Leader Election in Shared Spectrum Radio Networks

Sebastian Daum[1]
sebastian.daum@usi.ch

Seth Gilbert[2]
seth.gilbert@comp.nus.edu.sg

Fabian Kuhn[1]
fabian.kuhn@usi.ch

Calvin Newport[3]
cnewport@cs.georgetown.edu

[1]Faculty of Informatics, University of Lugano, 6904 Lugano, Switzerland
[2]Department of Computer Science, National University of Singapore, Singapore
[3]Department of Computer Science, Georgetown University, Washington D.C., USA

Abstract

In this paper, we study the *leader election problem* in the context of a congested single-hop radio network. We assume a collection of N synchronous devices with access to a shared band of the radio spectrum, divided into $\mathcal{F}$ frequencies. To model unpredictable congestion, we assume an abstract *interference adversary* that can choose up to $t < \mathcal{F}$ frequencies in each round to disrupt, preventing communication. The devices are individually activated in arbitrary rounds by an adversary. On activation, a device does not know how many other devices (if any) are also active. The goal of the leader election problem is for each active device to output the id of a leader as soon as possible after activation, while preserving the safety constraint that all devices output the *same* leader, with high probability.

We begin by establishing a lower bound of $\Omega\big(\frac{\log^2 N}{(\mathcal{F}-t)\log\log N} + \frac{\mathcal{F}t}{\mathcal{F}-t}\cdot\log N\big)$ rounds, through reduction to an existing result in this model [6]. We then set out to prove this bound tight (within $\log\log N$ factors). For the case where $t = 0$, we present a novel randomized algorithm, based on a strategy of recruiting *herald nodes*, that works in $O\big(\frac{\log^2 N}{\mathcal{F}} + \log N\big)$ time. For $1 \leq t \leq \mathcal{F}/6$, we present a variant of our herald algorithm in which multiple real (potentially disrupted) frequencies are used to simulate each non-disrupted frequency from the $t = 0$ case. This algorithm works in $O\big(\frac{\log^2 N}{\mathcal{F}} + t\log N\big)$ time. Finally, for $t > \mathcal{F}/6$ we show how to improve the *trapdoor protocol* of [6], used to solve a similar problem in a non-optimal manner, to solve leader election in optimal $O\big(\frac{\log N + \mathcal{F}t}{\mathcal{F}-t}\cdot\log N\big)$ time, for (only) these large values of t. We also observe that if $\mathcal{F} = \omega(1)$, $t = o(\log N)$ and $t \leq (1-\epsilon)\mathcal{F}$ for a constant $\epsilon > 0$, our protocols beat the classic $\Omega(\log^2 N)$ bound on wake-up in a single frequency radio network, underscoring the observation that more frequencies in a radio network allows for more algorithmic efficiency—even if devices can each only participate on a single frequency at a time, and a significant fraction of these frequencies are disrupted adversarially.

1 Introduction & Related Work

Due to the growing number of wireless devices, the systems community has placed new emphasis on developing shared spectrum networks [19]. Such networks allow multiple unrelated protocols to share the same band of the radio spectrum, each dynamically adjusting its use based on the local behavior it observes. Many of the most widely-deployed wireless standards—including WiFi [1], Bluetooth [4], and Zigbee [2]—operate in shared spectrum networks, and with the recent opening of broadcast television bands for secondary use by networked devices, more such standards are sure to follow [3]. Shared spectrum networks, however, introduce new algorithmic challenges. A protocol operating in this environment encounters a communication medium that is being used concurrently in a dynamic and unpredictable fashion. Even tasks as basic as finding other nearby devices become complex in this unpredictable setting [22].

In this paper, we study the foundational problem of *leader election* in the shared spectrum setting. We argue that discovering nearby devices and then electing a leader to coordinate their behavior (e.g., by disseminating a frequency hopping pattern or spread spectrum code [20]) is a key building block in the construction of efficient protocols in these complex networks. In addition, techniques to solve leader election in single-hop radio networks have also proved useful for solving more basic problems like computing maximal independent sets for clustering or colorings to coordinate channel access in a multi-hop setting [16, 18]. Formally, we capture the dynamics of shared spectrum communication using the well-studied t*-disrupted* radio network model [5–9, 12, 17, 20, 21]. We present the first known optimal solution to leader election in this setting (within $\log \log$ factors). As detailed below, this optimal solution also beats a classic lower bound on communication in a non-shared, single-frequency radio network—providing further evidence of the surprising computational power of multiple frequency network models (see also [9]).

Leader Election Results. We study leader election in the t-disrupted network model, which describes a congested single-hop synchronous radio network consisting of $\mathcal{F} > 0$ communication frequencies. In each round, each active device can choose a single frequency on which to participate. Concurrent broadcasts on the same frequency lead to message loss on that frequency, due to collision. We assume no collision detection. To capture the unpredictable interference caused by unrelated protocols using the same shared spectrum, we introduce an abstract *interference adversary* that can choose up to $t < \mathcal{F}$ frequencies in each round to *disrupt*—preventing communication. We assume t is a known upper bound. The leader election problem assumes N devices that are activated in an arbitrary pattern by an adversary. On being activated, a device has no *a priori* knowledge of which other devices (if any) are also active. Some devices might never be activated. The goal is to output the id of a leader as soon as possible, while maintaining the safety property that all active devices output the same id, with high probability. The time complexity of a leader election algorithm is measured as the maximum number of rounds from when a device is activated to when it outputs a leader id.

We start by establishing a lower bound of $\Omega\big(\frac{\log^2 N}{(\mathcal{F}-t)\log\log N} + \frac{\mathcal{F}t}{\mathcal{F}-t} \cdot \log N\big)$ rounds, through reduction to our previous bound on the *wireless synchronization problem* in this same model [6]. Notice, the *trapdoor protocol* presented in [6] can be adapted to solve leader election in $O\big(\frac{\mathcal{F}}{\mathcal{F}-t}\log^2 N + \frac{\mathcal{F}t}{\mathcal{F}-t}\log N\big)$ time, which is not tight with respect to this lower bound. In the remainder of the paper, we set out to close this gap.

To accomplish this goal, we present three different algorithms, each optimal with respect to a different subset of the range of possible t values. For the case where $t = 0$, we present a randomized algorithm that works in $O\big(\frac{\log^2 N}{\mathcal{F}} + \log N\big)$ time. This algorithm uses a novel strategy of recruiting *herald* nodes to advance a leadership case on behalf of a potential leader. To understand the intuition behind this strategy, imagine that $\mathcal{F} = \log N$. Our algorithm, in this case, assigns an exponential distribution of probabilities to the channels, and active devices choose their channels according to this distribution; they then broadcast on the selected channel with constant probability. At a high-level, we can show that with constant probability, there will be

a favored channel in this round that has only a single broadcaster. The receivers on this favored channel are then responsible for *heralding* this single broadcaster by announcing it on a special announcement channel, also with constant probability. We can prove that $O(\log N)$ rounds will be sufficient to ensure that with high probability, a single such announcement is promulgated to the whole active network, and subsequently to all devices that are eventually activated.

For $1 < t \leq \mathcal{F}/6$, we present a variant of our herald algorithm in which multiple real (potentially disrupted) frequencies are used to simulate each undisrupted frequency from the $t = 0$ case. This algorithm still works in $O\big(\frac{\log^2 N}{\mathcal{F}} + t \log N\big)$ time, despite the extra channels needed for simulation. Finally, for $t > \mathcal{F}/6$ we improve the trapdoor protocol of [6] to work in $O\big(\frac{\log N + \mathcal{F}t}{\mathcal{F}-t} \cdot \log N\big)$ time for these values of t. This improvement matches the lower bound (within $\log \log$ factors). It is important to note that it requires t to be large. That is, for smaller t, this particular algorithm is no longer optimal—whereas our above herald-style algorithms are.

The most relevant existing result is the trapdoor protocol [6] mentioned above. This protocol can be adapted to solve the leader election problem in our model. As the time complexity is not tight with the lower bound, our protocols close this gap. In fact, the original problem solved by the trapdoor protocol is "wireless synchronization," and our protocols can be adapted to solve this problem as well. Thus, not only do we present the first known tight leader election bounds in this model, we also present the first known tight wireless synchronization bounds. Also relevant is the work of Meier et al. [15], which studies bounds on device discovery in the t-disrupted model. They focus primarily on the case where there are only 2 devices that start simultaneously, but t is unknown—showing unknown t algorithms that can come within a $\log^2 \mathcal{F}$ competitive ratio of the optimal known t algorithms, in terms of expected performance. We focus instead on a case where there are an unknown number of nodes started adversarially, but t is known.

The Computational Power of Frequency Diversity. Notice, for $t = 0$ and $\mathcal{F} = \omega(1)$, our herald protocol beats the classic $\Omega(\log^2 n)$ bound on a single device broadcasting alone in a single frequency, non-disrupted radio model [11,14] (i.e., our model with $\mathcal{F} = 1$ and $t = 0$). Indeed, our algorithms show that this advantage can be maintained even for relatively large amounts of disruption (i.e., for $t \leq (1 - \epsilon)\mathcal{F}$, $t = o(\log N)$ and $\mathcal{F} = \omega(1)$). Put another way, the presence of multiple communication frequencies adds non-trivial power to a radio network model, even if devices can each only use one frequency per round *and* a significant fraction of frequencies are disrupted. (A similar result was presented in [9], which proved that global broadcast can be solved faster in the t-disrupted radio network model than in the classical undisrupted, single-frequency model.)

2 Model & Definitions

We model randomized distributed algorithms in a synchronous single hop radio networking consisting of multiple communication channels and bounded disruption. In more detail, we assume time is divided into synchronized slots, called *rounds*. We assume that N devices—which we call *nodes*—begin each execution *inactive*. At the beginning of each round, an adversary decides which devices (if any) to make *active*, at which point they start executing with a round number of 1 (that is, we assume no *a priori* knowledge of a global round number). The radio network consists of $\mathcal{F} \geq 1$ disjoint and distinguishable narrowband communication frequencies — throughout the paper we use the notion of frequency and channel interchangeably.

In each round, each active node can choose a single frequency on which to participate by either broadcasting or receiving. If a single node broadcasts on a given frequency in a given round, then all nodes receiving on that frequency receive its message. If two or more nodes broadcast on the same frequency in the same round, then the message is lost due to collision. Nodes are incapable of *collision detection*, i.e.,

nodes can not distinguish between a collision and no message being broadcasted. We emphasize that nodes learn nothing about the behavior on other frequencies.

The multi-frequency open spectrum networks we model here are prone to disruption generated by unrelated protocols and other source of electromagnetic emission. We capture this unpredictable disruption with an *interference adversary* that can disrupt up to $\mathcal{F}$ frequencies per round. By disrupting a frequency, the adversary prevents any node from receiving a message on that frequency. That is, a node receives a message on a frequency only if a single node broadcasts on that frequency *and* it is not disrupted. We assume the interference adversary knows the algorithm being executed by the nodes, and the entire history prior to the current round, but does not know the private randomness used by nodes to make random choices.

Each node knows the $\mathcal{F}$ frequencies, an upper bound to the number of nodes, N, and an upper bound t to the number of disrupted frequencies.

2.1 Mathematical Preliminaries

We frequently need to say that an event A happens with probability close to 1. If the probability that A does not occur is exponentially small in some parameter k, i.e., if $\mathbb{P}(A) = 1 - e^{-ck}$ for some constant $c > 0$, we say that A happens with very high probability w.r.t. k, abbreviated as w.v.h.p.(k). We say that an event happens with high probability w.r.t. a parameter k, abbreviated as w.h.p.(k), if it happens with probability $1 - k^{-c}$, where the constant $c > 0$ can be chosen arbitrarily (possibly at the cost of adapting some other involved constants). If an event happens w.h.p.(N), we just say it happens with high probability (w.h.p.).

In order to show concentration of random variables, we will make use of the notion of negative association as defined in [13]. For completeness, we added the definition, as well as some basic cases in Appendix A. In particular, consider an experiment in which weighted balls are thrown independently into n bins according to some given distribution. For $i \in [n]$, let X_i be the total weight of balls in bin i and for an arbitrary parameter $a \geq 0$ and $i \in [n]$, let Y_i and Z_i be indicator random variables such that $Y_i = 1$ iff $X_i \leq a$ and $Z_i = 1$ iff $X_i \geq a$. It can be shown that the following lemma holds (see e.g., [10, 13]):

Lemma 2.1. *The random variables $X_1, \ldots, X_n$ (or any subset of these random variables) are negatively associated. The same is true for the random variables $Y_1, \ldots, Y_n$ and for the random variables $Z_1, \ldots, Z_n$.*

Specifically, negative association is useful because as, e.g., shown in [10], for the sum of negatively associated random variables, the usual Chernoff bounds hold. We will make use of the following bounds:

Lemma 2.2. *For a parameter $a > 0$, let $X_1, \ldots, X_n$ be independent or negatively associated non-negative random variables with $X_i \leq a$. Further, let $X := X_1 + \cdots + X_n$ and $\mu := \mathbb{E}[X]$. For $\delta > 0$, it holds that*

$$\mathbb{P}\left(X \geq (1+\delta)\mu\right) \leq \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{\mu/a}.$$

For $\delta \leq 1$, the bound can by upper bounded by $\mathbb{P}(X \geq (1+\delta)\mu) \leq e^{-\delta^2\mu/3a}$, for $\delta > 1$, it holds that $\mathbb{P}(X \geq (1+\delta)\mu) \leq e^{-\delta \ln(1+\delta)\mu/2a}$. Further, for every $\delta \in (0,1)$,

$$\mathbb{P}\left(X \leq (1-\delta)\mu\right) \leq \left(\frac{e^{-\delta}}{(1-\delta)^{1-\delta}}\right)^{\mu/a} \leq e^{-\frac{\delta^2}{2a}\cdot\mu}.$$

Lemma 2.3. *Assume there are k bins and n balls with non-negative weights $w_1, \ldots, w_n \leq 1/4$, as well as a parameter $q \in (0,1]$. Assume that $\sum_{i=1}^{n} w_i = c \cdot k/q$ for some constant $c \geq 1$. Each ball is independently selected with probability q and each selected ball is thrown into a uniformly random bin. With probability w.v.h.p.(k), there are at least $k/4$ bins in which the total weight of all balls is between $c/3$ and $2c$.*

Proof. For $i \in [k]$, let X_i be the random variable that counts the total weight of balls in bin i. Since each ball is thrown into bin i with probability q/k, we have $\mathbb{E}[X_i] = c$. We define Bernoulli random variables $Y_1, \ldots, Y_k$ and $Z_1, \ldots, Z_k$, where for all $i \in [k]$, $Y_i = 1$ iff $X_i \geq c/3$ and $Z_i = 1$ iff $X_i \leq 2c$. Using the bound from Lemma 2.2, we can bound the probabilities that $Y_i = 1$ and $Z_i = 1$ as follows:

$$\mathbb{P}(Y_i = 1) = 1 - \mathbb{P}\left(X_i < \frac{c}{3}\right) \geq 1 - \left(\frac{e^{-2/3}}{(1/3)^{1/3}}\right)^{4c} \geq 1 - \left(\frac{3}{e^2}\right)^{4/3} > \frac{2}{3}, \quad (1)$$

$$\mathbb{P}(Z_i = 1) = 1 - \mathbb{P}(X_i > 2c) \geq 1 - \left(\frac{e^1}{2^2}\right)^{4c} \geq 1 - \frac{e^4}{2^8} > \frac{3}{4}. \quad (2)$$

The random variables Y_i and Z_i are not independent, however, as stated in Lemma 2.1, $Y_1, \ldots, Y_n$, as well as $Z_1, \ldots, Z_n$ are negatively associated. Let $Y = \sum_{i=1}^{k} Y_i$ be the number of bins with balls of total weight at least $c/3$ and let $Z = \sum_{i=1}^{k} Z_i$ be the number of bins with balls of total weight at most $2c$. Because $\mathbb{E}[Y] = k \cdot \mathbb{P}(Y_i = 1) > 2k/3 = 8k/12$ and $\mathbb{E}[Z] = k \cdot \mathbb{P}(Z_i = 1) > 3k/4 = 9k/12$, by applying Lemma 2.2, we therefore get that $Y \geq 7k/12$ and $Z \geq 8k/12$ with probability w.v.h.p.(k). Therefore, there is at least $3k/12 = k/4$ bins in which the total ball weight is between $c/3$ and $2c$. □

In addition, we need a few simple results in our analysis, which we sum up in the following proposition:

Proposition 2.4. *1.* $x \in \left[0, \frac{1}{2}\right) \Rightarrow e^{-\frac{3}{2}x} \leq 1 - x \leq e^{-x}$

2. Let $n, k \in \mathbb{N}$, $\lambda_i \in [0, \frac{1}{k}]$ *for* $i = 1, 2, \ldots, n$ *and* $\sum_{i=1}^{n} \lambda_i = 1$. *Then* $\sum_{i=1}^{n} \lambda^2 \leq \frac{1}{k}$.

3 Problem

The goal of the leader election problem is for active nodes to agree on a single active node to play the role of the leader. Formally, we say an algorithm *solves the leader election problem within* $f(\mathcal{F}, t, N)$ *rounds* if it guarantees the following properties are satisfied w.h.p.(N), when executed in a network with parameters $\mathcal{F}$,t,N:

1. *Liveness:* Every node that is activated outputs the id of an active node as leader within $f(\mathcal{F}, t, N)$ rounds of being activated.
2. *Well-Formedness:* Every node that is activated performs no more than a single output.
3. *Safety:* No two nodes output different leaders.

We call an algorithm that solves the leader election problem a *leader election algorithm*.

4 Lower Bound

We establish a lower bound on leader election that we will subsequently prove to be tight (within $\log \log$ factors) in the remainder of this paper. This bound, presented below, requires that the leader election algorithm in question is also *regular* [6]. An algorithm is regular if there exists a sequence of pairs $(F_1, b_1), (F_2, b_2), ...,$ where each F_i is a probability distribution over frequencies and b_i is a probability, such that for each node u and local round r, as u has not received a message through its first r rounds, it chooses its frequency and whether or not to broadcast according to F_r and b_r, respectively. Once u receives a message we no longer restrict its behavior. Notice, all the algorithms described in this paper are regular.

We continue with the main bound:

Theorem 4.1. *Let* $\mathcal{A}$ *be a regular algorithm that solves the leader election problem in* $f(\mathcal{F}, t, N)$ *rounds. It follows that* $f(\mathcal{F}, t, N) = \Omega\left(\frac{\log^2(N)}{(\mathcal{F}-t)\log\log(N)} + \frac{\mathcal{F}t}{\mathcal{F}-t} \cdot \log N\right)$.

To prove this theorem, our strategy is to first prove that any leader election algorithm must satisfy a specific communication property. We then leverage a lemma that was adapted from our study of a related problem [6], to bound such algorithms, yielding the result claimed by our theorem.

In more detail, the lemma we adapt from [6] bounds a type of algorithm that we call *vocal*. Formally, we say an algorithm is *vocal with probability* p, if for every activation pattern[1] that includes at least 2 nodes, with probability at least p some node receives a message from another. In the setting of non-disrupted single frequency radio networks (i.e., our model with $\mathcal{F} = 1$ and $t = 0$), Jurdzinski and Stachowiak [14] proved a classic bound of $\Omega\left(\frac{\log^2(N)}{\log\log(N)}\right)$ rounds for a vocal algorithm to deliver its first message (a problem they called *wake-up*). In fact, the lemma we adapt from [6], relies, in part, on a generalization of the Jurdzinski and Stachowiak argument to a setting with multiple frequencies and disruption. Notice that Farach-Colton et al. [11] later removed the $\log\log(N)$ factor, but the same techniques we used to generalize [14] did not apply.

It is tempting to claim it is obvious that any solution to leader election must be vocal, and therefore any bound on vocal algorithms applies to leader election. And in fact there are studies of other radio network problems that make this exact claim without justification. Here we emphasize that more care is needed. Though our intuition might tell us that *some* communication is required for meaningful coordination, this is not necessarily always the case. Indeed, for any number of advanced radio network problems, one can devise algorithms that, for some activation pattens, solve the problem with no messages ever being received. Silence, in other words, can convey information. Accordingly, our first task is to formally argue that to solve leader election with high probability requires that the algorithm is vocal with an almost as high probability:

Lemma 4.2. *If algorithm $\mathcal{A}$ solves the leader election problem with probability at least $1 - \epsilon$, then with probability at least $1 - 3\epsilon$, $\mathcal{A}$ is vocal, and at least one node waits to output a leader until after the first message is received.*

Proof. Assume for contradiction that $\mathcal{A}$ is vocal in the required manner with probability less than $1 - 3\epsilon$. It follows that there exists an activation pattern P that activates at least 2 nodes, such that with probability at least 3ϵ, $\mathcal{A}$ with pattern P generates a *silent* execution in which every node elects a leader before any node receives a message. Let u and v be two nodes activated in P. Our strategy is to argue that u and v cannot distinguish a silent execution with pattern P from executions where they are alone. This will lead to a non-trivial probability that at least one of these nodes fails to solve leader election.

To formalize this intuition we need to formalize our treatment of randomness. In more detail, assume that at the beginning of each execution, a sufficiently large collection of bits is generated for the system, where each bit is determined with independent randomness. These bits are then partitioned among all N processes, and processes that end up activated use their bits to resolve their probabilistic choices. Let B be the set of all possible bit strings that could be generated for an execution. By definition, every $s \in B$ is equally likely to be generated.

Let $B_S \subset B$ be the subset of strings that, when combined with activation pattern P, generate a silent execution. By our above assumption, $|B_S| > 3\epsilon|B|$. Let $B_S^L \subseteq B_S$ be the subset of strings from B_S that, when combined with activation pattern P, are not only silent but also solve leader election (that is, all nodes output the same node as leader, and it is an active node). Given our assumption about $\mathcal{A}$ solving leader election with probability $1 - \epsilon$, it follows:

$$|B_S^L| > 2\epsilon|B|.$$

Let S_P be the set of nodes activated in P. Define $\ell : B_S^L \to S_P$, such that $\forall s \in B_S^L$, $\ell(s)$ is the single node in S_P to be the elected leader when we run $\mathcal{A}$ with s and activation pattern P. A simple counting argument tells us:

[1] By *activation pattern*, we mean the description of which nodes are activated in an execution and during which global round.

$$\exists u \in S_P, \frac{|\{s \in B_S^L : \ell(s) = u\}|}{|B_S^L|} \leq 1/2.$$

That is, at least one node is not elected leader in more than half of these bit strings.

Moving on, let $B_S^{L,\bar{u}} \subseteq B_S^L$ be the strings in B_S^L where u is not elected leader. The key observation is that for any $s \in B_S^{L,\bar{u}}$, an execution of $\mathcal{A}$ with s with an activation pattern that only activates u is indistinguishable through u's leader output w.r.t. an execution with u in pattern P, as in both cases, u makes the same random choices and receives no messages, before outputting a leader. For these strings, when u is run alone, it elects some other node as leader, which violates the properties of leader election. To complete the proof, we note that by construction:

$$|B_S^{L,\bar{u}}| \geq (1/2)|B_S^L| > \epsilon.$$

We have, therefore, identified an activation pattern (u being activated alone) for which $\mathcal{A}$ solves leader election with probability *less* than $1 - \epsilon$. A contradiction. □

We can now present our main lemma regarding vocal algorithms. This lemma is adapted from the proof arguments presented for Theorems 1 and 4 from [6].

Lemma 4.3 (from [6])**.** *Let $\mathcal{A}$ be a regular vocal algorithm that guarantees that a message is received within $f(\mathcal{F}, t, N)$ rounds of the first activation, with probability at least $1 - 1/N$. It follows that $f(\mathcal{F}, t, N) = \Omega\left(\frac{\log^2(N)}{(\mathcal{F}-t)\log\log(N)} + \frac{\mathcal{F}t}{\mathcal{F}-t} \cdot \log N\right)$.*

Returning to Theorem 4.1, the proof follows from the combination of Lemmas 4.2 and 4.3.

5 Basic Herald Algorithm

We start our description of algorithms with the simplest case, when there are no disrupted frequencies, i.e., $t = 0$. This allows us to present the basic ideas and techniques without having to worry about many of the technical difficulties that arise in the more general setting.

The described protocol runs in $O\big(\frac{\log^2 N}{\mathcal{F}} + \log N\big)$, which is tight (up to $\log\log$ factors).

Algorithm Description. For convenience, we define $F := \mathcal{F} - 2$ and assume that the $\mathcal{F}$ frequencies are $1, \ldots, F$, as well as two special frequencies $\mathcal{H}$ and $\mathcal{L}$. W.l.o.g., we assume that $F \leq \log N$, otherwise, we just only use the first $2 + \log N$ channels. We also assume for simplicity that N is a power of 2 and that F divides $\log N$. W.l.o.g., we assume the first round in which any node wakes up to be round 1.

After awaking, a node u considers itself *waiting* (state W). It stays in this state for $\Theta(\log N)$ rounds (we call that Phase 0) in which it only listens on channels $\mathcal{L}$ and $\mathcal{H}$ (with prob. $1/2$ on each of them). If u does not receive any message during Phase 0, it switches to the state *competing* (C), where it behaves as follows.

The algorithm acts in phases. Each phase lasts for $l := c \log N$ rounds, where c is an appropriately chosen constant. There are a total of $2\frac{\log N}{F} + 1$ phases. If a node finishes its last phase while still competing, then it declares itself a leader (state L) and starts broadcasting on channel $\mathcal{L}$ with probability $\frac{1}{2}$ in each round.

In a given round, a competing node chooses one of the available channels—each of them with a different probability. The highest probability is assigned to channels $\mathcal{H}$ and $\mathcal{L}$, which have a special role among all channels. If a competing node u chooses channels $\mathcal{L}$ or $\mathcal{H}$ then it listens in that round, if it chooses a channel $i \in [F]$, u listens with constant probability $\pi_\ell \leq \frac{1}{2}$ and transmits with probability $1 - \pi_\ell$ (the value of π_ℓ is determined at the end of the proof of Lemma 5.2). Once awake (i.e., active), each node u keeps track of its age $age(u)$ by counting the number of rounds it has been awake.

Most of the time a node spends listening; if it ever hears a message, then it follows these rules:

Algorithm 1: Basic herald algorithm for the undisrupted case

State description: W – waiting, C – competing, H – herald, L – leader, E – eliminated

begin
  set *phase* $:= 0$; *count* $:= 0$; *age* $:= 0$; *state* $:= W$
  while *state* $\neq E$ **do**
    count $:=$ *count* $+ 1$; *age* $:=$ *age* $+ 1$
    if *count* $= c \log N + 1$ **then**
      count $:= 1$; *phase* $:=$ *phase* $+ 1$
    if *phase* $= 1$ *and state* $= W$ **then** *state* $:= C$
    if *phase* $= \frac{2 \log N}{F} + 2$ *and state* $= C$ **then** *state* $:= L$

    switch *state* **do**
      case W : With prob. $1/2$, listen on channel $\mathcal{H}$, otherwise listen on channel $\mathcal{L}$
      case C :
        Randomly pick $r \in [0, 1)$
        Let $I := \max\left\{i : r \geq (2^{i-F} \cdot 2^{(F/2)(phase-1)})/(4N)\right\}$
        if $I > F$ *or* $I = 0$ **then**
          With prob. $1/2$, listen on channel $\mathcal{H}$, otherwise listen on channel $\mathcal{L}$
        else
          On channel I with probability π_ℓ listen or otherwise broadcast (id, age)
      case H :
        Broadcast *bc* on channel $\mathcal{H}$
        if *bc* $\neq (id, age)$ **then** *state* $:= E$ **else** *state* $:= C$
      case L : Broadcast (id, age) on channel $\mathcal{L}$ with prob. $1/2$, otherwise listen on channel $\mathcal{L}$

Upon receiving a message *msg* $= (msg.id, msg.age)$:
if *current channel is* $\mathcal{H}$ *or* $\mathcal{L}$ **then**
  if *msg.age* $\geq$ *age and msg.id* $\neq$ *id* **then** *state* $:= E$
else *// can only happen if state* $= C$
  state $:= H$
  if *msg.age* $\geq$ *age* **then**
    bc $:= (msg.age + 1, msg.id)$
  else
    bc $:= (age + 1, id)$

1. First assume that a node u hears a message on channel $\mathcal{H}$ or $\mathcal{L}$. The message contains the name and age of a node v. Note that v is not necessarily the sender of that message. If $u \neq v$ and $age(v) \geq age(u)$, then u immediately considers itself *eliminated* (state E). Eliminated nodes only listen on channel $\mathcal{L}$ to learn of a leader election. If u is older than v then u does not react to the message.

2. If a node u hears a message containing the name and age of a node v on a channel other than $\mathcal{H}$ or $\mathcal{L}$, then u considers itself a *herald* (state H) for exactly one round (i.e., for the following round). In that round, u broadcasts a message on channel $\mathcal{H}$. If the herald u is strictly older than the age of v in the message it received, then u broadcasts its own name and (current) age on channel $\mathcal{H}$. If not,

then it broadcasts the name and age of v instead (adding 1 round to comprise the fact that there is a one-round-delay). A node u that heralds the name of a node $v \neq u$ also considers itself eliminated.

A competing node in the first phase chooses one of the $\mathcal{F}$ channels using the following probabilities: It selects channel $i \in [F]$ with probability $\frac{2^{i-F}}{4N}$ and with half of the remaining probability for channels $\mathcal{H}$ and $\mathcal{L}$, respectively. Each time a node progresses to a new phase all those probabilities (except of those for channels $\mathcal{H}$ and $\mathcal{L}$) are multiplied by $2^{\frac{F}{2}}$. Channels $\mathcal{H}$ and $\mathcal{L}$ always get the remaining probability. After $2\frac{\log N}{F} + 1$ phases, the chances for channels $F, F-1, F-2, \ldots$ are $\frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \ldots$, respectively. Channel $\mathcal{H}$ and $\mathcal{L}$ are therefore both always chosen with probability more than $\frac{1}{4}$.

Analysis. The goal of the remainder of Section 5 is to prove the main theorem of the section.

Theorem 5.1. *With high probability, Algorithm 1 elects exactly one leader and it does so in $O\left(\frac{\log^2 N}{F}\right)$ rounds after the first node wakes up.*

In any round, for a competing node v we denote with $p_v(m)$ the probability of v choosing channel m and we denote by P_m the sum of the probabilities of all nodes to choose channel m, i.e., $P_m := \sum_v p_v(m)$. Note that $P_F = 2^m P_{F-m}$ and therefore $P_F \approx \frac{1}{2}\sum_{m=1}^{F} P_m$. When making a statement about the probability mass, we refer to P_F. If we state that a constant fraction of the probability mass is eliminated we mean that a collection of notes, which contribute a constant fraction to the probability mass, are eliminated. Further, in any round, $\mathcal{A}$ is the age which separates younger and older nodes in such a way that both groups contribute roughly one half of the whole probability mass. In more precise mathematical terms:

$$\mathcal{A} := \min\left\{h \in \mathbb{N}_0 : \sum_{v:age(v)\leq h} p_v(F) > \frac{P_F}{2}\right\} = \max\left\{h \in \mathbb{N}_0 : \sum_{v:age(v)\geq h} p_v(F) \geq \frac{P_F}{2}\right\}.$$

A herald is called *old*, if it broadcasts on channel $\mathcal{H}$ the id of an *old* node u, where u is old if $age(u) \geq \mathcal{A}$. We call a round in which exactly one node u becomes an old herald a *successful round*. We next show that under the appropriate circumstances, each round is successful with constant probability. In the round following a successful round, all nodes listening on channel $\mathcal{H}$ receive the old herald's message. Since the herald's message contains the fingerprint of an old node and each other node listens on channel $\mathcal{H}$ at least with probability $\frac{1}{4}$, we are able to show that in expectation at least a constant fraction of the probability mass is eliminated when this happens.

Lemma 5.2. *A round r during an execution of Algorithm 1 for which $P_F \in \left[\frac{1}{2}, 2^{F-1}\right]$ is a successful round with constant probability.*

Proof. The conditions of the lemma imply that for one channel $\lambda \in \{1, \ldots, F\}$ it holds that P_λ is in $[\frac{1}{2}, 1)$. We will prove two claims:

W.l.o.g., assume that of all nodes $V = \{v_1, \ldots, v_n\}$ the nodes $v_1, \ldots v_\kappa$ are those being awake and that $age(v_1) \geq \cdots \geq age(v_\kappa)$. Let i_0 be the smallest value such that $age(v_{i_0}) < \mathcal{A}$, i.e., $v_1, \ldots, v_{i_0-1}$ are old and $v_{i_0}, \ldots, v_\kappa$ are not. Further, let c_{ij} be the probability that on channel λ exactly nodes v_i and v_j are present in the current round, but no other node. We denote by $p_l(m)$ the probability that node v_l is on channel m. We have

$$\begin{aligned} c_{ij} &= p_i(\lambda)p_j(\lambda)\prod_{k\neq i,j}(1-p_k(\lambda)) \geq p_i(\lambda)p_j(\lambda)\prod_{k}(1-p_k(\lambda)) \\ &\overset{\text{(Prop. 2.4.1)}}{\geq} p_i(\lambda)p_j(\lambda)e^{-\frac{3}{2}P_\lambda} =: c'_{ij}. \end{aligned}$$

So the probability of having exactly two nodes on channel λ is lower bounded by the sum over all c'_{ij} with $i < j$. Since $c'_{ij} = c'_{ji}$ for all i, j, we can sum over all c'_{ij} (without restrictions in i and j), which is $P_\lambda^2 e^{-\frac{3}{2}P_\lambda}$, deduct $\sum_i c'_{ii}$ and multiply the result with $\frac{1}{2}$. However, we need also to make sure that at least one of the nodes v_i, v_j is old, so we also refrain from including any c'_{ij} for which both $i, j \geq i_0, i \neq j$. Call the event of having exactly two nodes on channel λ with at least one of them being old D_λ. Then,

$$\mathbb{P}(D_\lambda) \geq \frac{1}{2}\Bigg(\underbrace{\sum_{i,j} c'_{ij}}_{=P_\lambda^2 e^{-\frac{3}{2}P_\lambda}} - \sum_{i<i_0} c'_{ii} - \underbrace{\Big(\sum_{i\geq i_0} c'_{ii} + \sum_{i,j\geq i_0, i\neq j} c'_{ij}\Big)}_{=:C'_0} \Bigg).$$

Set $\delta_i := \frac{p_i(\lambda)}{P_\lambda}$ for $i < i_0$ (i.e., all old nodes) and $\delta_{i_0} := \frac{\sum_{j\geq i_0} p_j(\lambda)}{P_\lambda}$ (all young nodes together). Note that the δ_i's are not defined for $i \in \{i_0 + 1, \ldots, n\}$, however, they still 'consider' the probabilities of all nodes. Now we have that $\delta_i \leq \frac{1}{2}$ for all i (including i_0) and our choice of δ_{i_0} also ensures that $\sum_{i=1}^{i_0} \delta_i = 1$, which allows us to apply Proposition 2.4.2. Furthermore it holds that $c'_{ii} = \delta_i^2 P_\lambda^2 e^{-\frac{3}{2}P_\lambda}$ and $C'_0 = \delta_{i_0}^2 P_\lambda^2 e^{-\frac{3}{2}P_\lambda}$. We get

$$\mathbb{P}(D_\lambda) \geq \frac{1}{2} P_\lambda^2 e^{-\frac{3}{2}P_\lambda}\left(1 - \sum_{i<i_0} \delta_i^2 - \delta_{i_0}^2\right) \geq \frac{1}{4} P_\lambda^2 e^{-\frac{3}{2}P_\lambda} = \Omega(1)$$

We call a transmission on a channel $m \in [F]$ *successful*, if *exactly one* node transmits on that channel, *exactly one* node listens and at least one of those nodes is old—we denote that event by A_m. The probability of having a successful transmission on some channel m conditioned on having exactly two nodes on that channel is $2\pi_\ell(1 - \pi_\ell) \geq \pi_\ell$. Since we also know that $\mathbb{P}(D_\lambda) = \Omega(1)$, we have

$$\mathbb{P}(A_\lambda) = \Omega(\pi_\ell).$$

That finishes the proof of our first claim. We continue with proving the second one.

To have a successful round, we have to have exactly one old herald being created in that round. This probability is lower bounded by the probability that this herald is created on channel λ while none is created on any other channel. Let B_m be the event that *exactly one* node transmits and *at least one* node listens on channel m:

$$\mathbb{P}(\text{successful round}) \geq \mathbb{P}(A_\lambda)\left(1 - \mathbb{P}\left(\bigcup_{m\in[F]\setminus\lambda} B_m \,\middle|\, A_\lambda\right)\right) \geq \mathbb{P}(A_\lambda)\left(1 - \sum_{m\in[F]\setminus\lambda} \mathbb{P}(B_m|A_\lambda)\right).$$

When we condition on A_λ, there are $\kappa - 2$ nodes that are on channels different from λ. Let $u_1, \ldots, u_{\kappa-2} \subset \{v_1, \ldots, v_\kappa\}$ be these nodes. Let i, j, k be in $\{1, \ldots \kappa - 2\}$ in the following calculations. Conditioning on A_λ the probability for any node u_k being on channel m increases by $(1 - p_k(\lambda))^{-1}$. That is,

$$\mathbb{P}(u_k \text{ is on channel } m|A_\lambda) = \frac{p_k(m)}{1 - p_k(\lambda)},$$

and moreover we have

$$\mathbb{P}(u_k \text{ is not transmitting on channel } m|A_\lambda) = 1 - \frac{(1 - \pi_\ell)p_k(m)}{1 - p_k(\lambda)}.$$

Let $B_m^{i,j}$ be the event that u_i is on channel m, no other node transmits on channel m, and u_j listens on

channel m. For $m \in [F] \setminus \lambda$, we have

$$\begin{aligned}
\mathbb{P}\big(B_m^{i,j}|A_\lambda\big) &= \frac{p_i(m)}{1-p_i(\lambda)}(1-\pi_\ell)\frac{p_j(m)}{1-p_j(\lambda)}\pi_\ell \prod_{k\notin\{i,j,\kappa-1,\kappa\}}\left(1-\frac{(1-\pi_\ell)p_k(m)}{1-p_k(\lambda)}\right)\\
&= p_i(m)p_j(m)\pi_\ell \overbrace{(1-\pi_\ell)(1-p_{\kappa-1}(\lambda))(1-p_\kappa(\lambda))}^{\leq 1}\\
&\quad \prod_{k\leq\kappa}\frac{1-p_k(\lambda)-(1-\pi_\ell)p_k(m)}{1-p_k(\lambda)} \prod_{k=i,j,\kappa-1,\kappa}(1-\underbrace{p_k(\lambda)-(1-\pi_\ell)p_k(m)}_{\leq\frac{1}{4}+\frac{1}{8}})^{-1}\\
&\overset{\text{(Prop. 2.4.1)}}{\leq} p_i(m)p_j(m)\,\pi_\ell\, e^{\frac{3}{2}P_\lambda}\, e^{-P_\lambda-P_m(1-\pi_\ell)}(8/5)^4\\
&\leq p_i(m)p_j(m)\,\pi_\ell\, e^3 e^{-\frac{P_m}{2}}.
\end{aligned}$$

Because $B_m = \bigcup_{i,j\in[\kappa-2],\, i\neq j} B_m^{i,j}$, applying a union bound yields

$$\mathbb{P}(B_m|A_\lambda) \leq \sum_{i,j\in[\kappa-2],\, i\neq j}\mathbb{P}(B_m^{i,j}|A_\lambda) \leq P_m^2 e^3\pi_\ell e^{-\frac{P_m}{2}} =: C_m.$$

For a fixed m this value is $O(\pi_\ell)$.

We now upper bound the sum over all C_m. We set $\lambda' := \min\{\lambda+4, F\}$. Remember that $P_{m+1} = 2P_m\ \forall m \in \{1,\dots,F-1\}$.

$$\frac{C_{m+1}}{C_m} = 4e^{-P_m+\frac{P_m}{2}} = 4e^{-\frac{P_m}{2}} < \frac{1}{2}, \quad \forall m \geq \lambda'$$

and also

$$\frac{C_{m-1}}{C_m} = \frac{1}{4}e^{-\frac{P_m}{4}+\frac{P_m}{2}} = \frac{1}{4}e^{\frac{P_m}{4}} < \frac{1}{2}, \quad \forall m \leq \lambda$$

That is, the further a frequency is away from λ, the smaller the probability that there is a herald created, more precisely, after some constant distance from channel λ, probabilities drop by at least $\frac{1}{2}$ with each step within the frequencies (actually way more). We thus make use of geometric series to upper bound all frequencies outside $[\lambda,\dots,\lambda']$ simultaneously by $C_\lambda + C_{\lambda'}$. As mentioned above $C_m = O(\pi_\ell)$ for a fixed m. Thus, in total, we have

$$\sum_{m\neq\lambda}\mathbb{P}(B_m|A_\lambda) \leq \sum_m C_m \leq C_\lambda + C_{\lambda'} + \sum_{m=\lambda}^{\lambda'} C_m = O(\pi_\ell).$$

Choosing π_ℓ small enough but still constant, then this value is less than 1. On the other hand, since we choose π_ℓ to be a constant the event A_λ happens with constant probability and thus the probability of a successful round is also a constant, concluding the proof. □

Because in every round with constant probability an old herald is created, we can argue that in expectation all listening nodes on channel $\mathcal{H}$ of the same or lower age as the age being heralded are eliminated. By the definition of $\mathcal{A}$ we have that every time a successful round happens, a constant fraction of the probability mass is eliminated with constant probability.

Lemma 5.3. *With high probability, at all times, it holds that $P_F \leq 2^{F-1}$. Further, with high probability, there are no $l/2$ consecutive rounds in which $P_F \geq \frac{1}{2}$.*

Proof. There are two ways for P_F to increase: Either a node is already actively contributing to P_F and finishes a phase or it switches from the waiting state to the competing state. The former allows an increase of that node's contribution by $2^{\frac{F}{2}}$ while the latter is an absolute increase from 0 to $\frac{1}{4N}$. However, there are at most N nodes that can switch from the waiting state to the competing state, so that contribution is at most $\frac{1}{4}$, thus comparably small.

Thus, within $l = c \log N$ rounds (i.e., within the time of one phase), P_F can not increase by more than a factor of $2^{\frac{F}{2}}$ and an additive value of $\frac{1}{4}$. This guarantees that for P_F to exceed 2^{F-1} it must hold for at least l rounds that $P_F \geq \frac{1}{2}$, so all the requirements for lemma 5.2 are fulfilled.

If the constant c in the algorithm is chosen large enough, applying Chernoff bounds guarantees that, w.h.p., there are $\Omega(\log N)$ successful rounds. Each of these rounds eliminates a constant fraction of the probability mass in expectation (by choice of $\mathcal{A}$ and since the probability of listening on channel $\mathcal{H}$ is at least $\frac{1}{4}$). But since $\log N = \Omega(F)$, w.h.p., l rounds are thus sufficient have P_F drop back to a value smaller than $\frac{1}{2}$, a contradiction.

Using an identical argument and that $P_F \leq 2^{F-1}$ at all times, it also follows that $P_F \geq 1/2$ cannot hold for $l/2$ consecutive rounds. □

We are now ready to prove Theorem 5.1.

Proof of Theorem 5.1. The running time is clear from the construction of the algorithm as well is the fact that at least one leader is elected since among the nodes that wake up in round 1, there is at least one node that finishes its last phase without being eliminated.

For the sake of contradiction, assume that more than one leader is elected with probability more than $1/N$ (or any polynomial in N). W.l.o.g., let v_1 and v_2 be elected leaders with $age(v1) \geq age(v_2)$.

Assume first that $age(v_2) \geq age(v_1)+l/2$. Then they both are in the last phase for at least $l/2$ rounds. In that time interval it holds that $p_1(F) = p_i(F) = \frac{1}{4}$, i.e., $P_F \geq \frac{1}{2}$ for $l/2$ consecutive rounds, a contradition to the second claim of Lemma 5.3. Hence, w.h.p., $age(v_2) < age(v_1) + l/2$. But then, v_1 is already a leader for $l/2$ rounds before any other node becomes a leader. W.h.p. this is sufficiently long for v_2 to hear node v_1 on channel $\mathcal{L}$ before becoming a leader. Consequently, there is only one leader w.h.p. □

6 Herald Algorithm Tolerating Disruption

In this section, we address the problem of up to $t \in [1, \rho\mathcal{F}]$ nonfunctional frequencies, for some constant $\rho < 1/3$. We will show that a slight adaption of the basic herald algorithm provides an asymptotically optimal algorithm under the assumption that ρ is known to all nodes. For simplicity[2], we assume that the number of disrupted channels is $t \leq \mathcal{F}/6$. Also for simplicity, we assume that $2t$ divides $\mathcal{F}$. As discussed in Section 2, the adversary can freely decide which channels to disrupt just before a round starts.

The algorithm we present works similar to Algorithm 1, except that each used channel in the original algorithm is replaced by a *block* of $2t$ channels with a total of $F + 2 = \frac{\mathcal{F}}{2t}$ blocks. Similarly to Section 5, we again assume that $F \leq \log N$. If it is not, we can just only use the first $2t(\log N + 2)$ channels. Analogously to the algorithm in Section 5, we name two of the blocks $\mathcal{H}$ and $\mathcal{L}$ and the remaining blocks $1, 2, \ldots, F$. Each block consists of $2t$ *sub-channels* and in a particular block b for a particular sub-channel s the corresponding channel is denoted by *channel* (b, s). Each round nodes choose a block b in the same manner as they choose channels in Algorithm 1. On the selected block b, they choose a uniformly random sub-channel s. After choosing a channel (b, s), nodes continue in the same manner as in Algorithm 1. The second change is that for a competing node each phase now only lasts for $\Theta\big(F + \frac{\log N}{t}\big)$ rounds. The third change is that a node that moves to state L does not immediately consider itself a leader, but it first becomes

[2] A natural generalization yields $t < \rho\mathcal{F}$ for any $\rho < 1/3$: divide channels into blocks of $t(1 + \epsilon)$ instead of blocks of size $2t$.

Algorithm 2: Herald algorithm for disrupted channels

Major changes to the $t = 0$ case (besides notational changes regarding the switch from channels to blocks) are marked with an $\rightarrow$ symbol.

State description: W – waiting, C – competing, H – herald, L – leader, E – eliminated

begin
  set *phase* $:= 0$; *count* $:= 0$; *age* $:= 0$; *state* $:= W$
  while *state* $\neq E$ **do**
    count $:=$ *count* $+ 1$; *age* $:=$ *age* $+ 1$
→ **if** *count* $= c(F + \log N/t) + 1$ **then**
      count $:= 1$; *phase* $:=$ *phase* $+ 1$
    if *phase* $= 1$ *and state* $= W$ **then** *state* $:= C$
    if *phase* $= \frac{2 \log N}{F} + 2$ *and state* $= C$ **then** *state* $:= L$
→ Pick s uniformly at random out of $\{1, 2, \ldots, 2t\}$
    switch *state* **do**
      case W : With prob. $1/2$, listen on channel $(\mathcal{H}, s)$, otherwise listen on channel $(\mathcal{L}, s)$
      case C :
        Randomly pick $r \in [0, 1)$
        Let $I := \max\left\{i : r \geq (2^{i-F} \cdot 2^{(F/2)(\textit{phase}-1)})/(4N)\right\}$
        if $I > F$ *or* $I = 0$ **then**
          With prob. $1/2$, listen on channel $(\mathcal{H}, s)$, otherwise listen on channel $(\mathcal{L}, s)$
        else
          On channel (I, s) with probability π_ℓ listen or otherwise broadcast $(\textit{id}, \textit{age})$
      case H :
        Broadcast *bc* on channel $(\mathcal{H}, s)$
        if $\textit{bc} \neq (\textit{id}, \textit{age})$ **then** *state* $:= E$ **else** *state* $:= C$
      case L :
→ **if** $\textit{age} > c'(\log^2 N/\mathcal{F} + t \log N)$ **then** Consider yourself a leader
        On channel $(\mathcal{L}, s)$ with prob. $1/2$ listen, otherwise broadcast $(\textit{id}, \textit{age})$

Upon receiving a message ***msg*** $= (\textit{msg.id}, \textit{msg.age})$**:**
if *current block is* $\mathcal{H}$ *or* $\mathcal{L}$ **then**
  if *msg.age* $\geq$ *age and msg.id* $\neq$ *id* **then** *state* $:= E$
else *// can only happen if state* $= C$
  state $:= H$
  if *msg.age* $\geq$ *age* **then**
    bc $:= (\textit{msg.age} + 1, \textit{msg.id})$
  else
    bc $:= (\textit{age} + 1, \textit{id})$

a *candidate*. Candidates listen or broadcast their id and age on a uniformly random channel on block $\mathcal{L}$. If a candidate does not get eliminated after being in state L for $\Theta(t \log N)$ rounds, it considers itself a *leader*. Any node that receives a message on $\mathcal{L}$ can calculate itself from the age being broadcasted whether that node is already a leader or not.

Analysis. Analogously to section 5 we denote with P_m the sum of all nodes' probabilities to choose block m. $\mathcal{A}$ and the notion of an old herald are also analogously defined. With $l = c\big(F + \frac{\log N}{t}\big)$ we denote the length of one phase. Our primary goal of this subsection is to prove the following main theorem of the section.

Theorem 6.1. *With high probability, Algorithm 2 elects exactly one leader and it does so in $O\big(\frac{\log^2 N}{\mathcal{F}} + t \log N\big)$ rounds after the first node wakes up.*

First we prove in Lemma 6.2 that once the total probability mass exceeds a certain threshold, w.v.h.p.(t) each round $\Omega(t)$ heralds of age $\mathcal{A}$ or higher are created. In Lemma 6.3 we show that also the total number of heralds created each round is of order $O(t)$ w.v.h.p.(t). Both together provide for the fact that each round, w.v.h.p.(t), a constant fraction of the total probability mass is eliminated. Since that happens w.v.h.p.(t), a phase does not need to last longer than $\Theta\big(F + \frac{\log N}{t}\big)$ rounds to guarantee a probability mass reduction of order $\Omega(2^F)$ w.h.p.(N). We finally show that, w.h.p., at any time the number of nodes in state L is in $O(t)$. $\Theta(t \log N)$ rounds after the first nodes move to state L, one of them can safely declare itself a leader.

Lemma 6.2. *If $P_F \in [2t, 2^F t]$ in any round, then w.v.h.p.(t) $\Omega(t)$ heralds are created, which herald the name and age of a node of age at least $\mathcal{A}$ on one of the channels in block $\mathcal{H}$ in the following round.*

Proof. Following the lines of the proof for Lemma 5.2 we get that there is a block λ for which $P_\lambda \in [2t, 4t]$. Let $Dis_\lambda \subset \{1, 2, \ldots, 2t\}$ be the subset of disrupted sub-channels in block λ and let $t_\lambda := |Dis_\lambda| \leq t$. Then define $q := 1 - \frac{t_\lambda}{2t} \in (0, 1]$, $k := 2t - t_\lambda$, $c := \frac{qP_\lambda}{k} = \frac{P_\lambda}{2t} \geq 1$ and $w_i := p_i(\lambda)$ for $i = 1, 2, \ldots, n$, where $p_i(\lambda)$ denotes the probability of node i to choose block λ. Note that $\sum_{i=1}^{n} w_i = P_\lambda$. We now apply Lemma 2.3, where the bins are the set of $k := 2t - t_\lambda \geq t$ undisrupted sub-channels in block λ, i.e., the channels $[2t] \setminus Dis_\lambda$. W.v.h.p.($t$) we get that on at least $k/4 \geq t/4$ of these sub-channels the total probability mass to choose block λ is in $[1/3, 4]$.

Clearly the age of a node has no impact regarding the choice of a sub-channel. Thus, due to symmetry and with analogous reasoning as in Lemma 5.2, we get that, independently, on each of those $k/4$ channels, with constant probability a herald is created that broadcasts the age of an old node in the following round. Hence, in expectation, the number of heralds created on block λ is $\Omega(t)$. Using a standard Chernoff bound (cf. Lemma 2.2), we also get that w.v.h.p.(t) that number is $\Omega(t)$. □

In Algorithm 1 we needed to make sure that there is a constant chance to have exactly one herald being created. Here we need to be more thorough to maintain an optimal running time: we need to make sure that $\Theta(t)$ heralds are created w.v.h.p.(t) in *each round*.

Lemma 6.3. *If $P_F \in [2t, 2^F t]$ in any round, then w.v.h.p.(t) $O(t)$ heralds are created.*

Proof. Let λ denote the block in which $P_\lambda = \rho 2t$ with $\rho \in [1, 2)$. For $b \in \{1, \ldots, F\}$, $s \in \{1, \ldots, 2t\}$ and $v \in V := \{1, \ldots, n\}$ let $Z_{b,s,v}$ be the random variable that indicates whether node v chooses channel (b, s), let $Z_{b,s} = \sum_{v \in V} Z_{b,s,v}$ count the number of nodes on channel (b, s) and for $i \geq 1$ let $X_{b,s}^{(i)}$ be an indicator random variable for which $X_{b,s}^{(i)} = 1$ iff $Z_{b,s} \leq 2^i$ and $X_{b,s}^{(i)} = 0$ otherwise. By Lemma 2.1 the random variables $Z_{b,s,v}$ are negatively associated. The same is also true for the random variables $Z_{b,s}$ and the random variables $X_{b,s}^{(i)}$ for a fixed i.

We define the *node-range* $R_i := [2^{i-1} + 1, 2^i]$ and say that a channel (b, s) is in node-range i iff $X_{b,s} \in R_i$. We also let H_i be the number of heralds created on all channels in range i and set $\gamma := \ln(1/\pi_\ell)$. The proof is now carried out as follows:

(I) First, we show that for $\xi \geq 0$, the total number of nodes in all blocks b for $b \leq \lambda + \xi$ is $O(2^\xi t)$ w.v.h.p.(t).

(II) We then show that w.v.h.p.(t), for every $i \geq 0$, all blocks b with $b > \lambda + i$ together have at most $O(t)$ channels with at most 2^i nodes.

(III) Next, we reason that w.v.h.p.(t), for every $i \geq 1$, there are at most νt channels in node-range R_i for a sufficiently large constant ν.

(IV) In the fourth part, we show that w.v.h.p.(t), channels in node-ranges larger than $\log t$ do not create any heralds at all.

(V) Next we show that $\mathbb{P}(H_i \geq 2^i e^{1-\frac{\gamma}{2}2^{i-1}}\nu t + x + 2^i) \leq e^{-\gamma x/4}$. Essentially, x measures how much H_i exceeds a safe value which would guarantee a total of $O(t)$ heralds over all node ranges.

(VI) In the last step, we use a probabilistic argument to union bound over all 'bad cases' to show that w.v.h.p.(t) on channels in node-range between 1 and $\log t$, in total only $O(t)$ heralds are created.

Part (I): The total probability for nodes to land on channels in all blocks $b \leq \lambda + \xi$ for $\xi \geq 0$ is at most $\sum_{i=-\xi}^{\infty} P_\lambda \cdot 2^{-i} = O(2^\xi P_\lambda) = O(2^\xi t)$. Therefore, the number of nodes $\sum_{b=1}^{\lambda+\xi}\sum_{s=1}^{2t}\sum_{v\in V} Z_{b,s,v}$ on channels (b, s) in all blocks $b \leq \lambda + \xi$ is $O(2^\xi t)$ in expectation. Because the $Z_{b,s,v}$ are negatively associated 0/1 random variables, we can apply a standard Chernoff bound to get that the number of nodes on channels (b, i), $b \leq \lambda + \xi$ is also $O(2^\xi t)$ w.v.h.p.(t).

Part (II): Assume that $i \geq 0$, $b > \lambda+i$, and $s \in [2t]$. Further, let $\mu := \mathbb{E}[Z_{b,s}] = \rho 2^{b-\lambda}$, $k := (b-\lambda)-i \geq 1$ and $\delta := 1 - \rho^{-1}2^{-k} > 1/2$. By applying Lemma 2.2, we then get

$$p_{k,i} := \mathbb{P}(X_{b,s}^{(i)} = 1) = \mathbb{P}(Z_{b,s} \leq 2^i) = \mathbb{P}(Z_{b,s} \leq (1-\delta)\mu) \leq e^{-\delta^2\mu/2} \leq e^{-2^{k+i-3}}.$$

Let $X^i = \sum_{b=\lambda+i+1}^{F}\sum_{s=1}^{2t} X_{b,s}^{(i)}$ be the number of channels on blocks $b > \lambda + i$ with at most 2^i nodes. Then:

$$\mathbb{E}[X^{(i)}] = \sum_{k=1}^{F-\lambda-i}\sum_{s=1}^{2t}\mathbb{P}\big(X_{b,s}^{(i)} = 1\big) < 2t \cdot \sum_{k=1}^{\infty} e^{-2^{k+i-3}} = O(t).$$

As before, because $X^{(i)}$ is the sum of negatively associates 0/1 random variables, we can apply a standard Chernoff bound to get that $X^{(i)} = O(1)$ w.v.h.p.(t).

Part (III): Let C_i be the number of channels in node-range i, i.e., C_i is the number of channels with at least $2^{i-1} + 1$ and at most 2^i nodes. By Part (I), w.v.h.p.(t), the number of nodes on blocks $b \leq \lambda + i$ is at most $O(2^i t)$. Since each channel in node-range i has at least $2^{i-1} + 1$ nodes, w.v.h.p.(t), at most $O(t)$ channels on blocks $b \leq \lambda + i$ are in node-range i. By Part (II), w.v.h.p.(t), at most $O(t)$ channels on blocks $b > \lambda + i$ have at most 2^i nodes. Hence, overall, there are at most $O(t)$ channels in node-range i.

Part (IV): Let Y_i be the number of channels in node-range i on which heralds are created. The probability of creating a herald on a specific channel with $m \in R_i$ nodes can be computed as $m(1-\pi_\ell)\pi_\ell^{m-1} \leq 2^i\pi_\ell^{2^{i-1}} = 2^i e^{-\gamma 2^{i-1}} =: p_i$, where $\gamma = \ln(1/\pi_\ell) = \Theta(1)$.

$$\mathbb{P}\Big(\sum_{i=\log t+1}^{\infty} Y_i \geq 1\Big) \leq \sum_{i=\log t+1}^{\infty} C_i p_i \leq \sum_{i=\log t+1}^{\infty} O(t)2^i e^{-\gamma 2^{i-1}} = e^{-\gamma' t}, \quad \text{for some constant } \gamma' > 0.$$

The first inequality follows from a union bound over all C_i channels in node-range i, the second inequality follows from Part (III).

Part (V): Recall that p_i is an upper bound on the probability that heralds are created on a specific channel in node-range i. By Part (III), the number of channels in node-range i is at most $C_i \leq \nu t$ w.v.h.p.(t). If $C_i \leq \nu t$, the random variable Y_i is dominated by a binomial random variable $\hat{Y}_i \sim \mathrm{Bin}(\nu t, p_i)$. If on a channel with m nodes any heralds are created, then there are $m-1$ heralds created at once, so we have that the random variable H_i is dominated by $2^i Y_i$. We therefore have $\mathbb{P}(H_i \geq h) \leq \mathbb{P}(Y_i \geq h/2^i) \leq \mathbb{P}(\hat{Y}_i \geq h/2^i)$. For $k \geq 0$, we thus have

$$\mathbb{P}(H_i \geq 2^i k) \leq \mathbb{P}(\hat{Y}_i \geq k) \leq \binom{\nu t}{k} p_i^k \leq \left(\frac{e\nu t}{k}\right)^k p_i^k = \left(\frac{e 2^i \nu t}{k e^{\gamma 2^{i-1}}}\right)^k. \tag{3}$$

Let $\alpha_i := 2^i e \cdot e^{-\frac{\gamma}{2} 2^{i-1}} \nu t$ and $k_i := \left\lceil \frac{x}{2^i} + \alpha_i \right\rceil$. Applying (3), we then get

$$\mathbb{P}(H_i \geq 4^i e^{1-\frac{\gamma}{2}2^{i-1}} \nu t + x + 2^i) \leq \mathbb{P}(H_i \geq 2^i k_i) \leq \left(\frac{e 2^i \nu t}{k_i e^{\gamma 2^{i-1}}}\right)^{k_i \ (k_i \geq \alpha_i)} \leq \left(\frac{1}{e^{\frac{\gamma}{2} 2^{i-1}}}\right)^{k_i \ (k_i \geq x/2^i)} \leq e^{-\gamma x/4}.$$

Part (VI): As before let $\alpha_i := 2^i e^{1-\frac{\gamma}{2}2^{i-1}} \nu t$ and note that $\hat{\alpha} := \sum_{i=1}^{\log t} 2^i \alpha_i = O(t)$. Let η be chosen such that $\eta t - \hat{\alpha} \geq 6t$ and such that ηt is an integer.

By Part (IV), w.v.h.p.(t), heralds are only created on channels in node-ranges $i \leq \log t$. For the remainder of the proof, we condition on the fact that this is indeed the case. We call a vector $h = (h_1, \ldots, h_{\log t})$ for $h_i \in \mathbb{N}_0$ a *herald vector* with *weight* $w_h = h_1 + \cdots + h_{\log t}$; we call such a herald vector *heavy* iff $w_h = \eta t$. Consider the herald vector $H = (H_1, \ldots, H_{\log t})$, then its weight w_H counts the total number of heralds created. We say that H is lower bounded by h if $H_i \geq h_i$ for all $i \in [\log t]$, written as $H \geq h$.

If more than ηt heralds are created, then clearly $H \geq \bar{h}$ for some specific heavy herald vector $\bar{h}$. We will show that the probability of such an event is exponentially small in t. Because there are at most $(\eta t+1)^{\log t} = e^{O(\log^2 t)}$ different heavy herald vectors, the lemma then follows by a union bound over all heavy herald vectors.

Let $\bar{h} = (h_1, \ldots, h_{\log t})$ be a heavy herald vector. We define $x_i := \max\left\{0, h_i - 2^i \alpha_i - 2^i\right\}$, where $i \in [\log t]$. By Part (V) we have

$$\mathbb{P}(H \geq \bar{h}) \leq \prod_{i=1}^{\log t} e^{-\gamma x_i/4} = e^{-\frac{\gamma}{4} \cdot \sum_{i=1}^{\log t} x_i}.$$

We therefore need to show that $\sum_{i=1}^{\log t} x_i = \Omega(t)$. However since $x_i \geq h_i - 2^i \alpha_i - 2^i$, we have

$$\sum_{i=1}^{\log t} x_i \geq \sum_{i=1}^{\log t} \left(h_i - 2^i \alpha_i - 2^i\right) = \eta t - \hat{\alpha} - 2t \geq 4t.$$

We therefore get that $\mathbb{P}(H \geq \bar{h}) \leq e^{-\gamma t}$. Choosing γ appropriately this concludes the proof since

$$\mathbb{P}(w_H \geq \eta t) \leq \sum_{\bar{h} \text{ heavy}} \mathbb{P}(H \geq \bar{h}) = e^{-\bar{\gamma} t}, \quad \text{for some constant } \bar{\gamma} > 0.$$

□

Both lemmas provide that for $P_F \geq 2$ the total number of heralds created in a single round is in $\Theta(t)$.

Lemma 6.4. *With high probability, at all times, it holds that $P_F < t2^F$. Further, for appropriately chosen constant c, with high probability, there are no $l/2$ consecutive rounds in which $P_F \geq 2t$.*

Proof. As in the case of the analysis in Section 5, there are two ways in which P_F can increase. Either some nodes switch to a new phase or there are new nodes switching from the waiting into the competing state. Analogously to the argument in the proof of Lemma 5.3, within the time of one phase, P_F can only increase by a factor of $2^{F/2}$ and a small additive amount. In order for P_F to exceed $t2^F$, P_F therefore has to be at least $2t$ for l consecutive rounds (i.e., for the duration of one phase). We show that for c sufficiently large, w.h.p., this cannot be the case.

Consider some round in which $P_F \in [2t, t2^F]$ and assume that there are $\hat{t} \in \{t, \ldots, 2t\}$ channels on block $\mathcal{H}$ that are undisrupted. Let $X_1, X_2, \ldots, X_{\hat{t}}$ be the random variables counting the number of heralds on each of the $\hat{t}$ undisrupted channels in block $\mathcal{H}$. By Lemma 2.1 the random variables $X_1, \ldots, X_{\hat{t}}$ are negatively associated. The same is true for the indicator random variables $X_{i,\leq\alpha}$ that take value 1 iff $X_i \leq \alpha$, where $i \in [\hat{t}]$ and $\alpha \geq 0$. We define $X_{\leq\alpha} := \sum_{i=1}^{\hat{t}} X_{i,\leq\alpha}$ and we let $p_{\leq\alpha}$ denote the probability that at most α heralds are on a specific single channel. Let $2\eta t$ be the number of heralds on block $\mathcal{H}$ for this round for some $\eta = \Theta(t)$. We then have

$$\begin{aligned} p_0 &:= p_{\leq 0} = \left(1 - \frac{1}{2t}\right)^{2\eta t} \overset{\text{(Prop. 2.4.1)}}{\geq} e^{-\frac{3}{2}\eta}, \\ p_{\leq 1} &= p_0 + \eta t \cdot \frac{1}{2t}\left(1 - \frac{1}{2t}\right)^{2\eta t - 1} = p_0\left(1 + \frac{2\eta t}{2t-1}\right) \geq p_0\,(1+\eta)\,. \end{aligned}$$

Since η is a constant, the two probabilities p_0 and $p_{\leq 1}$ are constant as well. Therefore

$$\mu_{\leq\alpha} := \mathbb{E}[X_{\leq\alpha}] = p_{\leq\alpha}\hat{t}.$$

We can apply Lemma 2.2 to get that for $\delta \leq 1$,

$$\mathbb{P}(|X_{\leq\alpha} - \mu_{\leq\alpha}| \geq \delta\mu_{\leq\alpha}) \leq e^{-\delta^2\Theta(t)}.$$

In other words, for constant δ the $X_{\leq\alpha}$ are close to their expected values w.v.h.p.(t). Therefore if we choose δ small enough ($\delta < \eta/(2\eta+4)$), w.v.h.p.(t), we obtain

$$\begin{aligned} X_{\leq 1} - X_{\leq 0} &\geq (1-\delta)\mu_{\leq 1} - (1+\delta)\mu_{\leq 0} \geq \big((1-\delta)(1+\eta) - (1+\delta)\big) \cdot p_0\hat{t} \\ &= \big(\eta - \delta(2+\eta)\big) \cdot p_0\hat{t} = \Theta(t). \end{aligned}$$

Therefore, w.v.h.p.(t), a constant fraction of the undisrupted channels on block $\mathcal{H}$ have exactly one herald. By Lemmas 6.2 and 6.3, w.v.h.p.(t), there are $\Theta(t)$ heralds and a constant fraction of these heralds has age at least $\mathcal{A}$. By symmetry, the $\Theta(t)$ heralds that broadcast alone on some undisrupted channel of block $\mathcal{H}$ are a uniformly random subset of all heralds. Hence, w.v.h.p.(t), a constant fraction of the heralds that broadcast alone on an undisrupted channel of $\mathcal{H}$ are old. Each node that listens on $\mathcal{H}$ therefore has a constant probability of picking a sub-channel with exactly one old herald.

Because each node has a constant probability to listen on $\mathcal{H}$, w.v.h.p.(t), a constant fraction of the total probability contributing to P_F is eliminated in *each* round for which $P_F \geq 2t$. Assume that for constants $\hat{\gamma} > 0$ and $\hat{s} > 1$, with probability $p := 1 - e^{-\hat{\gamma}t}$, a $1/\hat{s}$-fraction of the total probability mass is eliminated and let us call a round successful if an $1/\hat{s}$-fraction of the total probability mass is eliminated. In order to get from some $P_F < t2^F$ to $P_F < 2t$ w.h.p., we need $\Theta(F)$ successful rounds in a time span of l rounds w.h.p.(N).

If $1 - p = e^{-\hat{\gamma}t}$ is more than e^{-3}, then t is less than $3/\hat{\gamma} = O(1)$, having a phase lasting $\Omega(\log N)$ rounds, each being successful with a constant probability. A standard Chernoff argument gives us that $\Theta(l)$ of them are successful w.h.p.(N).

If $1-p$ is less than e^{-3}, then we apply Chernoff again to show that less than a constant fraction of l rounds are *not* successful. By choosing $\delta = e^{\hat{\gamma}t-1} \geq e^2$ and letting X count the number of unsuccessful rounds we get:

$$\begin{array}{rcl} \mathbb{P}(X \geq l/2) & < & \mathbb{P}(X \geq e \cdot e^{-\hat{\gamma}t} l) = \mathbb{P}(X \geq (1+\delta)\mu) \leq e^{-\mu \overbrace{((\delta+1)\ln(\delta+1)-\delta)}^{>0}} \\ & \overset{\mu = e^{-\hat{\gamma}t} l}{\leq} & e^{-l(e^{-1}(\hat{\gamma}t-1)-e^{-1}+e^{-\hat{\gamma}t})} \overset{l \geq c\frac{\log N}{t}}{\leq} N^{-\frac{c}{t}\frac{\hat{\gamma}t-2}{e}} \leq N^{-c\frac{\hat{\gamma}}{3e}} \end{array}$$

That is, w.h.p.(N), a constant fraction of l rounds are successful. If c is chosen large enough, it follows in the same way that, w.h.p., the maximum number of consecutive rounds in which $P_F \geq 2t$ is less than $l/2$. □

Lemma 6.5. *With high probability, at all times, the number of candidates is $O(t)$.*

Proof. We show that at all times, the number of candidates is less than $16t$ with probability at least $1 - N^{-d}$, where the constant d can be chosen arbitrarily. For the sake of contradiction, assume that there is a first round r_0 in which the number of candidates is at least $16t$ with probability more than N^{-d}.

We first show that between round $r' = r_0 - l/2$ and r_0 at most $8t$ of all active nodes switch to state L. Assume that this is not the case. All these $8t$ nodes are therefore together in the last competing phase for at least $cF/2$ rounds. Because in these rounds, all $8t$ nodes choose channel F with probability $1/4$, this implies that there are $cF/2$ consecutive rounds in which $P_F \geq 2t$, something that does not happen w.h.p. according to Lemma 6.4.

Hence, w.h.p., more than $8t$ of the candidates have already been active in round r'. Because we also assumed that r_0 is the first time, where the number of candidates is at least $16t$, this also implies that overall between rounds r' and r_0, there are less than $24t$ different candidates, i.e., $n \in [8t, 24t]$ different candidates.

Consider some round $r \in [r', r_0]$. Assume that in round r, $\hat{t} \in [t, 2t]$ of the $2t$ channels in block $\mathcal{L}$ are not disrupted. Each of the n candidates picks a uniformly random channel. Applying Lemma 2.3 with parameters $q = \hat{t}/(2t)$, $k = \hat{t}$, and $w_i = 1$ for $i \in [n]$—which implies $c = \frac{n}{2\hat{t}} \in [4, 12]$— we get that w.v.h.p.($t$), on at least $\hat{t}/4$ channels, there are between $\lceil 4/3 \rceil = 2$ and 24 nodes. On all these channels, independently, there is a constant probability that exactly one candidate broadcasts. Hence, w.v.h.p.(t), there are $\Theta(t)$ channels on $\mathcal{L}$ on which exactly one candidate broadcasts. Let a be the median age of the candidates. As the set of the candidates broadcasting on these channels is a uniformly random subset of all candidates in the given round, w.v.h.p.(t), there are also $\Theta(t)$ channels on which a candidate of age at least a broadcasts. Independently, each of the $\Theta(t)$ candidates of age at most a listens with constant probability on one of these channels. Therefore, w.v.h.p.(t), a constant fraction of the candidates is eliminated in each round $r \in [r', r_0]$. For c large enough, this implies that in the $l/2$ rounds in the interval $[r', r_0]$, w.h.p., more than $8t$ candidates are eliminated, a contradiction to the assumption that the number of candidates in round r_0 is at least $16t$. □

We now can prove our main theorem.

Proof of Theorem 6.1. Analogously to the proof of Theorem 5.1 we have that at least one node v becomes a leader, i.e., it moves through $\Theta\left(\frac{\log N}{F}\right)$ phases of length $\Theta\left(F + \frac{\log N}{t}\right)$ each and is a candidate for $\Theta(t \log N)$ rounds. In total that gives a running time of

$$\Theta\left(\frac{\log N}{F}\left(F + \frac{\log N}{t}\right) + t \log N\right) = \Theta\left(t \log N + \frac{\log^2 N}{\mathcal{F}}\right).$$

Assume that besides v another node v' becomes a leader. Then v' is a candidate for $\Theta(t \log N)$ rounds in which also v is either a candidate or a leader, that broadcasts its id and age on some channel $(\mathcal{L}, s)$ with

probability $1/2$ each of those rounds. Since the number of candidates is in $O(t)$ during all those rounds, w.h.p., there is a constant probability that v broadcasts alone on an undisrupted channel in block $\mathcal{L}$. On the other hand v' listens with probability $1/2$ on block $\mathcal{L}$ and thus listens on the same channel on which v broadcasts with probability $\Omega(1/t)$, i.e., in $O(t)$ rounds there is a constant chance for v' to hear v. Thus, w.h.p., v' is eliminated within $O(t \log N)$ rounds, contradicting the initial assumption.

□

7 Improved Trapdoor Protocol

In this section, we provide a leader election protocol that is tight (within $\log\log$ factors) for $t > \mathcal{F}/6$. Our strategy is to modify the trapdoor protocol of [6] to produce a new protocol, which we call the *truncated trapdoor protocol*, that behaves in an optimal manner for t values in this range.

In the trapdoor protocol, when a node is activated, it attempts to make it through $\log N$ *phases* without being knocked out. For each round of each phase i a node chooses a channel with uniform randomness. If the node is *inactive* (i.e., has been knocked out already) it listens. Otherwise it is *active*, and it broadcasts its id and age (i.e., rounds it has been active) with probability $\frac{2^{i-1}}{N}$. If an active node receives a message from another active node of the same age or older, the node is knocked out and becomes inactive. If a node makes it through all $\log N$ phases without being knocked out, it declares itself the leader, and subsequently selects a channel at random in each round, broadcasting its id with probability $1/2$. If any node hears from a leader, it outputs the leader id and halts.

In the version of the protocol presented in [6], the first $\log N - 1$ phases were of length $\Theta\big(\frac{\mathcal{F}}{\mathcal{F}-t}\log N\big)$ and the final phase had length $O\big(\frac{\mathcal{F}t}{\mathcal{F}-t}\cdot\log N\big)$. Here, we consider a more efficient variant where each of the first $\log N - 1$ phases is reduced to length $l = O\big(\frac{\log(N)}{\mathcal{F}-t}\big)$. Below, we prove that the truncated trapdoor protocol still solves leader election. Our analysis requires that $t > \mathcal{F}/6$. We assume w.l.o.g. that $\mathcal{F} - t = O(\log N)$. For a given round let $p(v)$ be the probability that v broadcasts in that round (i.e., $p(v) = 2^{i-1}/N$ if node v is in phase i). For each round, we define $P := \sum_v p(v)$. We prove that P is bounded:

Lemma 7.1. *With high probability, at all times, it holds that $P \leq 4\mathcal{F} + 1$.*

Proof Sketch. For the sake of contradiction, assume that P exceeds $4\mathcal{F} + 1$ with probability larger than $1/N^d$, for a given constant d. In l rounds, P can only increase by a factor of 2 and by an additive amount of 1 contributed by newly activated nodes (which start in phase 1 with broadcast probability $1/N$). Hence, if P exceeds $4\mathcal{F} + 1$, it must have been between $2\mathcal{F}$ and $4\mathcal{F} + 1$ for l consecutive rounds before this point.

Consider some such round r, during which $P \in [2\mathcal{F}, 4\mathcal{F} + 1]$. Let $\hat{t} \geq \mathcal{F} - t$ be the number of non-disrupted channels in round r. We can apply Lemma 2.3 to show that, w.v.h.p.$(\mathcal{F} - t)$, in round r there are at least $(\mathcal{F} - t)/4$ channels on which the total broadcast probability of all nodes is between $2/3$ and 10. The parameters for the lemma are $k = \hat{t}$, $q = \hat{t}/\mathcal{F}$, $w_v = p(v)$ and consequently, $c \in [2, 5]$.

Now consider one such channel with a total broadcast probability between $2/3$ and 10. We can choose an age a such that at least half of the total broadcast probability comes from nodes of age at most a, and at least half of the total broadcast probability comes from nodes of age at least a. We also note that because the total broadcast probability is at least $2/3$, and because an individual node cannot broadcast with probability more than $1/2$, we know we are dealing with at least 2 nodes. With constant probability, therefore, there is exactly one node v of age at least a that transmits on our selected channel. Assume this event occurs. It follows: all nodes of age at most a (except v) receive the message and are knocked out. Therefore, with constant probability, a constant fraction of the probability mass on the channel is eliminated. This happens independently on all channels that match our broadcast probability bound. As proved above, there at least $(\mathcal{F} - t)/4$ such channels, w.v.h.p.$(\mathcal{F} - t)$.

By repeating this argument $c\log(N)/(\mathcal{F}-t)$ times, for a sufficiently large constant c, we get that with high probability, at least $2/3$ of the initial probability mass of P is eliminated and hence P becomes less than $2\mathcal{F}$. By choosing the constant c large enough, we therefore get a contradiction to the assumption that P exceeds $4\mathcal{F}+1$ with probability more than N^{-d}. □

The above lemma replaces Lemma 9 of [6], which we can then combine with the argument for Theorem 10 of that same paper to get the following:

Theorem 7.2. *With high probability, the truncated trapdoor protocol elects exactly one leader and it does so in* $O\left(\frac{\log N+\mathcal{F}t}{\mathcal{F}-t}\cdot\log N\right)$ *rounds after the first node wakes up.*

References

[1] I. 802.11. Wireless LAN MAC and Physical Layer Specifications, June 1999.

[2] Z. Alliance. Zigbee specification. *ZigBee Document 053474r06*, 1, 2005.

[3] P. Bahl, R. Chandra, T. Moscibroda, R. Murty, and M. Welsh. White Space Networking with Wi-Fi like Connectivity. In *Proceedings of the ACM SIGCOMM Conference*, 2009.

[4] Bluetooth Consortium. *Bluetooth Specification Version 2.1*, July 2007.

[5] S. Dolev, S. Gilbert, R. Guerraoui, D. R. Kowalski, C. Newport, F. Kuhn, and N. Lynch. Reliable Distributed Computing on Unreliable Radio Channels. In *the Proceedings of the 2009 MobiHoc S^3 Workshop*, 2009.

[6] S. Dolev, S. Gilbert, R. Guerraoui, F. Kuhn, and C. Newport. The wireless synchronization problem. In *Proceedings of the Principles of Distributed Computing*, pages 190–199, 2009.

[7] S. Dolev, S. Gilbert, R. Guerraoui, and C. Newport. Gossiping in a Multi-Channel Radio Network: An Oblivious Approach to Coping with Malicious Interference. In *Proceedings of the International Symposium on Distributed Computing*, 2007.

[8] S. Dolev, S. Gilbert, R. Guerraoui, and C. Newport. Secure Communication Over Radio Channels. In *Proceedings of the International Symposium on Principles of Distributed Computing*, 2008.

[9] S. Dolev, S. Gilbert, M. Khabbazian, and C. Newport. Leveraging Channel Diversity to Gain Efficiency and Robustness for Wireless Broadcast. In *Proceedings of the International Symposium on Distributed Computing*, 2011.

[10] D. Dubhashi and D. Ranjan. Balls and bins: A study in negative dependence. *Random Structures & Algorithms*, 13(2):99–124, 1998.

[11] M. Farach-Colton, R. J. Fernandes, and M. A. Mosteiro. Lower Bounds for Clear Transmissions in Radio Networks. In *Proceedings of the Latin American Symposium on Theoretical Informatics*, 2006.

[12] S. Gilbert, R. Guerraoui, D. Kowalski, and C. Newport. Interference-Resilient Information Exchange. In *the Proceedings of the Conference on Computer Communication*, 2009.

[13] K. Joag-Dev and F. Proschan. Negative association of random variables, with applications. *Annals of Statistics*, 11(1):286–295, 1983.

[14] T. Jurdzinski and G. Stachowiak. Probabilistic Algorithms for the Wakeup Problem in Single-Hop Radio Networks. In *Proceedings of the International Symposium on Algorithms and Computation*, pages 535–549, 2002.

[15] D. Meier, Y. A. Pignolet, S. Schmid, and R. Wattenhofer. Speed Dating Despite Jammers. In *Proceedings of the International Conference on Distributed Computing in Sensor Systems*, 2009.

[16] T. Moscibroda and R. Wattenhofer. Maximal independent sets in radio networks. In *Proceedings of the Symposium on the Principles of Distributed Computing (PODC)*, pages 148–157, 2005.

[17] C. Newport. *Distributed Computation on Unreliable Radio Channels*. PhD thesis, MIT, 2009.

[18] J. Schneider and R. Wattenhofer. Coloring unstructured wireless multi-hop networks. In *Proc. 28th Symp. on Principles of Distributed Computing (PODC)*, pages 210–219, 2009.

[19] M. Sherman, A. Mody, R. Martinez, C. Rodriguez, and R. Reddy. IEEE Standards Supporting Cognitive Radio and Networks, Dynamic Spectrum Access, and Coexistence. *IEEE Communications Magazine*, 46(7):72–79, 2008.

[20] M. Strasser, C. Pöpper, and S. Capkun. Efficient Uncoordinated FHSS Anti-jamming Communication. In *Proceedings International Symposium on Mobile Ad Hoc Networking and Computing*, 2009.

[21] M. Strasser, C. Pöpper, S. Capkun, and M. Cagalj. Jamming-resistant Key Establishment using Uncoordinated Frequency Hopping. In *the Proceedings of the IEEE Symposium on Security and Privacy*, 2008.

[22] R. Zhang, Y. Zhang, and X. Huang. JR-SND: Jamming-Resilient Secure Neighbor Discovery in Mobile Ad Hoc Networks. In *Proceedings of the International Conference on Distributed Computing Systems*, 2011.

Appendices

A Negatively Associated Random Variables

The notion of negatively associated random variables has been introduced in [13]. Informally, a set of random variables $\{X_1, \ldots, X_n\}$ is negatively associated if any two random variables that can be constructed by monotonic functions from disjoint subsets of $\{X_1, \ldots, X_n\}$ are negatively correlated. Formally, the concept is defined as follows.

Definition A.1 (Negative Association)**.** [13] Random variables $X_1, \ldots, X_n$ are called negatively associated iff for any two disjoint sets $I, J \subset [n]$ and for every two functions $f : \mathbb{R}^{|I|} \to \mathbb{R}$ and $g : \mathbb{R}^{|J|} \to \mathbb{R}$ such that either f and g are both componentwise non-decreasing or both componentwise non-increasing, $f(X_i, i \in I)$ and $g(X_j, j \in J)$ are negatively correlated, i.e.,

$$\mathbb{E}\big[f(X_i, i \in I) \cdot g(X_j, j \in J)\big] \leq \mathbb{E}\big[f(X_i, i \in I)\big] \cdot \mathbb{E}\big[g(X_j, j \in J)\big].$$

It is straight-forward that independent random variables $X_1, \ldots, X_n$ are negatively associated. It is also clear from the definition that the random variables of any subset of a set of negatively associated random variables are negatively associated. In [10, 13], additional basic properties w.r.t. negatively associated random variables are given.

Lemma A.1. *[10, 13] The following statements are true:*

(A) Let $X_1, \ldots, X_n$ be negatively associated random variables. For any pairwise disjoint sets $I_1, \ldots, I_k \subseteq [n]$ and for any collection of k function $f_i : \mathbb{R}^{|I_i|} \to \mathbb{R}$ so that either all f_i are non-decreasing or all f_i are non-increasing, the random variables $Z_1, \ldots, Z_k$, where $Z_i = f_i(X_j, j \in I_i)$ are negatively associated.

(B) Let $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$ be two independent collections of negatively associated random variables. Then, all the random variables $X_1, \ldots, X_n, Y_1, \ldots, Y_m$ are negatively associated.

(C) Let $X_1, \ldots, X_n$ be a set of Bernoulli random variables such that $X_i = 1$ for exactly one index $i \in [n]$. Then, $X_1, \ldots, X_n$ are negatively associated.

As shown in [10], negative association turns out to be a useful notion to study certain properties of balls-into-bins processes, as shown by the following lemma.

Lemma A.2. *Assume that m weighted balls with positive weights $w_1, \ldots, w_m > 0$ are independently thrown into n bins (each ball potentially using a different distribution). Let $a, b > 0$ be two positive constants. Further, let X_i be the total weight of balls in bin i, let Y_i be an indicator random variable for which $Y_i = 1$ iff $X_i \geq \alpha$ and let Z_i be an indicator random variable for which $Z_i = 1$ iff $X_i \leq \beta$. The random variables $X_1, \ldots, X_n$ are negatively associated. The same is true for the random variables $Y_1, \ldots, Y_n$, as well as for the random variables $Z_1, \ldots, Z_n$.*

Proof. For $(i, j) \in [n] \times [m]$, let $X_{i,j}$ be an indicator random variable for which $X_{i,j} = 1$ iff ball j lands in bin i. Because of statement (C) in Lemma A.1, for any fixed j, the random variables $X_{1,j}, \ldots, X_{n,j}$ are negatively associated. Because the random variables $X_{i,j}$ are independent for different i, by statement (B) of Lemma A.1, all the random variables $X_{i,j}$ are negatively associated. Further, because $X_i = w_1 \cdot X_{i,1} + \cdots + w_m \cdot X_{i,m}$, the random variables $X_1, \ldots, X_n$ are negatively associated by statement (A) of Lemma A.1. Finally, because Y_i can be computed from X_i by a non-decreasing function and because Z_i can be computed from X_i by a non-increasing function, also the random variables $Y_1, \ldots, Y_n$, as well as the random variables $Z_1, \ldots, Z_n$ are negatively associated. □