding pushes made at the deadline τ (so $P(\tau) = P(\tau^-)$). Clearly, the cost of an optimal adaptive scheme must equal $\mathbb{E}(n - K^*(\tau^-) + P(\tau^-))$. Hence, on average $\mathbb{E}(K^*(\tau^-) - P(\tau^-))$ nodes are activated via opportunistic communication.

We can now start formally stating results that will be useful later on. Our first claim is that an optimal adaptive scheme will not perform pushes once roughly half the network's nodes become active. The intuition is that making a push when $i \geq n/2$ nodes are active *reduces*, to something less than λ_i, the rate at which nodes become activated (implying higher expected time between successive node activations). Thus, in expectation, there will be less than 1 more active node at time τ^-, a saving that is less than the cost of the push.

Proposition 1. *Optimal non-adaptive never starts with more than* $\lceil (n-1)/2 \rceil$ *active nodes. Furthermore, an optimal adaptive strategy never pushes at some time* t *if* $K^*(t) \geq \lfloor (n-1)/2 \rfloor$, *i.e.,* $t_k^* = \tau$ *for all* $k \geq \lfloor (n-1)/2 \rfloor$.

We can think of the optimal adaptive scheme as having a target minimum number of active nodes that depends only on the current time t. Our next result essentially says that this target is not larger than $k_{\mathrm{N}}(\tau - t)$, the number of initial pushes for the optimal non-adaptive strategy with deadline $\tau - t$.

Proposition 2. *Let* $k \in [n-1]$ *and* $0 \leq t < \tau$. *If* $u_k(t) - u_{k+1}(t) < 1$, *then* $t_k^* > t$, *i.e., adaptive will not push at time* t *if* $K^*(t^-) \geq k$.

To prove these two propositions, we need some information about the optimal adaptive strategy. For this purpose, it is useful to consider $u_k^*(t)$, defined as the expected cost incurred by an optimal adaptive scheme $K^*(\cdot)$ in the remaining time, conditioned on $K^*(t^-) = k$. By exploiting certain recurrences involving the u_k^*'s and their derivatives, we are able to show that $u_k^*(t) - u_{k+1}^*(t) \leq u_k(t) - u_{k+1}(t)$ for all $k \in [n-1], 0 \leq t \leq t_k^*$. Intuitively, this is explained by the enhanced control an adaptive scheme has over the underlying process, since it could choose to push immediately after time t, hence the benefit of being given this extra active node for free at time t is not more than one. From this, the above propositions follow fairly easily.

3 Estimates on the Evolution of the Non-adaptive Process

In this section we give a number of useful estimates on the evolution of the non-adaptive process, as well as characterize the optimal non-adaptive strategy and its cost.

Proposition 3. *If* $t \in [0, \tau]$ *and* $k \in [n-1]$, *then* $u_k(t) = \dfrac{(1+o(1))n}{1 + \frac{k}{n-k} \cdot e^{\tau - t}} + o(1)$.

This result is essentially well-known (see e.g., [3,5]), so we only briefly sketch its proof. The evolution of the process starting from (say) $n/\log n$ active nodes, and all the way until all but $n/\log n$ nodes are active, is highly concentrated. With very high probability, it closely follows the solution of the deterministic differential equation $\frac{dx}{dt} = x(1-x)$, where $x(t)$ denotes the proportion of active nodes at time t. This yields the logistic curve of the above proposition. When there are very few active nodes, $\lambda_i \approx i$, and the process is well approximated by a linear birth process, for which exact analytic results are available. A similar approximation holds when there are very few inactive nodes; stitching together these estimates yields Proposition 3.

We need some more refined estimates on how $u_k(t)$ varies with k and t. These do not follow from Proposition 3, but notice that they would follow immediately if $u_k(t)$ was *exactly* described by the logistic curve.

Lemma 1. *For all $k < n/2$, $u_k(t+h) \leq u_k(t)e^h$ for all $h \leq \tau - t$. Also, if $\tau - t = \omega(1)$, then $u_{k+1}(t) = u_k(t)\big(1 - \frac{1+o(1)}{\lambda_k}\big)$.*

The expected cost of a non-adaptive strategy starting with k pushes is $k + u_k(0)$. This cost is in fact a convex function of k. Again, this would follow immediately if $u_k(t)$ was precisely described by the logistic formula.

Lemma 2. *For every $t \in [0, \tau]$ the sequence $\{k + u_k(t)\}_{k \in [n]}$ is convex. As a consequence, k_{N} can be taken to be the smallest k such that $u_k(0) - u_{k+1}(0) < 1$.*

Now we obtain an estimate of the optimal non-adaptive strategy, i.e., the number k_{N} of nodes activated at the start. The rates λ_k are unimodal (increasing until $n/2$, and then decreasing). Intuitively then, the optimal non-adaptive strategy aims to have $n/2$ active nodes at time $\tau/2$, so that the rates are on average as large as possible during the evolution. The expected amount paid at the end should be roughly the same as the cost paid at the start; cf. [23] (the proof follows immediately by optimizing using the estimate of Proposition 3).

Proposition 4. *Given a deadline τ, the optimal non-adaptive pick is such that*

$$k_{\mathrm{N}} = (1 + o(1))\frac{n}{1 + e^{\tau/2}} \quad \textit{and} \quad u_{k_{\mathrm{N}}}(0) = k_{\mathrm{N}}(1 + o(1)).$$

Thus, the total expected cost of the optimal non-adaptive strategy is $2(1+o(1))k_{\mathrm{N}}$.

4 Additive Gap for Small τ

In this section we consider the case in which τ is small, specifically, $\tau \leq 2 \log\log n$. In this situation, thanks to Proposition 4, the optimal non-adaptive strategy activates $k_{\mathrm{N}} = k_{\mathrm{N}}(\tau) = (1 + o(1))\frac{n}{1+e^{\tau/2}} = \Omega(\frac{n}{\log n})$ nodes initially. This implies that the non-adaptive evolution is highly concentrated. Intuitively, this should be enough to conclude that adaptive cannot obtain a significant advantage; we use a coupling argument to make this precise.

Let $\mathcal{S}$ be any countable collection of points on $\mathbb{R}_+$, with an infinite number of points, with at least one point at 0, and denote by $\mathcal{S}_i$ the position of the i-th point. Associate to $\mathcal{S}$ a counting process $(K^{\mathcal{S}}(t))_{0 \leq t \leq \tau}$ as follows. Let $X_i^{\mathcal{S}} := (\mathcal{S}_{i+1} - \mathcal{S}_i)/\lambda_i$, and let $T_i^{\mathcal{S}} := \sum_{j=1}^{i-1} X_j^{\mathcal{S}}$, so $T_1^{\mathcal{S}} = 0$. Then, for $i \leq n-1$, set $K^{\mathcal{S}}(t) = i$ for all $t \in [T_i^{\mathcal{S}}, T_{i+1}^{\mathcal{S}}) \cap [0, \tau]$, and $K^{\mathcal{S}}(t) = n$ for all $t \in [T_n^{\mathcal{S}}, \tau]$.

Now, let $\mathcal{N}$ be a Poisson point process of unit intensity, and let $\mathcal{N}'$ be obtained by adding k_{N} additional points at the origin to $\mathcal{N}$. Since the inter-activation times $X_i^{\mathcal{N}'}$ are 0 for $i < k_{\mathrm{N}}$, and distributed exponentially of rate λ_i for $k_{\mathrm{N}} \leq i < n$, we have that the law of $(K^{\mathcal{N}'}(t) : t \in [0, \tau])$ is precisely that of the evolution of the non-adaptive process with k_{N} pushes at time 0.

We can interpret an adaptive strategy directly in this perspective. For each $s \in \mathbb{R}_+$, it can decide whether to add a new point at position s, but based only on $\mathcal{N} \cap [0, s]$. In other words, it is a map φ that takes a set of points $\mathcal{S}$ and returns $\varphi(\mathcal{S}) \supseteq \mathcal{S}$, with $0 \in \varphi(\mathcal{S})$, and where $\varphi(\mathcal{S}) \cap [0, t]$ depends only on $\mathcal{S} \cap [0, t]$, for any t.

The resulting evolution is simply $K^{\varphi(\mathcal{S})}$, where the points in $\varphi(\mathcal{S}) \setminus \mathcal{S}$ correspond to pushes. One can see that this has the correct law of an evolution of an adaptive process, and that any adaptive strategy can be so described.

This provides a (somewhat non-obvious, but natural) coupling between the evolution of non-adaptive and adaptive. To exploit this, we relax the provision that adaptive may only look at the past when making its decisions. We define a *clairvoyant* strategy as *any* function ξ where $\xi(\mathcal{S}) \supseteq \mathcal{S}$ and $0 \in \xi(\mathcal{S})$. Clearly the optimal clairvoyant strategy has lower cost than the optimal adaptive one.

Lemma 3. *There is an optimal clairvoyant strategy which adds points only at the origin.*

So we are comparing the optimal non-adaptive strategy, which picks some number k_{N} of initial pushes without any knowledge of $\mathcal{N}$, to the optimal clairvoyant strategy, which picks some optimal number of initial pushes based on $\mathcal{N}$. A concentration argument shows that the extra information is very unlikely to be useful. More precisely, we argue that for any number k of initial pushes, the probability that the total cost paid is less than $2k_{\mathrm{N}} - O(\sqrt{n}\ \mathrm{polylog}(n))$ is polynomially small, and then apply a union bound to conclude the following result.

Lemma 4. *Let* $\frac{\log^2 n}{\sqrt{n}} < \tau \leq 2 \log\log n$. *Then the expected cost of the optimal clairvoyant strategy applied to* $\mathcal{N}$ *is* $2k_{\mathrm{N}} - O(\sqrt{n}\,\mathrm{polylog}(n))$.

Consequences for the Profit Maximization Version. Note that the previous result already implies that for the profit maximization version of the problem the adaptivity gap is $1 + o(1)$. Indeed, if $\tau \geq 2 \log\log n$ then $k_{\mathrm{N}} = o(n)$ and thus the profit of non-adaptive is $n - o(n)$ while adaptive certainly gets at most n. On the other hand, if $\tau \leq 2 \log\log n$ is at least a constant, Lemma 4 implies that the activations that adaptive and non-adaptive make differ by a sub linear term, and since both get a profit which is linear in n the ratio is $1 + o(1)$. Finally, if $\tau = o(1)$ then $k_{\mathrm{N}} = n/2 - o(n)$ so that throughout the rumor spreading process the λ_i's equal $4/(n + o(n))$ and this cannot be changed by an adaptive strategy.

5 Bounding the Adaptivity Gap

We now consider the case where $\tau \geq 2 \log\log n$. Here we need to do more than exploit the concentration of the evolution of the process. If the optimal non-adaptive scheme starts with relatively few pushes at the starting time, there will be a substantial amount of randomness at the beginning of the process, before the epidemic phase transition. Our goal is to show that an adaptive scheme cannot substantially exploit this.

The optimal adaptive strategy is difficult to handle, for example, the optimal choices of t_k^* are determined via an intricate recurrence. As in the last section, it will be very useful to rely on a more tractable lower bound, however, the lower bound we use here is quite different from the clairvoyant lower bound of Sect. 4.

As seen in Proposition 2, conditioned on $K^*(t^-) = k$, adaptive does not push at t if $u_k(t) - u_{k+1}(t) < 1$. We consider a modified set of rules for adaptive. Suppose it may push *for free*, however, if there are k active nodes at some time t, it may only push if $k < n/2$ and $u_k(t) - u_{k+1}(t) \geq 1$. Obviously the optimal adaptive strategy satisfies these restrictions, hence just pays less under these new rules. So the optimal strategy under these modified rules pays even less. The optimal "modified adaptive" strategy is very simple to describe: Since pushes are free, it will push whenever it is allowed to. We will show that the cost of the optimal modified adaptive strategy, which is simply the expected number of inactive nodes at time τ, is within a constant factor of the cost of non-adaptive.

Let $\tilde{K}(t)$ be the number of active nodes at time t using this optimal modified adaptive strategy. Let $\tilde{T}_i := \min\{t : \tilde{K}(t) \geq i\}$, and let $\tilde{P}(t)$ denote the number of pushes up to and including time t. Observe that $\tilde{K}(0) = k_{\mathrm{N}}$, since k_{N} is the first k such that $u_k(t) - u_{k+1}(t) < 1$ by Lemma 2.

If one considers the non-adaptive evolution $K(t)$, $u_{K(t)}(t)$ is precisely the Doob martingale for the number of inactive nodes at time τ. It will be useful to look at a variant of this for the modified adaptive process. Specifically at $\tilde{U}(t) := u_{\tilde{K}(t)}(t)$, i.e., the expected number of inactive nodes at time τ, *given that no pushes are made between time t and τ*. It is a supermartingale, rather than a martingale, since any pushes made by the modified adaptive strategy will decrease the end payment. Since $\tilde{U}(0) = u_{k_{\mathrm{N}}}(0) = (1 + o(1))k_{\mathrm{N}}$, we will be interested in how much smaller (in expectation) $\tilde{U}(\tau)$ is compared to $\tilde{U}(0)$.

Define $\tilde{t}_k := \inf\{t \in [0, \tau] : u_k(t) - u_{k+1}(t) \geq 1\}$, or $\tilde{t}_k = \tau$ in the case the infimum is taken over the empty set. Then, if $\tilde{K}(t) = k$ and $t < \tilde{t}_k$, the modified adaptive strategy clearly cannot push, so $\tilde{t}_k$ is the first time when it is convenient to activate more than k nodes. Since $u_k(t)$ is a strictly increasing function of t, we can equivalently state this as: No push will occur at time t if $\tilde{U}(t^-) < \phi_{\tilde{K}(t^-)}$, where $\phi_k := u_k(\tilde{t}_k)$ for all k. Conversely, if $\tilde{U}(t^-) = \phi_{\tilde{K}(t^-)}$, then the optimal modified adaptive strategy will certainly push. This causes $\tilde{U}(t)$ to jump down by precisely 1 unit. So we will refer to the values ϕ_k as *thresholds*; the process $\tilde{U}(t)$ is always below the current threshold $\phi_{\tilde{K}(t)}$ and, if the threshold is reached, a push will be performed. Moreover, by convexity of the sequence $\{u_k(\cdot)\}_k$, the times $\tilde{t}_k$ are increasing, so only a single push occurs at any moment in time.

The following proposition connects the number of pushes made, i.e., the number of times $\tilde{U}$ reaches the current threshold, with the cost saved by the modified adaptive strategy.

Proposition 5. *The optimal modified adaptive strategy saves one unit of cost with respect to non-adaptive for each push after $t = 0$, i.e.,*

$$\mathbb{E}(n - \tilde{K}(\tau)) = u_{k_{\mathrm{N}}}(0) - \mathbb{E}(\tilde{P}(\tau) - \tilde{P}(0)).$$

We use this as follows. Suppose $\mathbb{E}(\tilde{P}(\tau) - \tilde{P}(0)) \leq C$. Call the non-adaptive cost $c_N := k_{\mathrm{N}} + u_{k_{\mathrm{N}}}(0)$, from Proposition 4 we know that $c_N = 2(1 + o(1))u_{k_{\mathrm{N}}}(0)$.

The adaptivity gap $\rho(n, \tau)$ is clearly bounded by $c_N/\mathbb{E}(n - \tilde{K}(\tau))$, so by Proposition 5

$$\rho(n, \tau) \leq \frac{c_N}{u_{k_{\mathrm{N}}}(0) - \mathbb{E}(\tilde{P}(\tau) - \tilde{P}(0))} = 2(1 + o(1)) \left(1 + \frac{1}{u_{k_{\mathrm{N}}}(0)/C - 1}\right).$$

We also have the trivial upper bound $\rho(n, \tau) \leq c_N = 2(1+o(1))k_{\mathrm{N}}$, just because an adaptive strategy will certainly need to push at least once. The required constant bound on $\rho(n, \tau)$ for the case $\tau \geq 2 \log\log n$ thus follows.

The aim for the rest of the section is to bound $\mathbb{E}(\tilde{P}(\tau) - \tilde{P}(0))$ by a constant. To exploit the characterization of modified adaptive we will use some of the estimates on $u_k(t)$ that we derived in Sect. 3. Recall Lemma 1, which states that $u_{k+1}(t) = u_k(t)(1 - \frac{1+o(1)}{\lambda_k})$ and $u_k(t + h) \leq u_k(t)e^h$ for $t + h \leq \tau$. This has a very straightforward interpretation in terms of $\tilde{U}(t)$: Between activations $\tilde{U}(t)$ grows sub-exponentially, but if at time t there was an activation, then roughly $\tilde{U}(t)$ is multiplied by the factor $1 - 1/\lambda_{\tilde{K}(t^-)}$.

We have now all the ingredients to bound how many times the process $\tilde{U}(t)$ hits the thresholds $\phi_{\tilde{K}(t)}$, which is exactly the number of pushes. We proceed by transforming the process in a number of ways. First, given the (sub)exponential growth, taking logarithms yields a process that between jumps grows no faster than a linear function with unit slope. Secondly, we locally shift the resulting process so that the threshold at any moment of time is moved to zero. This process will always be non-positive; we will be interested in the number of times that it hits zero. Finally, we locally rescale time, as in Sect. 4, so that the distribution of the times of random activations are described by a Poisson point process of unit intensity. We locally rescale the value at the same time, so that still the process increases linearly at unit rate in between jumps. Formally, we define the following transformed process $H(s)$:

$$H(L(t)) := \lambda_{\tilde{K}(t)}\big(\log \tilde{U}(t) - \log \phi_{\tilde{K}(t)}\big), \qquad \text{where } L(t) := \int_0^t \lambda_{\tilde{K}(x)} dx.$$

An illustration of the evolution of $\tilde{U}$ and the corresponding transformed process H is shown in Fig. 1. Notice how upon each random activation or push the threshold increases, while $\tilde{U}$ jumps down.

We are interested in the number of times that H reaches 0, as this corresponds to the number of pushes after time 0. It is convenient to consider instead

$$H'(L(t)) := H(L(t)) + \tilde{P}(t) - \tilde{P}(0).$$

The process H jumps down immediately whenever it reaches 0, and in fact the size of this jump is larger than 1. Very roughly speaking, H' cancels out these jumps (actually it may still jump down, but by a smaller amount), while the jumps corresponding to random activations are unaffected. It can easily be shown that the number of pushes is bounded by $\max\{0, 1 + \sup_{0 \leq r \leq t} H'(L(r))\}$.

So all that remains is to bound the expected supremum of H'. The reason that this is possible is very simple: Through most of its evolution, the process has

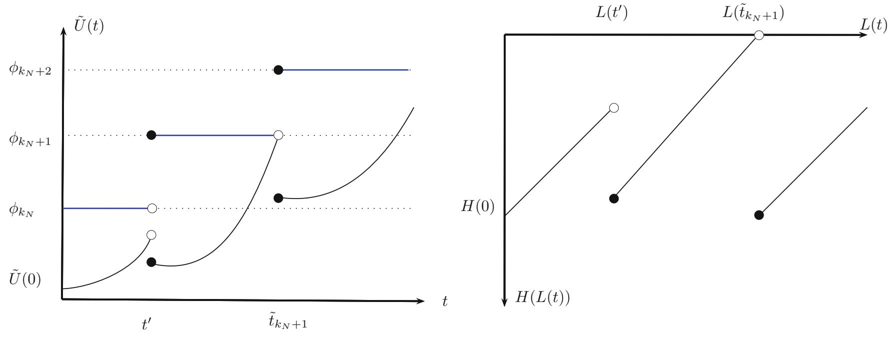

Fig. 1. A sample evolution of $\tilde{U}$, and the corresponding evolution of the transformed process H. A random activation occurred at time t', and a push occurred at time $\tilde{t}_{k_N+1}$.

a negative drift. More precisely, we have the following proposition. Here and for the remainder of this section, N will denote a Poisson process of unit intensity.

Proposition 6. *For any constant* $0 < c \leq 1/2$*, there is a* $c' = (1 + o(1))c$ *so that for any* $\sigma \leq L(\tilde{T}_{cn})$*,* $H'(s + \sigma) - H'(\sigma)$ *is stochastically dominated by* $s - 2(1 - c')N(s)$ *on the interval* $[0, L(\tilde{T}_{cn}) - \sigma]$*.*

Note in particular that $s - 2(1 - c)N(s)$ has negative drift for any $c < 1/2$.

Unfortunately, the bound we obtain on the drift gets worse as the number of active nodes increases. So some care is required; here we will sketch the argument. First, on the interval $[0, L(\tilde{T}_{n/4})]$, $H'(s)$ is dominated by $s - \frac{3}{2}N(s)$. This process starts from 0 and has negative drift (since $\mathbb{E}(s - \frac{3}{2}N(s)) = -s/2$). Proving that its expected maximum value is constant reduces to a straightforward concentration bound. Since $\tau \geq 2 \log \log n$, so that $k_{\mathrm{N}} = O(n/\log n)$, this negative drift also implies that the process will be very negative (below say $-n/64$) at time $L(\tilde{T}_{n/4})$, with very high probability. On the interval $[L(\tilde{T}_{n/4}), L(\tilde{T}_{n/2})]$, $H'(s)$ can be dominated by $s - N(s)$ conditioned on $N(L(\tilde{T}_{n/4})) = H'(L(\tilde{T}_{n/4})) \leq -n/64$. Again a concentration argument shows that with very high probability the process remains negative in this interval, and so the expected number of further pushes is again at most a constant.

Acknowledgments. We thank Albert Banchs, Antonio Fernández, Domenico Giustiniano, Nicole Immorlica, Julia Komjáthy, Brendan Lucier, and Yaron Singer for stimulating discussions and helpful pointers to the literature.

References

1. Andersson, H., Britton, T.: Stochastic Epidemic Models and Their Statistical Analysis. Lecture Notes in Statistics. Springer, New York (2000)
2. Asadpour, A., Nazerzadeh, H., Saberi, A.: Stochastic submodular maximization. In: Papadimitriou, C., Zhang, S. (eds.) WINE 2008. LNCS, vol. 5385, pp. 477–489. Springer, Heidelberg (2008)

3. Bailey, N.: A simple stochastic epidemic. Biometrika **37**, 193–202 (1950)
4. Barry, K.: Ford bets the fiesta on social networking, September 2009. www.wired.com/2009/04/how-the-fiesta/
5. Bartlett, M.: An Introduction to Stochastic Processes, with Special Reference to Methods and Applications. Cambridge University Press, Cambridge (1978)
6. Bollobás, B., Kohayakawa, Y.: On Richardson's model on the hypercube. In: Bollobás, B., Thomason, A. (eds.) Combinatorics, Geometry and Probability. Cambridge University Press, Cambridge (1997)
7. Borgs, C., Brautbar, M., Chayes, J., Lucier, B.: Maximizing social influence in nearly optimal time. In: SODA (2014)
8. Boyd, S., Arpita, G., Prabhakar, B., Shah, D.: Randomized gossip algorithms. IEEE Trans. Inf. Theory **52**(6), 2508–2530 (2006)
9. Chen, N.: On the approximability of influence in social networks. In: SODA (2008)
10. Chen, W., Wang, C., Wang, Y.: Scalable influence maximization for prevalent viral marketing in large-scale social networks. In: KDD (2010)
11. Chen, Y., Krause, A.: Near-optimal batch mode active learning and adaptive submodular optimization. In: ICML (2013)
12. Cisco: VNI: Global mobile data traffic forecast update, 2013–2018 (2014). www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.html. Accessed 28 October 2014
13. Doerr, B., Künnemann, M.: Tight analysis of randomized rumor spreading in complete graphs. In: ANALCO (2014)
14. Domingos, P., Richardson, M.: Mining the network value customers. In: KDD (2001)
15. Golovin, D., Krause, A.: Adaptive submodularity: theory and applications in active learning and stochastic optimization. J. Artif. Intell. Res. **42**, 427–486 (2011)
16. Grimmett, G.: Probability on Graphs: Random Processes on Graphs and Lattices. Cambridge University Press, Cambridge (2010). Institute of Mathematical Statistics Textbooks
17. Han, B., Hui, P., Kumar, A., Marathe, M., Pei, G., Srinivasan, A.: Cellular traffic offloading through opportunistic communications: a case study. In: CHANTS (2010)
18. Horel, T., Singer, Y.: Scalable methods for adaptively seeding a social network. In: WWW (2015)
19. Ioannidis, S., Chaintreau, A., Massoulie, L.: Optimal and scalable distribution of content updates over a mobile social network. In: IEEE INFOCOM (2009)
20. Kempe, D., Kleinberg, J., Tardos, E.: Maximizing the spread of influence through a social network. In: SIGKDD (2003)
21. Kempe, D., Kleinberg, J., Tardos, E.: Influential nodes in a diffusion model. In: ICALP (2005)
22. Kleinberg, J.: Cascading behavior in social and economic networks. In: ACM EC (2013)
23. Sciancalepore, V., Giustiniano, D., Banchs, A., Picu, A.: Offloading cellular traffic through opportunistic communications: analysis and optimization. ArXiv preprint 1405.3548 (2014)
24. Seeman, L., Singer, Y.: Adaptive seeding in social networks. In: FOCS (2013)
25. Whitbeck, J., Amorim, M., Lopez, Y., Leguay, J., Conan, V.: Relieving the wireless infrastructure: When opportunistic networks meet guaranteed delays. In: WOWMOM (2011)