MANETS: High mobility can make up for low transmission power [⋆]

Andrea E.F. Clementi[1], Francesco Pasquale[1⋆], and Riccardo Silvestri[2]

[1] Dipartimento di Matematica, Università di Roma "Tor Vergata"
{clementi,pasquale}@mat.uniroma2.it
[2] Dipartimento di Informatica, Università di Roma "La Sapienza" silvestri@di.uniroma1.it

Abstract. We consider a *Mobile Ad-hoc NETworks (MANET)* formed by n nodes that move independently at random over a finite square region of the plane. Nodes exchange data if they are at distance at most r within each other, where $r > 0$ is the *node transmission radius*. The *flooding time* is the number of time steps required to broadcast a message from a *source* node to every node of the network. Flooding time is an important measure of the speed of information spreading in dynamic networks.

We derive a nearly-tight upper bound on the *flooding time* which is a *decreasing* function of the maximal *velocity* of the nodes.

It turns out that, when the node velocity is "sufficiently" high, even if the node transmission radius r is far below the *connectivity threshold*, the flooding time does not asymptotically depend on r. This implies that flooding can be very fast even though every *snapshot* (i.e. the static random geometric graph at any fixed time) of the MANET is fully disconnected. Data reach all nodes quickly despite these ones use very low transmission power.

Our result is the first *analytical* evidence of the fact that high, random node mobility strongly speed-up information spreading and, at the same time, let *nodes save energy*.

Keywords: Mobile Ad-Hoc Networks, Evolving Graphs, Random Processes, Flooding.

1 Introduction

The impact of node mobility in data propagation is currently one of the major issues in Network Theory. The new trend is to consider node mobility as a *resource* for data forwarding rather than a *hurdle* [14, 10]. This is well-captured by the model known as *opportunistic Mobile Ad-Hoc NETworks* (*opportunistic MANET*), an interesting recent evolution of MANET [18, 11, 10, 22]. In opportunistic MANETS, mobile nodes are enabled to communicate even if a route connecting them never exists. Furthermore, nodes are not supposed to have or acquire any knowledge of the network topology (this one being highly-dynamic). Such two features make data communication in opportunistic networks a new challenging research topic from both foundational and practical view-points.

Inspired by opportunistic MANET, we here consider the *Evolving Graph* yielded by a set of n nodes moving over a finite square. Two nodes can exchange data, at a given time step, if their relative distance at that time step is not larger than a fixed *transmission radius* $r > 0$.

The aim of this work is to investigate the speed of data propagation in the above evolving graph. It is not hard to see that this study must cope with technical problems which are far to be trivial. Just to get an insight of such difficulties, we observe that classic *static* concepts like global *connectivity* and network *diameter* are almost meaningless in this context. On one hand, we can easily construct a sequence of node configurations such that *every* corresponding *snapshot* (i.e. the communication graph at any fixed time step) of the network is not connected while the broadcast task can be performed in a logarithmic number of time steps. On the other hand, it is easy to construct another temporal sequence of node configurations where the snapshot diameter is always

---

3 while the broadcast task requires $\Theta(n)$ time. Previous experimental works in this topic show that data communication can benefit from node mobility even though *all* the *snapshots* of the MANETs are not connected [11, 18, 21, 14, 6]. However, to our knowledge, the only *analytical* evidence of this phenomenon is that proved in [11] concerning *network capacity*, an information-theoretic concept which is not known to be related to the *speed* of data propagation.

We thus believe that, in order to investigate data propagation on evolving graphs - such as MANETS - a different concept and/or performance measure should be adopted and investigated.

In our opinion, a fundamental role in this dynamic world is played by the *Flooding Time*: this is in fact the desired crucial concept/measure. The flooding mechanism is the simple broadcast communication protocol where every *informed* node sends the source message at every time step (a node is said to be *informed* if it knows the source message). The flooding time of an evolving graph is the first time step in which all nodes are informed. The flooding time is a natural lower bound for any broadcast protocol and it represents the maximal speed of data propagation: the same role of the diameter in static networks. Flooding time of some classes of *Markovian Evolving Graphs* [3] has been recently studied in [7, 8].

Our work provides an analytical study of the flooding time of a natural and relevant class of MANETS: we prove a nearly-tight bound on the flooding time showing that high, random node mobility can *dramatically speed-up* data propagation with respect to the corresponding static model. This can be seen as a strong and somewhat final improvement of our previous work [8] where we (only) proved that (low) random node mobility does not *slow-down* data propagation.

Our MANET model: An informal definition. We consider a model of mobile networks called *geometric Markovian Evolving Graphs*, i.e., *geometric-MEG* [8]. It is the discrete version of the well-known *random-walk* model [5, 9, 12] and can also be viewed as the *walkers* model [9] on the square.

In *geometric-MEG*, nodes (i.e. radio stations) move over a finite square region of the plane and each node performs, independently from the others, a sort of *Brownian* motion. In our model we make time and space discrete (see Section 2 for details). The mobility parameter is the *move radius* ρ. At every time step, a node moves uniformly at random to any point which is within distance ρ from its current position; so ρ determines the *maximal node velocity*. At any time there is an edge (i.e. a bidirectional link) between two nodes if they are at distance at most r.

It turns out that a Geometric-MEG is a temporal sequence of random disk graphs (*random geometric graphs* [19]). When r is over the *connectivity threshold*, such random graphs are with high probability[3] (in short, *w.h.p.*) connected and have diameter $D < n$ that depends on n and r. In our previous work [8], we proved that when $\rho < r$ (so we are in the case of *low node mobility*) and the latter is over the connectivity threshold, then the flooding time of geometric-MEG is w.h.p. asymptotically equivalent[4] to the diameter of the corresponding snapshots. Under the assumption $\rho < r$, it is not hard to show [8] this is *the best the flooding can achieve*: the result is in fact asymptotically tight. It thus follows that, in the *slow* case, random node mobility does not significantly affect the flooding time with respect to the static case.

Our results. In this work, we consider the case $\rho > r$. This can viewed as a model for opportunistic MANETS where message transmission is very difficult due to critic environment conditions and/or poor node transmission power while node velocity is high and random. We are thus motivated by the futuristic scenario of mobile wireless sensor environments composed of myriads of tiny nodes dispersed in the environment and subject to high, unpredictable mobility as a consequence of environmental dynamics such as wind, storms or water streams. To our knowledge, the impact of such a high node mobility on the speed of data propagation has never been considered, at least from a foundational perspective.

---

[3] As usual, we say that an event $\mathcal{E}$ occurs *with high probability* if $\mathbf{P}(\mathcal{E}) \geqslant 1 - 1/n^{\Theta(1)}$.

[4] Actually, our previous bound leaves an $O(\log \log n)$ gap in a small range of the network parameters.

We provide a nearly-tight bound on the flooding time for such a case. Let G be a Geometric-MEG of n nodes over a square of edge size[5] $\sqrt{n}$, transmission radius r and move radius ρ such that $r \geq r_0$ and $\rho \geq c\sqrt{\log n}$, where r_0 and c are sufficiently large positive constants. Then, w.h.p., flooding in G is completed within

$$\mathcal{O}\left(\frac{\sqrt{n}}{\rho} + \log n\right) \quad \text{time steps.}$$

It is not hard to show that, for $\rho \geqslant r$, the expected flooding time is $\Omega(\sqrt{n}/\rho)$ [8], so our upper bound is nearly-tight and becomes tight whenever the flooding time is $\Omega(\log n)$.

When the transmission radius r is over the connectivity threshold (i.e. $\Theta(\sqrt{\log n})$), our bound implies that if the move radius ρ is (asymptotically) higher than r, the flooding time is (asymptotically) smaller than the diameter of the snapshots. When r is smaller than the connectivity threshold (say, it is a constant r_0) while $\rho \geqslant c\sqrt{\log n}$, then our bound implies that flooding can be efficiently completed despite the snapshot of every time step is formed by several connected components of small sizes [19, 13].

In general, our upper bound says that, in this case, *the flooding time does not asymptotically depends on the transmission radius.*

This fact has important technological consequences in the futuristic scenario of large, high-mobile MANETS. The two major goals in MANETS are: i) guarantee good and fast data communication, and ii) minimize node energy-consumption (which is clearly an increasing function of r). It is well-known that in classic, (static or low-mobile) MANETS such two goals are in contrast with each other and, thus, a suitable trade-off must be determined (actually, optimizing this trade-off is currently a major research issue in ad-hoc networking [2, 15, 20]). In particular, we know that [19, 13], in order to guarantee global connectivity (and thus data communication) in static random geometric graphs, the transmission radius must be $\Omega(\sqrt{\log n})$, so it must *increase with the network size*.

In this context, our bound is a strong mathematical evidence of the fact that, when node mobility is relatively high and random, the two above goals are not competing anymore. We can achieve fast data-forwarding by using small transmission radius (so, saving node energy). More importantly, the transmission radius can be an *absolute constant* and, so, it does not need to increase as the network size does. The technology of node transmitter devices can be thus *scalable*. Observe that node mobility in such opportunistic networks is due to the *host* mobility which is often fully independent from sensor devices: high sensor mobility does not (necessarily) imply high energy consumption [10, 18].

Adopted Techniques. The bound in [8] on the flooding time for the *slow* case is achieved thanks to the expanding properties of the *connected* snapshots of the geometric-MEG which are, in turn, guaranteed by two facts: 1) The stationary node distribution at every time step is almost uniform; and 2) the transmission radius r is over the connectivity threshold. In particular, they imply that, starting from the second time step, the number of informed nodes is large enough to apply standard Chernoff-like bounds. This allowed us to evaluate the number of *new* informed nodes at any successive step. The role of node mobility is thus shown to have a negligible impact on the flooding process.

This scenario is no longer true when r is below the connectivity threshold (say constant). The snapshots of the geometric MEG are very sparse and disconnected and, hence, their expanding properties are very scarce. This results into a relatively-long initial phase (called *Bootstrap*) of the flooding process where the number of informed nodes is not large enough to get useful concentration results from Chernoff-like bounds.

When $\rho >> r$, the flooding process is mainly due to *node mobility* that, roughly speaking, brings the source information *outside* the small connected components of the sparse snapshots: We provide a clean analytical description of this phenomenon. A key-ingredient here is a probabilistic analysis of the Bootstrap. We present a set of probabilistic lemmas for *almost-increasing random processes*

---

[5] For clarity's sake, we choose here to keep node density constant as the number of nodes grows.

that allows us to evaluate, at every time step, the number of new informed nodes even when the latter has a small expected value. The rather general form of such lemmas might result useful in other similar situations where Chernoff's-like bounds are useless.

2   The Node Mobility Model

We consider a model of dynamic graphs, introduced in [8], that is a discrete version of the *random walk mobility* model for radio networks [5]. In the latter model, nodes (i.e. radio stations) move on a bounded region of the plane (typically a square region) and each node performs, independently from the others, a sort of Brownian motion. At any time there is an edge (i.e. a bidirectional connection link) between two nodes if they are at distance at most r (r represents the *transmission radius*). In our model we discretize time and space. We choose to keep the density constant (i.e. the ratio between the number of nodes and the area) as the number n of nodes grows. The region in which nodes move is a square of side $\sqrt{n}$ and the density equals to 1. We remark that this choice is only for clarity's sake, and all our results can be scaled to any density $\delta(n)$. The nodes can assume positions whose coordinates are integer multiple of a resolution coefficient $\epsilon > 0$. Formally, nodes move on the following set of points

$$L_{n,\epsilon} = \left\{ (i\epsilon, j\epsilon) \mid i, j \in \mathbb{N} \wedge i, j \leqslant \frac{\sqrt{n}}{\epsilon} \right\}$$

At any time step, a node can move to one of the positions of $L_{n,\epsilon}$ within distance ρ from the previous position. The positive real number ρ is a fixed parameter that we call *move radius*. It can be interpreted as the maximum velocity of a node[6]. Formally, we introduce the *move graph* $M_{n,\rho,\epsilon} = (L_{n,\epsilon}, E_{n,\rho,\epsilon})$, where

$$E_{n,\rho,\epsilon} = \{(\mathbf{x}, \mathbf{y}) \mid \mathbf{x}, \mathbf{y} \in L_{n,\epsilon} \ \ d(\mathbf{x}, \mathbf{y}) \leqslant \rho\}$$

and $d(\cdot, \cdot)$ is the Euclidean distance. A node in position $\mathbf{x}$, in one time step, can move in any position in $\Gamma(\mathbf{x})$, where $\Gamma(\mathbf{x}) = \{\mathbf{y} \mid (\mathbf{x}, \mathbf{y}) \in E_{n,\rho,\epsilon}\}$. The nodes are identified by the first n positive integers $[n]$. The time-evolution of the movement of a single node i is represented by a Markov chain $\{P_{i,t} \ ; \ t \in \mathbb{N}\}$ where $P_{i,t}$ are random variables (in short *r.v.*) whose state-space is $L_{n,\epsilon}$ and

$$\mathbf{P}\left(P_{i,t+1} = \mathbf{x}\right) \quad = \quad \begin{cases} \frac{1}{|\Gamma(P_{i,t})|} & \text{if } \mathbf{x} \in \Gamma(P_{i,t}) \\ 0 & \text{otherwise} \end{cases}$$

In other words, $P_{i,t}$ is the position of node i at time t. Thus, the time evolution of the movements of all the nodes is represented by a Markov chain $\mathcal{P}(n, \rho, \epsilon) = \{P_t \ : \ t \in \mathbb{N}\}$ whose state space is $L_{n,\epsilon} \times L_{n,\epsilon} \times \cdots \times L_{n,\epsilon}$ (n times) and

$$P_t = (P_{1,t}, P_{2,t}, \ldots, P_{n,t})$$

Let us fix a *transmission radius* $r > 0$. A *geometric-MEG* is a sequence of random variables $\mathcal{G}(n, \rho, r, \epsilon) = \{G_t \ : \ t \in \mathbb{N}\}$ such that $G_t = ([n], E_t)$ with

$$E_t = \{(i, j) \mid d(P_{i,t}, P_{j,t}) \leqslant r\}$$

As for the stationary case, we observe that the stationary distribution π_i of Markov chain $\{P_{i,t} \ ; \ t \in \mathbb{N}\}$ is (see [1])

$$\pi_i(\mathbf{x}) \quad = \quad \frac{|\Gamma(\mathbf{x})|}{\sum_{\mathbf{y} \in L_{n,\epsilon}} |\Gamma(\mathbf{y})|}$$

Moreover, the stationary distribution of $\mathcal{P}(n, \rho, \epsilon)$ is the product of the independent distributions π_i for all $i \in [n]$. We say that a geometric-MEG $\mathcal{G}(n, \rho, r, \epsilon) = \{G_t \ : \ t \in \mathbb{N}\}$ is a *stationary*

---

[6] Indeed, a node can run through a distance of at most ρ in a unit of time.

geometric-MEG if the underlying P_0 is random with the stationary distribution of the Markov chain $\mathcal{P}(n, \rho, \epsilon) = \{P_t \; : \; t \in \mathbb{N}\}$. Notice that if $\mathcal{G}(n, \rho, r, \epsilon) = \{G_t \; : \; t \in \mathbb{N}\}$ is a stationary geometric-MEG then all r.v. G_t are random with the same probability distribution that we call *stationary distribution* of $\mathcal{G}(n, \rho, r, \epsilon)$.

In the rest of the paper, we will always assume that the move radius ρ is not larger than $\sqrt{n}$.

3 Bounding the Flooding Time

In the flooding mechanism, every node that is informed sends the source message at every time step: so, all nodes that are within distance r from an informed node will be informed at the next time step. For the sake of simplicity, every time step is divided into two consecutive actions: i) the *move action*, where nodes make their random move, and ii) the *transmission action*, where the informed nodes send the source message. Clearly this assumption does not affect the asymptotical bound on the flooding time.

Our result can be formally stated as follows.

Theorem 1. *Let $\mathcal{G}(n, \rho, r, \epsilon)$ be a stationary geometric-MEG. If $r \geqslant r_0$ and $\rho \geqslant c\sqrt{\log n}$ for sufficiently large constants r_0 and c, then the flooding time is w.h.p.*

$$\mathcal{O}\left(\frac{\sqrt{n}}{\rho} + \log n\right)$$

3.1 Proof's Overiew

The proof consists of a probabilistic analysis of the number of new informed nodes at every time step of the flooding process. In order to cope with this analysis, the temporal process is organized in three consecutive *phases*. Even though it is likely that in the real process these phases happen simultaneously rather than consecutively, our analysis yields the desired upper bound. The phases depend on the current number of informed nodes and on the "locality-degree" of the process. As for the latter, we need to partition the square into equal *supercells*, i.e., subsquares of side length $L = \Theta(\rho^2)$. This partition guarantees that any node v in a supercell S, after the move-action, can reach *any* position in *any* neighboring supercell with *almost-uniform* probability. Another crucial property yielded by the partition is that, for the first - say - $\mathcal{O}(n)$ time steps, every supercell will contain $\Theta(\rho^2)$ nodes, w.h.p.

The Bootstrap Phase. In this initial phase, we start our analysis focussing on what happens inside the neighborhood of the supercell S_0 containing the source, i.e., the supercell set $N(S_0)$ formed by S_0 and its adjacent supercells. We can say that, with positive-constant probability, S_0 contains $\Theta(r^2)$ informed nodes after the first time step. Observe that this is the crucial analysis point where we need to go from *positive-constant probability* to *high probability* and we cannot use Chernoff-like bounds. Indeed, in the successive time steps $t > 0$ of this phase, we consider the flooding-rate inside the supercell S_t' having the maximal number of informed nodes at time step t. We will then prove that, after $t = \mathcal{O}(\log n)$ time steps, w.h.p., there will be (at least) one supercell *quasi-informed*, i.e., it will have $\Theta(\rho^2)$ informed nodes.

The Spreading Phase. After the Bootstrap, we can thus assume (w.h.p.) that there is (at least) one supercell quasi-informed. We can thus look at the flooding from a quasi-informed supercell to its adjacent ones. We show that, w.h.p., if a supercell is quasi-informed at a given time step, then all its adjacent supercells will be quasi-informed within the next time step. Since we prove that the *boundary* of any supercell set D has size at least $\Omega(\sqrt{|D|})$, it turns out that this flooding phase makes all the supercells quasi-informed within $\mathcal{O}(\sqrt{n}/\rho)$ time steps.

The Filling Phase. At the end of the previous phase, we thus have w.h.p. all supercells quasi-informed. The Filling phase consists of the sequence of time steps required to get all supercells informed. We prove that, w.h.p., this final process can be completed in $O(\log n)$ time steps.

3.2 Preliminaries

We need to introduce the following notions.

- The square is partitioned into squared *supercells* of side length L with

$$\frac{\rho}{3\sqrt{2}} \leqslant L \leqslant \frac{\rho}{2\sqrt{2}}$$

- Every supercell is partitioned into squared *cells* of side length ℓ with

$$\frac{r}{1+\sqrt{2}} \leqslant \ell \leqslant \frac{r}{\sqrt{2}}$$

- The neighborhood $N(S)$ of a supercell S is the set of supercells formed by S and all its adjacent supercells.
- A supercell is *quasi-informed* at time t if it contains $\gamma\rho^2$ informed nodes at that time, where γ is a suitable positive constant that will be specified later.

We say that the *density condition* holds at time t if, for every supercell S, the number of nodes in S at time t is at least $\eta\rho^2$, for a suitable constant $0 < \eta < 1$. Let $\mathcal{D}$ be the following event: the density condition holds for every time step $t = 0, 1, \ldots, n$.

The proof of the following lemma is omitted since it is an easy consequence of the almost uniformity of the stationary distribution of geometric-MEG.

Lemma 1. *Let $\mathcal{G}(n, \rho, r, \epsilon)$ be a stationary geometric-MEG. If $\rho \geqslant c\sqrt{\log n}$ for a sufficiently large constant c, then the probability of event $\mathcal{D}$ is at least $1 - 1/n^4$.*

In the rest of this section, we will tacitly assume that event $\mathcal{D}$ holds. Thanks to the previous lemma, since we are conditioning w.r.t. an event that holds w.h.p., the corresponding unconditional probabilities are affected by only a negligible factor.

For the sake of simplicity, we will use the following probability notations. For an event $\mathcal{E}$ and a r.v. X, the notation

$$\mathbf{P}\left(\mathcal{E} \mid X\right) \leqslant p$$

means that, for every possible value x of X, it holds

$$\mathbf{P}\left(\mathcal{E} \mid X = x\right) \leqslant p$$

3.3 The bootstrap

We now provide an upper bound on the time required to get at least one supercell quasi-informed. We will prove the bound for $r = r_0$ where r_0 is a sufficiently large constant. Observe that, since the flooding time is a non-increasing function of the transmission radius r, the same upper bound holds for any $r \geqslant r_0$ as well.

The following lemma will be used to evaluate the number of new informed nodes after an initial sequence of consecutive time steps. Notice that the r.v. X_ts can be mutually dependent.

Lemma 2 (Almost-increasing random processes). *Let $\{X_t : t \in \mathbb{N}\}$ be a sequence of random variables with $X_0 = 1$. Assume that two real values $\alpha > 1$, $0 < \beta < 1$, a positive integer M, and a probability $p \in (0, 1)$ exist such that for every $t \in \mathbb{N}$ it holds that*

$$\mathbf{P}\left(X_{t+1} < \alpha\, X_t \mid X_t < M,\, X_{t-1}, \ldots, X_1\right) \leqslant p \tag{1}$$

$$\mathbf{P}\left(X_{t+1} \geqslant \beta\, X_t \mid X_t,\, X_{t-1}, \ldots, X_1\right) = 1 \tag{2}$$

If $p < \frac{\log \alpha}{e \log(\alpha/\beta)}$ then for any $t \geqslant \frac{\log M}{\log \alpha - e\, p \log(\alpha/\beta)}$ it holds that

$$\mathbf{P}\left(\bigcap_{i=1}^{t}\{X_i < M\}\right) \leqslant \exp\left(-pt\right)$$

Proof. For every t define the binary r.v. Y_t whose value is 1 if $X_t < \alpha X_{t-1} \ \wedge \ X_{t-1} < M$ and 0 otherwise. Let us say that a time step t is *bad* if $Y_t = 1$ and *good* if $Y_t = 0$. By hypothesis, we have

$$\mathbf{P}\left(Y_t = 1 \mid X_{t-1} < M, X_{t-2}, \ldots, X_1\right) \leqslant p$$

Consider the first t time steps and observe that if there are τ bad time steps and $t - \tau$ good time steps, then either $X_t \geqslant \alpha^{t-\tau}\beta^\tau$ or a time step $i \leqslant t$ exists such that $X_i \geqslant M$. So if $X_i < M$ for every $i = 1, \ldots, t$, then the number of bad time steps must be at least τ_t where

$$\tau_t = \frac{t \log \alpha - \log M}{\log(\alpha/\beta)}$$

Hence, if $t \geqslant \frac{\log M}{\log \alpha - e\, p \log(\alpha/\beta)}$ it holds that

$$\mathbf{P}\left(\bigcap_{i=1}^{t}\{X_i < M\}\right) \leqslant \mathbf{P}\left(\sum_{i=1}^{t} Y_i \geqslant \tau_t\right) \leqslant \mathbf{P}\left(B(t,p) \geqslant \tau_t\right) \leqslant \exp\left(-pt\right)$$

where in the second inequality we used Lemma 15, in the last inequality Observation 14, and the fact that the hypothesis on p implies that $\tau_t \geqslant e\,p\,t$. □

The following lemmas allow us to apply the previous lemma to the flooding process.

We say that a cell C is *infected* at time t if, immediately after the move action of time step t, C contains at least one informed node. The next lemma will be used in the Filling phase as well.

Lemma 3. *Positive constants a and b exist such that, for any time step t and for any supercell S, if at the beginning of time step t a supercell $S' \in N(S)$ contains m informed nodes, then*

$$\mathbf{P}\left(Z \leqslant am'\right) \ \leqslant \ \exp\left(-bm'\right)$$

where Z is the random variable counting the number of infected cells of S at time t and $m' = \min\{m, L^2/\ell^2\}$.

Proof. Let S' be a supercell in $N(S)$, let I be the set of informed nodes in S', and set $m = |I|$. For any cell C of S, let Z_C be the 0-1 random variable that has value 1 iff immediately after the move action of time step t at least one of the nodes in I is in C. Let $Z' = \sum_C Z_C$. Clearly, it holds that $Z \geqslant Z'$. Since the Z_C's are not independent, we will use the method of bounded differences (Lemma 16) in order to prove that Z' is concentrated around its expected value. Firstly, observe that

$$\mathbf{P}\left(Z_C = 0\right) \ \leqslant \ \left(1 - \frac{\ell^2}{\pi\rho^2}\right)^m.$$

For the sake of convenience let $\lambda = \left(1 - \frac{\ell^2}{\pi\rho^2}\right)$. It holds that

$$\mathbf{E}\left[Z'\right] \ = \ \sum_C \mathbf{P}\left(Z_C = 1\right) \ \geqslant \ \frac{L^2}{\ell^2}(1 - \lambda^m).$$

Let indicate the nodes in I by the integers $1, 2, \ldots, m$. For any $i \in I$, let X_i be the random variable whose value is the position of node i immediately after the move action (at time step t). Let $f : \prod_{i=1}^{m} L_{n,\epsilon} \to \mathbb{R}$ be the function such that $f(\mathbf{x}_1, \ldots, \mathbf{x}_m)$ is the number of cells C of S such that, for some i, position $\mathbf{x}_i$ is in C. Clearly, Z' is equal to the random variable $f(X_1, \ldots, X_m)$. Consider any integer k with $1 \leqslant k \leqslant m$ and any positions $\mathbf{x}_1, \ldots, \mathbf{x}_{k-1}, \mathbf{x}_k, \mathbf{x}'_k \in L_{n,\epsilon}$. Let $\hat{X}_1, \ldots, \hat{X}_m$ be the random variables such that for $i = 1, \ldots, k$ $\hat{X}_i = \mathbf{x}_i$ and for $j = k + 1, \ldots, m$ $\hat{X}_j = X_j$. Moreover, define random variables Y and Y' as follows

$$Y \ = \ f(\hat{X}_1, \ldots, \hat{X}_m) \quad \text{and} \quad Y' \ = \ f(\hat{X}_1, \ldots, \hat{X}_{k-1}, \mathbf{x}'_k, \hat{X}_{k+1}, \ldots, \hat{X}_m).$$

It is immediate to see that

$$\mathbf{E}\left[Y\right] = \mathbf{E}\left[f(X_1\dots,X_m)\mid(X_1,\dots,X_{k-1})=(\mathbf{x}_1,\dots,\mathbf{x}_{k-1}),X_k=\mathbf{x}_k\right]$$
$$\mathbf{E}\left[Y'\right] = \mathbf{E}\left[f(X_1\dots,X_m)\mid(X_1,\dots,X_{k-1})=(\mathbf{x}_1,\dots,\mathbf{x}_{k-1}),X_k=\mathbf{x}'_k\right]$$

Thus, we have to bound $|\mathbf{E}\left[Y\right]-\mathbf{E}\left[Y'\right]|$. Let ζ and ζ' be the number of cells of S that contain some position in $\{\mathbf{x}_1,\dots,\mathbf{x}_{k-1},\mathbf{x}_k\}$ and $\{\mathbf{x}_1,\dots,\mathbf{x}_{k-1},\mathbf{x}'_k\}$, respectively. It is easy to see that if $\zeta=\zeta'$ then $\mathbf{E}\left[Y\right]=\mathbf{E}\left[Y'\right]$ and so $|\mathbf{E}\left[Y\right]-\mathbf{E}\left[Y'\right]|=0$. If, instead, $\zeta\neq\zeta'$ then it must be the case that $|\zeta-\zeta'|=1$. Without loss of generality, suppose that $\zeta=\zeta'+1$. Consider now the analogue of variables Z_C for the variables $\hat{X}_1,\dots,\hat{X}_m$ and $\hat{X}_1,\dots,\hat{X}_{k-1},\mathbf{x}'_k,\hat{X}_{k+1},\dots,\hat{X}_m$. Call them $\hat{Z}_C$ and $\hat{Z}'_C$, respectively. Clearly, $\mathbf{E}\left[Y\right]=\sum_C\hat{Z}_C$ and $\mathbf{E}\left[Y'\right]=\sum_C\hat{Z}'_C$. Since $\zeta=\zeta'+1$, there exists a cell $\hat{C}$ of S such that $\hat{C}$ does not contain any position in $\{\mathbf{x}_1,\dots,\mathbf{x}_{k-1},\mathbf{x}'_k\}$, it contains $\mathbf{x}_k$, and for every other cell $C\neq\hat{C}$ of S $\hat{Z}_C=\hat{Z}'_C$. This implies that

$$\mathbf{E}\left[Y\right]-\mathbf{E}\left[Y'\right] = \mathbf{P}\left(\hat{Z}_{\hat{C}}=1\right)-\mathbf{P}\left(\hat{Z}'_{\hat{C}}=1\right) = 1-\mathbf{P}\left(\hat{Z}'_{\hat{C}}=1\right).$$

Moreover, it holds that

$$1-\mathbf{P}\left(\hat{Z}'_{\hat{C}}=1\right) = \mathbf{P}\left(\hat{Z}'_{\hat{C}}=0\right) \leqslant \lambda^{m-k}.$$

We have thus proved that, in any case, it holds that

$$|\mathbf{E}\left[f(X_1\dots,X_m)\mid(X_1,\dots,X_{k-1})=(\mathbf{x}_1,\dots,\mathbf{x}_{k-1}),X_k=\mathbf{x}_k\right]$$
$$-\mathbf{E}\left[f(X_1\dots,X_m)\mid(X_1,\dots,X_{k-1})=(\mathbf{x}_1,\dots,\mathbf{x}_{k-1}),X_k=\mathbf{x}'_k\right]| \leqslant \lambda^{m-k}$$

By applying Lemma 16 with $\delta=\frac{\mathbf{E}[Z']}{2}$, we obtain that

$$\mathbf{P}\left(Z'\leqslant\frac{\mathbf{E}\left[Z'\right]}{2}\right) \leqslant \mathbf{P}\left(|Z'-\mathbf{E}\left[Z'\right]|\geqslant\frac{\mathbf{E}\left[Z'\right]}{2}\right) \leqslant 2\exp\left(-\frac{\mathbf{E}\left[Z'\right]^2}{2\sum_{k=1}^m(\lambda^{m-k})^2}\right).$$

Since $\mathbf{E}\left[Z'\right]\geqslant\frac{L^2}{\ell^2}(1-\lambda^m)$, by simple calcutations, we get

$$\frac{\mathbf{E}\left[Z'\right]^2}{2\sum_{k=1}^m(\lambda^{m-k})^2} \geqslant \frac{\rho^2(1-\lambda^m)}{4072\cdot\ell^2}$$

Moreover, since $m\geqslant m'$ and $m'\leqslant\frac{L^2}{\ell^2}$, it holds that

$$(1-\lambda^m) \geqslant \frac{\ell^2}{2\pi\rho^2}m'$$

This implies that

$$\frac{\mathbf{E}\left[Z'\right]^2}{2\sum_{k=1}^m(\lambda^{m-k})^2} \geqslant \frac{m'}{25586} \quad\text{and}\quad \frac{\mathbf{E}\left[Z'\right]^2}{2} \geqslant \frac{m'}{227}$$

Finally, we have that

$$\mathbf{P}\left(Z\leqslant am'\right) \leqslant \mathbf{P}\left(Z\leqslant am'\right) \leqslant \exp\left(-bm'\right)$$

where $a=\frac{1}{227}$ and $b=\frac{1}{25586}$.

$\square$

We can now fix the constant γ defining quasi-informed supercells: $\gamma=\frac{a\eta}{227}$. For any supercell S, let $m_t(S)$ be the number of informed in S at time step t.

For any time step t let $Y_t=\max\{m_t(S):S\text{ is a supercell}\}$.

Lemma 4. *For any time step t, it holds that*

$$\mathbf{P}\left(Y_{t+1} < 2Y_t \mid Y_t < \frac{r_0^2}{28}\right) \leqslant \exp\left(-\frac{r_0^2}{224}\right)$$

Proof. Consider an informed node, and let C be the cell containing that node immediately after the move action at time t. The expected number of nodes in C is $\mathbf{E}[X] = \ell^2 \geqslant \frac{r_0^2}{28}$ and by using Chernoff bounds it holds that

$$\mathbf{P}\left(X \leqslant \frac{r_0^2}{56}\right) \leqslant \exp\left(-\frac{r_0^2}{224}\right)$$

Clearly all the nodes in cell C will be informed after the transmission action, hence

$$\mathbf{P}\left(Y_{t+1} \leqslant 2Y_t \mid Y_t < \frac{r_0^2}{2 \cdot 56}\right) \leqslant \mathbf{P}\left(Y_{t+1} \leqslant \frac{r_0^2}{56}\right)$$

$$\leqslant \mathbf{P}\left(X \leqslant \frac{r_0^2}{56}\right) \leqslant \exp\left(-\frac{r_0^2}{224}\right)$$

$\square$

Lemma 5. *For any supercell S and for any time step t, it holds that*

$$\mathbf{P}\left(m_{t+1}(S) < 2m_t(S) \ \Big|\ \frac{r_0^2}{28} < m_t(S) < \gamma\rho^2\right) \leqslant 3\exp\left(-\frac{br_0^2}{56}\right)$$

Proof. For sake of convenience, define $\mathcal{H}_t$ to be the event

$$\mathcal{H}_t = \left\{\frac{r_0^2}{28} < m_t(S) < \gamma\rho^2\right\}$$

For every cell C in supercell S let $Z_C = 1$ if at least one of the $m_t(S)$ informed nodes will be in cell C immediately after the move action of time t and $Z_C = 0$ otherwise. Let $Z = \sum_{C \in S} Z_C$ be the number of cells in S *infected* by the $m_t(S)$ nodes. And let

$$A = \ell^2 Z \geqslant \frac{r_0^2}{(1+\sqrt{2})^2} Z$$

be the size of the *infected area* in S after the move action.

Now consider a supercell S' adjacent to S. For every node u in S', at the beginning of time t, let $X_u = 1$ if u will be in the infected area immediately after the move action and $X_u = 0$ otherwise, and let $X = \sum_{u \in S'} X_u$. Observe that all such nodes will be informed after the transmission action, hence $m_{t+1}(S) \geqslant X$. In what follows we will evaluate the probability that X is less than 2 times $m_t(S)$. Since, by assumption, event $\mathcal{D}$ holds, we need to evaluate the following

$$\mathbf{P}\left(X < 2m_t(S) \mid \mathcal{H}_t, \mathcal{D}\right)$$

We first write down the above probability in a more suitable way

$$\mathbf{P}\left(X < 2m_t(S) \mid \mathcal{H}_t, \mathcal{D}\right) =$$

$$= \sum_{\frac{r_0^2}{28} < m < \gamma\rho^2} \mathbf{P}\left(X < 2m \mid m_t(S) = m, \mathcal{D}\right)\mathbf{P}\left(m_t(S) = m \mid \mathcal{H}_t, \mathcal{D}\right)$$

If we prove that for every m between $r_0^2/28$ and $\gamma\rho^2$ it holds that

$$\mathbf{P}\left(X < 2m \mid m_t(S) = m, \mathcal{D}\right) \leqslant 3\exp\left(-\frac{r_0^2}{56}\right)$$

then the thesis follows. Let us then relate the above probability to the size of the infected area[7]

$$\mathbf{P}\left(X < 2m \mid m_t(S) = m, \mathcal{D}\right) \leqslant \mathbf{P}\left(X < 2m \mid m_t(S) = m, \mathcal{D}, A \geqslant \alpha m\right) +$$
$$+ \mathbf{P}\left(A < \alpha m \mid m_t(S) = m, \mathcal{D}\right) \qquad (3)$$

where α is a suitable constant that we will choose later. Now we evaluate the first term of the above sum. For every node u in S' it holds that

$$\mathbf{P}\left(X_u = 1 \mid m_t(S) = m, \mathcal{D}, A \geqslant \alpha m\right) \geqslant \frac{\alpha}{\pi \rho^2} m \qquad (4)$$

Hence the conditional expectation of X is

$$\mathbf{E}\left[X \mid m_t(S) = m, \mathcal{D}, A \geqslant \alpha m\right] \geqslant \sum_u \frac{\alpha}{\pi \rho^2} m \geqslant \eta \rho^2 \frac{\alpha}{\pi \rho^2} m = \frac{\alpha \eta}{\pi} m \qquad (5)$$

If the X_us were independent, by using Chernoff bound we could get

$$\mathbf{P}\left(X < \frac{\alpha \eta}{2\pi} m \mid m_t(S) = m, \mathcal{D}, A \geqslant \alpha m\right) \leqslant \exp\left(-\frac{\alpha \eta}{8\pi} m\right)$$

Unfortunately, since event $\mathcal{D}$ makes assumption *on the future*, r.v. X_us conditioned on $\mathcal{D}$ are not independent anymore. Anyway, this is easy to handle. Split the event $\mathcal{D} = \mathcal{D}_1 \cap \mathcal{D}_2$ where $\mathcal{D}_1$ is the event "The density condition holds since time step t" and $\mathcal{D}_2$ is the event "The density condition holds from time step $t+1$ to time step n". Since $\mathcal{D}$ holds with high probability, then events $\mathcal{D}_1 \setminus \mathcal{D}$ and $\mathcal{D}_2 \setminus \mathcal{D}$ have negligible probabilities, hence conditioning on $\mathcal{D}$ is *almost the same* as conditioning on $\mathcal{D}_1$. Formally, we can observe that for every event $\mathcal{A}$ it holds that $\mathbf{P}\left(\mathcal{A} \mid \mathcal{D}\right) \leqslant 2\mathbf{P}\left(\mathcal{A} \mid \mathcal{D}_1\right)$. By conditioning on $\mathcal{D}_1$ instead of $\mathcal{D}$, bounds (4) and (5) remains unchanged, and now we can apply Chernoff bound because the X_us conditioned on $\mathcal{D}_1$ are independent

$$\mathbf{P}\left(X < \frac{\alpha \eta}{2\pi} m \mid m_t(S) = m, \mathcal{D}_1, A \geqslant \alpha m\right) \leqslant \exp\left(-\frac{\alpha \eta}{8\pi} m\right) \leqslant \exp\left(-\frac{\alpha \eta}{224\pi} r_0^2\right)$$

And

$$\mathbf{P}\left(X < \frac{\alpha \eta}{2\pi} m \mid m_t(S) = m, \mathcal{D}, A \geqslant \alpha m\right) \leqslant 2\exp\left(-\frac{\alpha \eta}{224\pi} r_0^2\right)$$

By choosing $\alpha = \frac{4\pi}{\eta}$, we have that

$$\mathbf{P}\left(X < 2m \mid m_t(S) = m, \mathcal{D}, A \geqslant \alpha m\right) \leqslant 2\exp\left(-\frac{r_0^2}{56}\right) \qquad (6)$$

As for the second addend of (3), i.e. $\mathbf{P}\left(A < \alpha m \mid m_t(S) = m, \mathcal{D}\right)$, two cases may arise.
<u>Case 1:</u> $r_0^2/28 < m \leqslant L^2/\ell^2$.
From Lemma 3, it holds that

$$\mathbf{P}\left(Z \leqslant am \mid m_t(S) = m, \mathcal{D}\right) \leqslant \exp\left(-bm\right) \leqslant \exp\left(-b\frac{r_0^2}{28}\right)$$

where a and b are the constants fixed by Lemma 3. Hence

$$\mathbf{P}\left(A < \alpha m \mid m_t(S) = m, \mathcal{D}\right) = \mathbf{P}\left(Z < \frac{\alpha}{\ell^2} m \mid m_t(S) = m, \mathcal{D}\right)$$
$$\leqslant \mathbf{P}\left(Z < \frac{\alpha(1+\sqrt{2})^2}{r_0^2} m \mid m_t(S) = m, \mathcal{D}\right)$$

---

[7] Here and in the sequel we will repeatedly use that for three events $\mathcal{A}, \mathcal{B},$ and $\mathcal{C}$ we can write $\mathbf{P}\left(\mathcal{A} \mid \mathcal{C}\right) \leqslant \mathbf{P}\left(\mathcal{A} \mid \mathcal{B}, \mathcal{C}\right) + \mathbf{P}\left(\overline{\mathcal{B}} \mid \mathcal{C}\right)$.

Thus, for $r_0 \geqslant (1 + \sqrt{2})\sqrt{\alpha/a}$, it holds that

$$\mathbf{P}\left(Z < \frac{\alpha(1 + \sqrt{2})^2}{r_0^2} m \mid m_t(S) = m, \mathcal{D}\right) \leqslant \mathbf{P}\left(Z \leqslant am \mid m_t(S) = m, \mathcal{D}\right) \leqslant \exp\left(-b\frac{r_0^2}{28}\right)$$

<u>Case 2:</u> $L^2/\ell^2 < m < \gamma\rho^2$.
From Lemma 3, it holds that

$$\mathbf{P}\left(Z \leqslant aL^2/\ell^2 \mid m_t(S) = m, \mathcal{D}\right) \leqslant \exp\left(-b\frac{L^2}{\ell^2}\right)$$

where a and b are the constants fixed by Lemma 3. Hence

$$\begin{aligned}
\mathbf{P}\left(A < \alpha m \mid m_t(S) = m, \mathcal{D}\right) &\leqslant \mathbf{P}\left(A < \alpha\gamma\rho^2 \mid m_t(S) = m, \mathcal{D}\right) \\
&\leqslant \mathbf{P}\left(A < 18\alpha\gamma L^2 \mid m_t(S) = m, \mathcal{D}\right) \\
&= \mathbf{P}\left(Z < 18\alpha\gamma\frac{L^2}{\ell^2} \mid m_t(S) = m, \mathcal{D}\right)
\end{aligned}$$

Since $\gamma \leqslant \frac{a}{18\alpha}$, it holds that

$$\mathbf{P}\left(Z < 18\alpha\gamma\frac{L^2}{\ell^2} \mid m_t(S) = m, \mathcal{D}\right) \leqslant \exp\left(-b\frac{L^2}{\ell^2}\right)$$

In both cases, since $L^2/\ell^2 \geqslant r_0^2/28$, we get

$$\mathbf{P}\left(A < \alpha m \mid m_t(S) = m, \mathcal{D}\right) \leqslant \exp\left(-b\frac{r_0^2}{28}\right) \tag{7}$$

By using (6) and (7) in (3), we obtain

$$\mathbf{P}\left(X < 2m \mid m_t(S) = m, \mathcal{D}\right) \leqslant 3\exp\left(-b\frac{r_0^2}{56}\right)$$

for every $r_0^2 < m < \gamma\rho^2$. $\qquad\square$

Lemma 6. *Within $\mathcal{O}(\log n)$ time steps there is a quasi-informed supercell w.h.p.*

Proof. By using simple geometric arguments, if we choose $\beta = 1/121$, then it holds that $Y_{t+1} \geqslant \beta Y_t$ with probability one, thus satisfying Hypothesis (2) of Lemma 2. From Lemmas 4 and 5, it is easy to see that the r.v. Y_ts satisfy Hypothesis (1) of Lemma 2 with constants $\alpha = 2$, $M = \gamma\rho^2$ and $p = 3\exp\left(-b\frac{r_0^2}{224}\right)$. Hence, for a sufficiently large constant r_0, the thesis follows by applying Lemma 2. $\qquad\square$

3.4 The spreading phase

Lemma 7 (Local-supercell spreading). *For any supercell S, if $m_t(S) \geqslant \gamma\rho^2$ then the event $m_{t+1}(S') \geqslant \gamma\rho^2$ holds for any supercell $S' \in N(S)$ with probability at least $1 - 1/n^4$.*

Proof. Let S' be any supercell in $N(S)$. Since $m_t(S) \geqslant \gamma\rho^2 \geqslant \rho^2/r_0^2$ (remind that r_0 is a constant that we can set sufficiently large), from Lemma 3, it follows that

$$\mathbf{P}\left(Z \leqslant a\frac{\rho^2}{r_0^2}\right) \leqslant \exp\left(-b\frac{\rho^2}{r_0^2}\right)$$

where Z is the random variable counting the number of infected cells of S' at time t.

Let $A = \ell^2 Z \geqslant \frac{r_0^2}{7} Z$ be the size of the *infected area*. From the above inequality, it holds that

$$\mathbf{P}\left(A \leqslant a\frac{\rho^2}{7}\right) \leqslant \exp\left(-b\frac{\rho^2}{r_0^2}\right) \tag{8}$$

Now consider a super-cell S'' adjacent to S' and different from S. We want to evaluate how many nodes from S'' move to the infected area. For every node u in S'' let X_u be the binary r.v. whose value is 1 iff node u, after the move action, is in the infected area of S'. Since we are assuming the density event $\mathcal{D}$ holds, the r.v. X_u's are not independent. We thus first bound the relative probabilities under the weaker assumption that the density condition holds till time step $t-1$: This makes r.v. X_u's independent. Then, we bring this bound under the density condition $\mathcal{D}$.

It is easy to see that

$$\mathbf{P}\left(X_u = 1 \;\middle|\; A \geqslant a\frac{\rho^2}{7}\right) \geqslant a\frac{1}{7\pi}$$

Since the density condition holds before the move action of time t, we get

$$\mathbf{E}\left[X \;\middle|\; A \geqslant a\frac{\rho^2}{7}\right] \geqslant a\frac{\eta\rho^2}{7\pi}$$

where X is the r.v. counting the number of nodes of S'' moving to the infected area of S'. Observe that X is a lower bound on the number of informed node in cell S'' after the transmission action of time step t. By applying Chernoff's bound, we obtain

$$\mathbf{P}\left(X \leqslant a\frac{\eta\rho^2}{14\pi} \;\middle|\; A \geqslant a\frac{\rho^2}{7}\right) \leqslant \exp\left(-a\frac{\eta\rho^2}{56\pi}\right)$$

Then, from (8), we get

$$\mathbf{P}\left(X \leqslant a\frac{\eta\rho^2}{14\pi}\right) \leqslant \mathbf{P}\left(X \leqslant a\frac{\eta\rho^2}{14\pi} \;\middle|\; A \geqslant a\frac{\rho^2}{7}\right) + \mathbf{P}\left(A \leqslant a\frac{\rho^2}{7}\right)$$

$$\leqslant \exp\left(-a\frac{\eta\rho^2}{56\pi}\right) + \exp\left(-b\frac{\rho^2}{r_0^2}\right) \leqslant \frac{1}{n^5}$$

for a suitable choice of constant c such that $\rho \geqslant c\sqrt{\log n}$. Since $\gamma \leqslant a\frac{\eta}{14\pi}$, we obtain

$$\mathbf{P}\left(X \geqslant \gamma\rho^2 \mid \mathcal{D}\right) \geqslant \left(1 - \frac{1}{n^5}\right) - 2\mathbf{P}\left(\overline{\mathcal{D}}\right) \geqslant 1 - \frac{1}{n^4}$$

$\square$

In order to prove a bound on the number of time steps that are sufficient to guarantee (with high probability) that one quasi-informed supercell spreads the information to all the supercells, we need two lemmas. The first one will provide a bound on the number of supercells that are adjacent to an arbitrary set of supercells.

Let Q be a $m \times m$ square grid, that is, Q is a square partitioned into $m \times m$ congruent sub-squares, called cells. For any subset B of the cells of Q, define the *boundary* ∂B of B as the set of all the cells that do not belong to B and that are adjacent to some cell in B:

$$\partial B = \{c \mid c \notin B \ \wedge \ \exists c' \in B : \ c' \text{ is adjacent to } c\}.$$

Lemma 8 (Boundary size). *Let Q be a $m \times m$ square grid and let B be any subset of the cells of Q. It holds that*

$$|\partial B| \geqslant \sqrt{\min\{|B|, m^2 - |B|\}}.$$

Proof. For the sake of convenience, we say that a cell in B is a black cell and all the cells not in B are white cells. We say that a row of the grid Q is black if all the cells of the row are black. Similarly, we define a black column. Moreover, a row or a column which contains both at least one black cell and at least one white cell is said to be gray. Let b_r and b_c be, respectively, the number of black rows and the number of black columns. To prove the bound on $|\partial B|$ we distinguish four cases.

$b_r = 0 \wedge b_c \geqslant 1$: In this case, every row is gray. This implies that every row contains at least one cell in ∂B (if the leftmost cell of the row is black then the leftmost white cell belongs to ∂B otherwise the cell immediately to the left of the leftmost black cell belongs to ∂B). Since the rows are m, it follows that $|\partial B| \geqslant m$. It immediately derives that $|\partial B| \geqslant \sqrt{\min\{|B|, m^2 - |B|\}}$, since $\sqrt{\min\{|B|, m^2 - |B|\}} \leqslant \sqrt{m^2} = m$.

$b_r \geqslant 1 \wedge b_c = 0$: This case is symmetric to the previous one.

$b_r \geqslant 1 \wedge b_c \geqslant 1$: Without loss of generality, assume that $b_r \leqslant b_c$. Since $b_c \geqslant 1$, there are $m - b_r$ gray rows. Thus, $|\partial B| \geqslant m - b_r$. It is easy to see that the number of cells belonging to either black rows or black columns is $m \cdot b_r + m \cdot b_c - b_r \cdot b_c$. Since these are black cells, it holds that

$$|B| \;\geqslant\; m \cdot b_r + m \cdot b_c - b_r \cdot b_c \;=\; m \cdot b_r + b_c(m - b_r) \;\geqslant\; m \cdot b_r + b_r(m - b_r).$$

Now, it easy to verify that the inequality $|B| \geqslant m \cdot b_r + b_r(m - b_r)$ implies that

$$b_r \;\leqslant\; m - \sqrt{m^2 - |B|}.$$

Hence, it holds that $|\partial B| \geqslant m - b_r \geqslant m - (m - \sqrt{m^2 - |B|}) = \sqrt{m^2 - |B|}$, and thus $|\partial B| \geqslant \sqrt{\min\{|B|, m^2 - |B|\}}$.

$b_r = 0 \wedge b_c = 0$: Let y_r and y_c be, respectively, the number of gray rows and the number of gray columns. Since there are neither black rows nor black columns, it must be the case that every black cell belongs to both a gray row and a gray column. This implies that

$$y_r \cdot y_c \;\geqslant\; |B|.$$

Without loss of generality, assume that $y_r \geqslant y_c$. It follows that $y_r^2 \geqslant |B|$, and so $y_r \geqslant \sqrt{|B|}$. Since every gray row contains at least a cell in ∂B, it holds that $|\partial B| \geqslant \sqrt{|B|} \geqslant \sqrt{\min\{|B|, m^2 - |B|\}}$.
$\qquad\square$

The second lemma will allow us to prove an upper bound on the number of steps to get all the supercells quasi-informed when, in one time step, the information propagates from all the quasi-informed supercells to their adjacent ones.

Lemma 9 (Spreading time I). *Let K be any integer with $K \geqslant 1$ and let $\{q_t \mid t \in \mathbb{N}\}$ be a succession of integers such that $q_0 \geqslant 1$, for every $t \geqslant 0$, $q_t \leqslant K$ and $q_{t+1} \geqslant q_t + \sqrt{\min\{q_t, K - q_t\}}$. Then, it holds that, for every $t \geqslant 5\sqrt{K}$, $q_t = K$.*

Proof. If $K \leqslant 2$ it is immediate to see that the thesis holds, so in the sequel of the proof we assume that $K \geqslant 3$. Firstly observe that if $q_t \leqslant K - 1$ then $q_{t+1} \geqslant q_t + 1$. This implies that an integer $\tau \geqslant 0$ exists such that $\tau = \min\{t \mid q_t \geqslant K/2\}$. Now we prove, by induction on t, that, for every t with $0 \leqslant t \leqslant \tau$, $q_t \geqslant \frac{t^2}{8}$. For $t = 0, 1, 2$ it is trivially true. Assume that it is true for $t \geqslant 2$ (and $t + 1 \leqslant \tau$). Then, since $q_t < K/2$, it holds that

$$q_{t+1} \;\geqslant\; q_t + \sqrt{q_t} \;\geqslant\; \frac{t^2}{8} + \sqrt{\frac{t^2}{8}} \;=\; \frac{t^2 + \sqrt{8}t}{8} \;\geqslant\; \frac{(t+1)^2}{8}$$

and this completes the proof by induction. From this it follows that if $\tau \geqslant 1$ then $q_{\tau-1} \geqslant \frac{(\tau-1)^2}{8}$. Thus, it is easy to verify that $\tau \leqslant 2\sqrt{K} + 1$. Now, we prove, by induction on t, that

$$\forall t \;:\; \tau \leqslant t \leqslant \left\lceil 5\sqrt{K} \right\rceil \qquad q_t \;\geqslant\; K - \left\lceil \frac{\left(\left\lceil 5\sqrt{K}\right\rceil - t\right)^2}{9} \right\rceil.$$

For $t = \tau$ it holds that

$$K - \left\lceil \frac{\left(\left\lceil 5\sqrt{K} \right\rceil - \tau \right)^2}{9} \right\rceil \leqslant K - \frac{\left(5\sqrt{K} - 2\sqrt{K} - 1 \right)^2}{9} \qquad \text{(since } \tau \leqslant 2\sqrt{K} + 1\text{)}$$

$$= \frac{9K - \left(3\sqrt{K} - 1 \right)^2}{9} = \frac{2\sqrt{K}}{3} - \frac{1}{9} \leqslant \frac{K}{2} \leqslant q_\tau.$$

Assume now that it is true for $t \geqslant \tau$ (and $t + 1 \leqslant \lceil 5\sqrt{K} \rceil$). This implies that an integer $\delta \geqslant 0$ exists such that

$$q_t = K - \left\lceil \frac{h^2}{9} \right\rceil + \delta$$

where we set $h = \lceil 5\sqrt{K} \rceil - t$. Since $q_t \geqslant K/2$, it holds that

$$q_{t+1} \geqslant q_t + \sqrt{K - q_t} = K - \left\lceil \frac{h^2}{9} \right\rceil + \delta + \sqrt{\left\lceil \frac{h^2}{9} \right\rceil - \delta}.$$

From the inequality $\forall a \geqslant b \geqslant 0 \ \sqrt{a - b} \geqslant \sqrt{a} - \sqrt{b}$, it derives that

$$\delta + \sqrt{\left\lceil \frac{h^2}{9} \right\rceil - \delta} \geqslant \delta + \sqrt{\left\lceil \frac{h^2}{9} \right\rceil} - \sqrt{\delta} \geqslant \sqrt{\left\lceil \frac{h^2}{9} \right\rceil}$$

where the last inequality (i.e. $\delta \geqslant \sqrt{\delta}$) holds since δ is an integer. Thus,

$$q_{t+1} \geqslant K - \left\lceil \frac{h^2}{9} \right\rceil + \sqrt{\left\lceil \frac{h^2}{9} \right\rceil}.$$

Now, we prove that, for every integer $j \geqslant 1$, the following inequality holds

$$\sqrt{\left\lceil \frac{j^2}{9} \right\rceil} \geqslant \left\lceil \frac{j^2}{9} \right\rceil - \left\lceil \frac{(j-1)^2}{9} \right\rceil.$$

For $j = 1, 2, \ldots, 7$, it can be exhaustively proved case by case. Let $j \geqslant 8$. It holds that

$$\sqrt{\left\lceil \frac{j^2}{9} \right\rceil} \geqslant \sqrt{\frac{j^2}{9}} \geqslant \frac{j}{3} \geqslant \frac{8 + 2j}{9}$$

where the last inequality is due to the assumption that $j \geqslant 8$. Moreover,

$$\left\lceil \frac{j^2}{9} \right\rceil - \left\lceil \frac{(j-1)^2}{9} \right\rceil \leqslant \frac{j^2}{9} + 1 - \frac{(j-1)^2}{9} = \frac{8 + 2j}{9}.$$

Thus the inequality is proved. By using it with $j = h$, we obtain that

$$q_{t+1} \geqslant K - \left\lceil \frac{h^2}{9} \right\rceil + \sqrt{\left\lceil \frac{h^2}{9} \right\rceil} \geqslant K - \left\lceil \frac{(h-1)^2}{9} \right\rceil = K - \left\lceil \frac{\left(\left\lceil 5\sqrt{K} \right\rceil - (t+1) \right)^2}{9} \right\rceil$$

and this completes the proof by induction. Thus, we have proved that for $\hat{t} = \lceil 5\sqrt{K} \rceil$, it holds that

$$q_{\hat{t}} \geqslant K - \left\lceil \frac{\left(\left\lceil 5\sqrt{K} \right\rceil - \hat{t} \right)^2}{9} \right\rceil = K.$$

Hence, for every $t \geqslant \lceil 5\sqrt{K} \rceil$, it holds that $q_t = K$. $\qquad\square$

By combining Lemmas 7, 8 and 9 we get the following bound.

Lemma 10 (Spreading time II). *If at time $t_1 \leqslant n/2$, there is at least one quasi-informed supercell then, with probability at least $1 - \frac{1}{n^2}$, at every time t with $t_1 + 22\frac{\sqrt{n}}{\rho} \leqslant t \leqslant n$ all the supercells are quasi-informed.*

Proof. In the sequel, when we write time t we mean time $t_1 + t$. For any $t \geqslant 0$, let $\mathcal{Q}_t$ be the set of quasi-informed supercells at time t. By hypothesis $|\mathcal{Q}_0| \geqslant 1$. In virtue of Lemma 8, if all the supercells in $\mathcal{Q}_t$ and all their adjacent supercells get quasi-informed, at time $t+1$, then

$$|\mathcal{Q}_{t+1}| \geqslant |\mathcal{Q}_t| + \sqrt{\min\{|\mathcal{Q}_t|, N - |\mathcal{Q}_t|\}}$$

where $N = \frac{n}{L^2}$ is the number of supercells. This implies that if the above inequality does not hold then there exists a supercell $S \in \mathcal{Q}_t$ such that either S or one adjacent supercell of S is not quasi-informed at time $t+1$. It follows that

$$\mathbf{P}\left(|\mathcal{Q}_{t+1}| < |\mathcal{Q}_t| + \sqrt{\min\{|\mathcal{Q}_t|, N - |\mathcal{Q}_t|\}}\right) \leqslant \mathbf{P}\left(\exists S \in \mathcal{Q}_t : \mathcal{E}_{S,t+1}\right)$$

where $\mathcal{E}_{S,t+1}$ is the event that occurs if S or one adjacent supercell of S is not quasi-informed at time $t+1$. By the union bound, it holds that

$$\mathbf{P}\left(\exists S \in \mathcal{Q}_t : \mathcal{E}_{S,t+1}\right) \leqslant \sum_S \mathbf{P}\left(S \in \mathcal{Q}_t \wedge \mathcal{E}_{S,t+1}\right) = \sum_S \mathbf{P}\left(\mathcal{E}_{S,t+1} \mid S \in \mathcal{Q}_t\right)\mathbf{P}\left(S \in \mathcal{Q}_t\right).$$

From Lemma 7, for every supercell S,

$$\mathbf{P}\left(\mathcal{E}_{S,t+1} \mid S \in \mathcal{Q}_t\right) \leqslant \frac{1}{n^4}.$$

It follows that

$$\mathbf{P}\left(|\mathcal{Q}_{t+1}| < |\mathcal{Q}_t| + \sqrt{\min\{|\mathcal{Q}_t|, N - |\mathcal{Q}_t|\}}\right) \leqslant \sum_S \mathbf{P}\left(\mathcal{E}_{S,t+1} \mid S \in \mathcal{Q}_t\right)\mathbf{P}\left(S \in \mathcal{Q}_t\right)$$

$$\leqslant \sum_S \mathbf{P}\left(\mathcal{E}_{S,t+1} \mid S \in \mathcal{Q}_t\right)$$

$$\leqslant \frac{N}{n^4} \leqslant \frac{1}{n^3}.$$

Thus, by the union bound, with probability at least $1 - \frac{1}{n^2}$, it holds that

$$\forall t = 0, 1, \ldots, n/2 \qquad |\mathcal{Q}_{t+1}| \geqslant |\mathcal{Q}_t| + \sqrt{\min\{|\mathcal{Q}_t|, N - |\mathcal{Q}_t|\}}.$$

By applying Lemma 9 with $q_t = |\mathcal{Q}_t|$ and $K = N$, we obtain that, for every t with $5\sqrt{N} \leqslant t \leqslant n/2$, $|\mathcal{Q}_t| = N$. The thesis follows by taking into account that $t_1 \leqslant n/2$, t means time $t_1 + t$, and $5\sqrt{N} \leqslant 5\frac{\sqrt{n}}{3\sqrt{2}\rho} \leqslant 22\frac{\sqrt{n}}{\rho}$. $\qquad\square$

3.5   The filling phase

We first prove that a node not yet informed and belonging to a quasi-informed supercell will get informed in one time step, with a constant probability.

Lemma 11. *There exists a constant $\beta > 0$ such that, for any node u, if at the beginning of a time step t the supercell that contains u is quasi-informed and node u is not informed then, with probability at least β, node u gets informed by the end of time step t.*

Proof. Assume that at the beginning of time step t node u is not informed and the supercell S that contains u is quasi-informed. Thus S contains at least $\gamma\rho^2 \geqslant \frac{\rho^2}{r_0^2}$ informed nodes at the beginning of time step t. From Lemma 3, it derives that

$$\mathbf{P}\left(Z \leqslant \frac{a\rho^2}{r_0^2}\right) \leqslant \exp\left(-\frac{b\rho^2}{r_0^2}\right)$$

where Z is the random variable counting the number of infected cells of S at time t. Now, let $\mathcal{U}$ be the event that occurs if node u gets informed by the end of time step t. It is immediate to see that, for any $k \geqslant 0$,

$$\mathbf{P}\left(\mathcal{U} \mid Z = k\right) \geqslant \frac{k\ell^2}{\pi\rho^2}.$$

It follows that

$$\begin{aligned}
\mathbf{P}\left(\mathcal{U}\right) &= \sum_{k \geqslant 0} \mathbf{P}\left(\mathcal{U} \mid Z = k\right)\mathbf{P}\left(Z = k\right) \\
&\geqslant \sum_{k \geqslant \frac{a\rho^2}{r_0^2}} \mathbf{P}\left(\mathcal{U} \mid Z = k\right)\mathbf{P}\left(Z = k\right) \\
&\geqslant \frac{a}{6\pi} \sum_{k \geqslant \frac{a\rho^2}{r_0^2}} \mathbf{P}\left(Z = k\right) \\
&= \frac{a}{6\pi}\mathbf{P}\left(Z \geqslant \frac{a\rho^2}{r_0^2}\right) \\
&\geqslant \frac{a}{6\pi}\left(1 - \exp\left(-\frac{b\rho^2}{r_0^2}\right)\right) \\
&\geqslant \beta
\end{aligned}$$

for a suitable constant $\beta > 0$. $\qquad\square$

From above lemma we can derive a logarithmic upper bound for the filling time.

Lemma 12. *If a time step $t_2 \leqslant \frac{3n}{4}$ exists such that at every time step t with $t_2 \leqslant t \leqslant n$ all the supercells are quasi-informed, then by time $t_2 + \mathcal{O}(\log n)$ all the nodes are informed, w.h.p.*

Proof. In the sequel, when we write time t we mean time $t_2 + t$. Let u be any node. For any time t, let $\mathcal{U}_t$ be the event that occurs if node u is not informed by the end of time step t. Let β be the constant of Lemma 11 and let k be a constant such that $\beta^{k \log n} < \frac{1}{n^2}$. By hypothesis, for every $t = 0, 1, \ldots k \log n$, all the supercells are quasi-informed. This implies that at each of these time steps the supercell that contains u is quasi-informed. Thus, from Lemma 11, it holds that

$$\mathbf{P}\left(\bigwedge_{t=0}^{k \log n} \mathcal{U}_t\right) = \prod_{t=0}^{k \log n} \mathbf{P}\left(\mathcal{U}_t \,\middle|\, \bigwedge_{j=0}^{t-1} \mathcal{U}_j\right) \leqslant \beta^{k \log n} < \frac{1}{n^2}.$$

This means that a node u is not informed within $k \log n$ time steps with probability at most $\frac{1}{n^2}$. Hence, a direct application of the union bound shows that, with probability at least $1 - \frac{1}{n}$, all the nodes are informed within $k \log n$ steps. $\qquad\square$

Finally, Theorem 1 follows from Lemma 6, Lemma 10, and Lemma 12.

4 Conclusions

Some interesting issues concerning the flooding time on geometric-MEG are still open. There is a logarithmic gap between our upper bound and the known lower bound [8] when the move radius ρ is very large. Closing this gap is an open problem.

A more challenging open issue is to extend our upper bound in the case where ρ and r are both very small (i.e. below $\sqrt{\log n}$).

Another open research work is to study the flooding time when the starting distribution is not the stationary one but is arbitrary (i.e. a worst-case analysis). We conjecture that the worst-case flooding time is not asymptotically larger than the stationary one for a large range of the network parameters.

Observe that our upper bound can be easily extended to the gossiping task (i.e. the all-to-all communication). It would be interesting to extend our analysis to other basic communication tasks such as data-gathering and routing.

Finally, a major challenge is to obtain results similar to our upper bounds for more realistic mobility models [17].

References

1. D. Aldous and J. Fill. *Reversible Markov Chains and Random Walks on Graphs*. http://stat-www.berkeley.edu/users/aldous/RWG/book.html, 2002.
2. C. Ambühl. An optimal bound for the MST algorithm to compute energy efficient broadcast trees in wireless networks. In *Proc. of 32th International Colloquium on Automata, Languages and Programming (ICALP)*, volume 3580 of *LNCS*, pages 1139–1150. Springer, 2005.
3. C. Avin, M. Koucky, and Z. Lotker. How to explore a fast-changing world. In *Proc. of 35th International Colloquium on Automata, Languages and Programming (ICALP'08)*, volume 5125 of *LNCS*, pages 121–132. Springer, 2008.
4. Y. Azar, A.Z. Broder, A.R. Karlin, and E. Upfal. Balanced allocations. *SIAM Journal on Computing*, 29(1):180–200, 1999.
5. T. Camp, J. Boleng, and V. Davies. A survey of mobility models for ad hoc network research. *Wireless Communication and Mobile Computing*, 2(5):483–502, 2002.
6. I. Chatzigiannakis, A. Kinalis, S. E. Nikoletseas, and J. D. P. Rolim. Fast and energy efficient sensor data collection by multiple mobile sinks. In *Proc. of MOBIWAC'07*, pages 25–32, 2007.
7. A. Clementi, C. Macci, A. Monti, F. Pasquale, and R. Silvestri. Flooding time in edge-markovian dynamic graphs. In *Proc. of 27th Annual ACM SIGACT-SIGOPS Symposium on Principles of Distributed Computing (PODC'08)*, pages 213–222. ACM Press, 2008.
8. A. Clementi, A. Monti, F. Pasquale, and R. Silvestri. Information spreading in stationary markovian evolving graphs. In *Proc. of the 23rd IEEE International Parallel and Distributed Processing Symposium*. IEEE Computer Society, 2009.
9. J. Diaz, D. Mitsche, and X. Perez-Gimenez. On the connectivity of dynamic random geometric graphs. In *Proc. of 19th annual ACM-SIAM symposium on Discrete algorithms (SODA'08)*, pages 601–610, 2008.
10. S. Jain et Al. Exploiting mobility for energy efficient data collection in wireless sensor networks. *ACM/Kluwer Mobile Networks and Applications (MONET)*, 11(3), 2006.
11. M. Grossglauser and N.C. Tse. Mobility increases the capacity of ad-hoc wireless networks. *IEEE/ACM Trans. on Networking*, 10(4), 2002.
12. R.A. Guerin. Channel occupancy time distribution in a cellular radio system. *IEEE Trans. on Veichular Technology*, 36(3):89–99, 1987.
13. P. Gupta and P.R. Kumar. Critical power for asymptotic connectivity in wireless networks. *Stochastic Analysis, Control, Optimization and Applications*, pages 547–566, 1998.
14. A. Kinalis and S. E. Nikoletseas. Adaptive redundancy for data propagation exploiting dynamic sensory mobility. In *Proc. of ACM MSWIM'08*, pages 149–156, 2008.
15. L. GC02M. Kirousis, E. Kranakis, D. Krizanc, and A. Pelc. Power consumption in packet radio networks. *Theoretical Computer Science*, 243:289–305, 2000.
16. C. McDiarmid. On the method of bounded differences. In *(J. Siemons ed.), London Mathematical Society Lecture Note*, 141, pages 148–188. Cambridge University Press, 1989.
17. A. Mei and J. Stefa. Swim: a simple model to generate small mobile worlds. In *Proc. of IEEE INFOCOM'09*, 2009.
18. L. Pelusi, A. Passarella, and M. Conti. Beyond manets: Dissertation on opportunistic networking. *IIT-CNR Tech. Rep.*, 2006.
19. M. Penrose. *Random Geometric Graphs*. Oxford University Press, 2003.

20. P. Santi and D. M. Blough. The critical transmitting range for connectivity in sparse wireless ad hoc networks. *IEEE Transactions on Mobile Computing*, 2(1):25–39, 2003.
21. Z. Zhang. Routing in intermittently connected mobile ad-hoc networks and delay tolerant networks: overview and challenges. *IEEE Communication Surveys*, 8(1), 2006.
22. W. Zhao, M. Ammar, and E. Zegura. A message ferrying approach for data delivery in sparse mobile ad-hoc networks. In *Proc. of 5th ACM MobiHoc'04*, 2004.

A Basic Probability

Lemma 13 (Chernoff's bound). *Let be $X = \sum_{i=1}^{n} X_i$ where $X_1, \ldots, X_n$ are independent Bernoulli random variables and let be $0 < \varepsilon < 1$. If $0 < \mu \leqslant \mathbf{E}[X]$, then it holds*

$$\mathbf{P}\{X \leqslant (1-\varepsilon)\mu\} \leqslant e^{-\frac{\varepsilon^2}{2}\mu}$$

Observation 14 *Let X be a binomial random variable with $\mathbf{E}[X] = \mu$. Then*

$$\mathbf{P}(X \geqslant e\,\mu) \leqslant e^{-\mu}$$

We use the following standard probability bound (See [4]).

Lemma 15. *Let $X_1, \ldots, X_n$ be a sequence of random variables with values in an arbitrary domain, and let $Y_1, \ldots, Y_n$ be a sequence of binary random variables, with the property that $Y_i = Y_i(X_1, \ldots, X_i)$. If*

$$\mathbf{P}(Y_i = 1 \mid X_1, \ldots, X_{i-1}) \leqslant p$$

then

$$\mathbf{P}\left(\sum Y_i \geqslant k\right) \leqslant \mathbf{P}(B(n,p) \geqslant k)$$

where $B(n,p)$ denotes the binomially distributed random variable with parameters n and p.

The following lemma states one inequality of the method of bounded differences (see Corollary 6.10 in [16]).

Lemma 16. *Let $X_1, \ldots, X_n$ be random variables, with X_k taking values in a set A_k for each k, and let $\underline{X}$ denote the vector $(X_1, \ldots, X_n)$. Let $f : \prod_{k=1}^{n} A_k \to \mathbb{R}$ be an appropriately measurable function. Suppose that there are constants $c_1, \ldots, c_n$ so that*

$$\begin{aligned} |\mathbf{E}[f(\underline{X}) \mid (X_1, \ldots, X_{k-1}) = (x_1, \ldots, x_{k-1}), X_k = x_k] \\ - \mathbf{E}[f(\underline{X}) \mid (X_1, \ldots, X_{k-1}) = (x_1, \ldots, x_{k-1}), X_k = x_k']| \leqslant c_k \end{aligned}$$

for each $k = 1, \ldots, n$ and $x_i \in A_i$ ($i = 1, \ldots, k-1$) and $x_k, x_k' \in A_k$. Then for any $\delta > 0$,

$$\mathbf{P}(|f(\underline{X}) - \mathbf{E}[f(\underline{X})]| \geqslant \delta) \leqslant 2\exp\left(-\frac{2\delta^2}{\sum_{k=1}^{n} c_k^2}\right).$$