Leader Election in Shared Spectrum Radio Networks

Sebastian Daum  
University of Lugano  
Lugano, Switzerland  
sebastian.daum@usi.ch

Fabian Kuhn  
University of Freiburg  
Freiburg, Germany  
kuhn@informatik.uni-freiburg.de

Seth Gilbert  
NUS  
Singapore  
seth.gilbert@comp.nus.edu.sg

Calvin Newport  
Georgetown University  
Washington D.C.  
cnewport@cs.georgetown.edu

ABSTRACT
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 $F$ frequencies. To model unpredictable congestion, we assume an abstract interference adversary that can choose up to $t < F$ 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 $\Omega\left(\frac{\log^2 N}{F} + \frac{F}{\log \log N} + \frac{F}{t \log N}\right)$ rounds, through reduction to an existing result in this model [5]. 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 $O\left(\frac{\log^2 N}{F} + \log N\right)$ time. For $1 \leq t \leq F/6$, 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 $O\left(\frac{\log^2 N}{F} + t \log N\right)$ time. Finally, for $t > F/6$ we show how to improve the trapdoor protocol of [5], used to solve a similar problem in a non-optimal manner, to solve leader election in optimal $O\left(\frac{\log^2 N}{F} + \log N\right)$ time, for (only) these large values of $t$.

We also observe that if $F = \omega(1)$ and $t \leq (1 - \epsilon)F$ 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 allow 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.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

PODC’12, July 16–18, 2012, Madeira, Portugal.
Copyright 2012 ACM 978-1-4503-1450-3/12/07 ...$10.00.

Categories and Subject Descriptors
C.2.1 [Network Architecture and Design]: Wireless Communication;
F.2.2 [Analysis of Algorithms and Problem Complexity]: Non-numerical Algorithms and Problems—computations on discrete structures

General Terms
Algorithms, Theory

Keywords
wireless network, leader election, shared spectrum, disruption

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 [12], Bluetooth [2], and Zigbee [23]—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 [1]. 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 \([4\text{-}8, 11, 17, 20, 21]\). 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 \([8]\)).

**Leader Election Results.** We study leader election in the \( t \)-disrupted network model, which describes a congested single-hop synchronous radio network consisting of \( F > 0 \) 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 \( t < F \) 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 \( \Omega\left(\frac{\log^2 N}{(F-1)\log \log N} + \frac{F^2}{F-t} \log N\right) \) rounds, through reduction to our previous bound on the wireless synchronization problem in this same model \([5]\). Notice, the trapdoor protocol presented in \([5]\) can be adapted to solve leader election in \( O\left(\frac{F}{t} \log^2 N + \frac{F^2}{F-t} \log N\right) \) 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 \( O\left(\log^2 N + \log N\right) \) 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 \( F = \log N \). 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 \( 1 < t \leq F/6 \), 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 works in \( O\left(\frac{\log^2 N}{t} + t \log N\right) \) time. Finally, for \( t > F/6 \) we improve the trapdoor protocol of \([5]\) to work in \( O\left(\frac{F^2}{F-t} \log N + \log N\right) \) 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 \([5]\) 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 \( \log^2 F \) 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 that for \( t = 0 \) and \( F = \omega(1) \) 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 \([10, 14]\) (i.e., our model with \( F = 1 \) and \( t = 0 \)). Indeed, our algorithms show that this advantage can be maintained even for relatively large amounts of disruption (i.e., for \( F - t =: \epsilon F = \omega(1) \) and \( t/\epsilon = o(\log N) \)). 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 \([8]\), 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 \( F \geq 1 \) 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 \( F \) frequencies per round. By disrupting a frequency, the adversary
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 $\Pr(A) = 1 - e^{-ck}$ 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]. 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 $a > 0$ 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 $X_i > a$. It can be shown that the following lemma holds (see e.g., [9, 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 [9], 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

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

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

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

**Lemma 2.3.** Assume there are $k$ bins and $n$ balls with non-negative weights $u_1, \ldots, u_n \leq 1/4$, as well as a parameter $q \in (0, 1)$. Assume that $\sum_{i=1}^{n} u_i = c \cdot k / q$ 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 $2c$. Due to lack of space we only provide a proof sketch and refer to [3] for the full proof.

**Proof Sketch.** We introduce random variables $X_i$ that count the weight in each bin and Bernoulli random variables $Y_i$ ($Z_i$) that count the number of bins with more than $c/3$ (less than $2c$) total weight. Then we apply Chernoff bounds to the negatively associated $Y_i$ ($Z_i$). Note that $Y_i$ and $Z_i$ are not negatively associated. Careful assembling of the dependent variables $Y = \sum Y_i$ and $Z = \sum Z_i$ allows us to conclude that the number of bins, for which both conditions hold, are in $\Theta(k)$ w.v.h.p.(k).

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

**Proposition 2.4.**

1. Let $n, k \in \mathbb{N}$, $\lambda_i \in [0, \frac{1}{2}]$ for $i = 1, 2, \ldots, n$ and $\sum_{i=1}^{n} \lambda_i = 1$. Then $\sum_{i=1}^{n} \lambda_i^3 \leq \frac{1}{64}$.

2. Let $n, k \in \mathbb{N}, \lambda_i \in [0, \frac{1}{2}]$ for $i = 1, 2, \ldots, n$ and $\sum_{i=1}^{n} \lambda_i = 1$. Then $\sum_{i=1}^{n} \lambda_i e^{-\lambda_i} \leq \frac{1}{64}$.

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 $f(F, t, N)$ rounds if it guarantees the following properties are satisfied w.h.p.(N), when executed in a network with parameters $F$, $t$, $N$:

1. **Liveness:** Every node that is activated outputs the id of an active node as leader within $f(F, t, N)$ 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 [5]. An algorithm is regular if there exists a sequence of pairs $(F_1, b_1), (F_2, b_2), \ldots$, 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 $F_i$ and $b_i$, 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 $A$ be a regular algorithm that solves the leader election problem in $f(F, t, N)$ rounds. It follows that $f(F, t, N) = \Omega \left(\frac{\log^2(N)}{F \cdot t} + \frac{F \cdot t}{F - t} \cdot \log N\right)$.

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 [5], to bound such algorithms, yielding the result claimed by our theorem.

In more detail, the lemma we adapt from [5] 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 [5], 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. [10] 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 patterns, 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 $A$ solves the leader election problem with probability at least $1 - \epsilon$, then with probability at least $1 - 3\epsilon$, $A$ is vocal, and at least one node waits to output a leader until after the first message is received.

**Proof.** Assume for contradiction that $A$ 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 $3\epsilon$, $A$ 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 $B_\epsilon \subseteq B$ be the subset of strings that, when combined with activation pattern $P$, generate a silent execution. By our above assumption, $|B_\epsilon| > 3\epsilon |B|$. Let $B_\epsilon^1 \subseteq B_\epsilon$ be the subset of strings from $B_\epsilon$ 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 $A$ solving leader election with probability $1 - \epsilon$, it follows

$$|B_\epsilon^1| > 2\epsilon |B|.$$

Let $S_P$ be the set of nodes activated in $P$. Define $\ell : B_\epsilon^1 \rightarrow S_P$, such that $\forall s \in B_\epsilon^1$, $\ell(s)$ is the single node in $S_P$ to be the elected leader when we run $A$ with $s$ and activation pattern $P$. A simple counting argument tells us:

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

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

Moving on, let $B_{\epsilon, u}^1 \subseteq B_\epsilon^1$ be the strings in $B_\epsilon^1$ where $u$ is not elected leader. The key observation is that for any $s \in B_{\epsilon, u}^1$, an execution of $A$ 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:

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

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

\[\square\]

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 [5].

**Lemma 4.3 (From [5]).** Let $A$ 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 $f(\mathcal{F}, t, N) = \Omega\left(\frac{\log^2(N)}{(\mathcal{F}-t) \log \log(N)} + \frac{\mathcal{F}^2}{\mathcal{F} - t} \cdot \log(N)\right)$.

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 $O\left(\frac{\log^2(N)}{\mathcal{F}^2} + \log(N)\right)$, which is tight (up to $\log \log$ factors).

**Algorithm Description.** Detailed pseudo-code of the algorithm of this section and of Section 6 is given in Algorithm 1. For the case $t = 0$, which we analyze in this section, lines 15 and 29 can basically be omitted and in the preceding analysis we write $H$ and $\mathcal{L}$ instead of $(H, 0)$ and $(\mathcal{L}, 0)$ respectively.

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 $H$ 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 awakening, 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 $H$ (with prob $1/2$ on each of them). If $u$ does not receive any message during Phase 0, it switches to the state competing ($\mathcal{C}$), where it behaves as follows.

The algorithm acts in phases. Each phase lasts for $t := c \log N$ rounds, where $c$ is an appropriately chosen constant. There are a total of $2 \frac{\log N}{\mathcal{F}} + 1$ phases. If a node finishes its last phase while

---

\(^1\)By activation pattern, we mean the description of which nodes are activated in an execution and during which global round.
Algorithm 1:

t = 0: Basic herald algorithm for the undisrupted case
t > 0: Herald algorithm for disrupted channels


1 begin
2 \[ \text{if } t = 0 \text{ then} \]
3 \[ \text{phase-length} := c \log N + 1 \]
4 \[ \text{sub-channel-set} := \{0\} \]
5 \[ \text{else} \]
6 \[ \text{phase-length} := c(F + \log N/t) + 1 \]
7 \[ \text{sub-channel-set} := \{1, 2, \ldots, 2t\} \]
8 \[ \text{set phase} := 0; \text{count} := 0; \text{age} := 0; \text{state} := W \]
9 \[ \text{while } \text{state} \neq E \text{ do} \]
10 \[ \text{count} := \text{count} + 1; \text{age} := \text{age} + 1 \]
11 \[ \text{if } \text{count} = \text{phase-length} \text{ then} \]
12 \[ \text{count} := 1; \text{phase} := \text{phase} + 1 \]
13 \[ \text{if } \text{phase} = 1 \text{ and state} = W \text{ then } \text{state} := C \]
14 \[ \text{if } \text{phase} = \frac{2 \log N}{t} + 2 \text{ and state} = C \text{ then } \text{state} := L \]
15 Pick s uniformly at random out of sub-channel-set
16 switch state do
17 case W: With prob. 1/2, listen on channel \((H, s)\), otherwise listen on channel \((L, s)\)
18 case C:
19 \[ \text{Randomly pick } r \in [0, 1) \]
20 \[ \text{Let } I := \max \{i : r \geq (2^{i-F} \cdot 2^{(F/2)(\text{phase}-1)})/(4N)\} \]
21 \[ \text{if } I > F \text{ or } I = 0 \text{ then} \]
22 \[ \text{With prob. 1/2, listen on channel } (H, s), \text{ otherwise listen on channel } (L, s) \]
23 \[ \text{else} \]
24 \[ \text{On channel } (I, s) \text{ with probability } \pi_{R} \text{ listen or otherwise broadcast } (id, age) \]
25 case H:
26 \[ \text{Broadcast bc on channel } (H, s) \]
27 \[ \text{if } bc \neq (id, age) \text{ then } \text{state} := E \text{ else} \]
28 \[ \text{state} := C \]
29 case L:
30 \[ \text{if } t > 0 \text{ and age} > c(\log^{2} N/F + t \log N) \text{ then} \]
31 \[ \text{Consider yourself a leader} \]
32 \[ \text{On channel } (L, s) \text{ with prob. 1/2 listen, otherwise broadcast } (id, age) \]
33 \[ \text{Upon receiving a message } msg = (msg.id, msg.age): \]
34 \[ \text{if current channel/block is } H \text{ or } L \text{ then} \]
35 \[ \text{if } msg.age \geq \text{age} \text{ and msg.id} \neq \text{id} \text{ then } \text{state} := E \]
36 \[ \text{else} \]
37 \[ \text{state} := H \]
38 \[ \text{if } msg.age \geq \text{age} \text{ then} \]
39 \[ bc := (msg.age + 1, msg.id) \]
40 \[ \text{else} \]
41 \[ bc := (age + 1, id) \]

still competing, then it declares itself a leader (state \(L\)) and starts broadcasting on channel \(L\) 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 \(H\) and \(L\), which have a special role among all channels. If a competing node \(u\) chooses channels \(L\) or \(H\) then it listens in that round, if it chooses a channel \(v \in [F]\), \(u\) listens with constant probability \(\pi_{c} = \frac{1}{2}\) and transmits with probability \(1 - \pi_{c}\) (the value of \(\pi_{c}\) 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 \(\text{age}(u)\) 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:

1. First assume that a node \(u\) hears a message on channel \(H\) or \(L\). 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 \(\text{age}(v) \geq \text{age}(u)\), then \(u\) immediately considers itself eliminated (state \(E\)). Eliminated nodes only listen on channel \(L\) 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 \(H\) or \(L\), 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 \(H\). 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 \(H\). 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 \(F\) channels using the following probabilities: It selects channel \(i \in [F]\) with probability \(\frac{2^{i-F}}{F} \cdot 2^{(F/2)(\text{phase}-1)}/(4N)\) and with half of the remaining probability for channels \(H\) and \(L\) respectively. Each time a node progresses to a new phase all those probabilities (except of those for channels \(H\) and \(L\)) are multiplied by \(2^{F}\). Channels \(H\) and \(L\) always get the remaining probability. After \(\frac{2 \log N}{t} + 1\) phases, the chances for channels \(F, F - 1, F - 2, \ldots\) are \(\frac{1}{F}, \frac{1}{F+1}, \ldots\) respectively. Channel \(H\) and \(L\) are therefore both always chosen with probability more than \(\frac{1}{2}\).

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

**Theorem 5.1.** With high probability, for \(t = 0\) 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 \(p_{n}(m)\) 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., \(P_{m} := \sum p_{n}(m)\). Note that \(P_{r} = 2^{m} P_{r-1}^{m}\) and therefore \(P_{r} \approx \frac{1}{2} \sum_{m=1}^{F} P_{m}\). When making a statement about the probability mass, we refer to \(P_{r}\). 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, \(A\) 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:

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

A herald is called old, if it broadcasts on channel \( \mathcal{H} \) the id of an old node \( u \), where \( u \) is old if \( \text{age}(u) \geq A \). 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 node listens on channel \( \mathcal{H} \) at least with probability \( \frac{1}{2} \), 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 \( P_\lambda \) is in \( \left[ \frac{1}{2}, 1 \right] \). We will prove two claims:

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

\[ c_{ij} = p_i(\lambda)p_j(\lambda) \prod_{k \neq i, j} \left( 1 - p_k(\lambda) \right) \]
\[ \quad \quad \quad \quad \quad = p_i(\lambda)p_j(\lambda)e^{-\frac{2}{\lambda}P_\lambda} =: c_{ij}. \]

So the probability of having exactly two nodes on channel \( \lambda \) 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 \( P_\lambda^2e^{-\frac{2}{\lambda}P_\lambda} \), deduct \( \sum_{i<j} c_{ij} \), 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 \( \lambda \) with at least one of them being old \( D_\lambda \). Then,

\[ \mathbb{P}(D_\lambda) \geq \frac{1}{2} \left( \sum_{i,j} c_{ij} - \sum_{i<i_0} c_{ij} - \sum_{i_0<j} c_{ij} + \sum_{i_0<j, i<j} c_{ij} \right) =: c_{i_0}^2. \]

Set \( \delta_i := \frac{p_i(\lambda)}{P_\lambda} \) for \( i < i_0 \) (all old nodes) and \( \delta_{i_0} := \sum_{i=i_0}^{\infty} p_j(\lambda) \) (all young nodes together). Note that the \( \delta_i \)’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 \( \delta_{i_0} \) also ensures that \( \sum_{i=1}^{i_0} \delta_i = 1 \), which allows us to apply Proposition 2.4.2. Furthermore it holds that \( c_{ii} = \delta_i^2P_\lambda^2e^{-\frac{2}{\lambda}P_\lambda} \) and \( c_{00} = \delta_{i_0}^2P_\lambda^2e^{-\frac{2}{\lambda}P_\lambda} \). We get

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

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 \( 2\pi(1 - \pi) \geq \pi \). Since we also know that \( \mathbb{P}(D_\lambda) = \Omega(1) \), we have

\[ \mathbb{P}(A_\lambda) = \Omega(\pi). \]

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 \( \lambda \) 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 \( m \):

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

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

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

and moreover we have

\[ \mathbb{P}(u_k \text{ is not transmitting on channel } m \mid A_\lambda) = 1 - \frac{(1 - \pi) \pi_k(m)}{1 - p_k(\lambda)}. \]

Let \( B_i^m \) 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

\[ \mathbb{P}(B_i^m \mid A_\lambda) = \frac{p_i(m)}{1 - p_i(\lambda)} (1 - \pi) \frac{\pi_j(m)}{1 - p_j(\lambda)} \prod_{k \in \{1, \ldots, k - 2\}} \left( 1 - \frac{(1 - \pi) \pi_k(m)}{1 - p_k(\lambda)} \right) \]
\[ \leq \frac{p_i(m)}{1 - p_i(\lambda)} \prod_{k \in \{1, \ldots, k - 2\}} \left( 1 - \frac{(1 - \pi) \pi_k(m)}{1 - p_k(\lambda)} \right) \]
\[ \leq \frac{p_i(m)}{1 - p_i(\lambda)} \prod_{k \in \{1, \ldots, k - 2\}} \left( 1 - \frac{(1 - \pi) \pi_k(m)}{1 - p_k(\lambda)} \right) \]

Because \( B_m = \bigcup_{i,j \in \{1, \ldots, k - 2\}} B_i^m \) applying a union bound yields

\[ \mathbb{P}(B_m \mid A_\lambda) \leq \sum_{i,j \in \{1, \ldots, k - 2\}, i \neq j} \mathbb{P}(B_i^m \mid A_\lambda) \leq P_m^2 \pi \pi e^{-\frac{P_m}{2}} =: C_m. \]

For a fixed value \( m \) this value is in \( O(\pi) \).
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 v1 and v2 be elected leaders with age(v1) ≥ age(v2).

Assume first that age(v2) ≥ age(v1) + 1/2. 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_1(F) = \frac{1}{2}, i.e., P_F ≥ \frac{1}{2} for l/2 consecutive rounds, a contradiction to the second claim of Lemma 5.3. Hence, w.h.p., age(v2) < age(v1) + 1/2. But then, v1 is already a leader for l/2 rounds before any other node becomes a leader. W.h.p. this is sufficiently long for v2 to hear node v1 on channel 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 ∈ [1, ρF] nonfunctional frequencies, for some constant ρ < 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, we assume that the number of disrupted channels is t ≤ F/6. Also for simplicity, we assume that 2l divides 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 the algorithm in the case t = 0, except that each used channel in the original algorithm is replaced by a block of 2t channels with a total of F + 2 = \frac{F}{2} blocks. Similarly to Section 5, we again assume that F ≤ 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 H and L and the remaining blocks 1, 2, ..., 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 in the case t = 0. 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 the case t = 0. The second change is that for a competing node each phase now only lasts for \Theta(F + \log N) rounds. The third change is that a node that moves to state L does not immediately consider itself a leader, but it first becomes a candidate. Candidates listen or broadcast their id and age on a uniformly random channel on block 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 L can calculate itself from the age being broadcasted whether that node is already a leader or not.

Detailed pseudo-code of the algorithm is given in Algorithm 1 in Section 5.

Analysis. Analogously to section 5 we denote with P_m the sum of all nodes’ probabilities to choose block m. A and the notion of an old herald are also analogously defined. With \ell = \ell(F + \frac{\log N}{F}) 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 1 elects exactly one leader and it does so in O(\log N + t \log N) 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.h.p. (t) each round \Omega(t) heralds of age \mathcal{A} or higher are created. In Lemma 6.3 we show that also

\[ P_{F} \leq 2^{F-1} \] at all times, it also follows that \( P_{F} \geq 1/2 \) cannot hold for l/2 consecutive rounds.

Proof of Lemma 5.2. 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.

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

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

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

That is, the further a frequency is away from \( \lambda \), the smaller the probability that there is a herald created, more precisely, after some constant distance from channel \( \lambda \), 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 \([\lambda, \ldots, \lambda] \) simultaneously by \( C_\lambda + C_\lambda \). As mentioned above \( C_m = O(\pi_t) \) for a fixed \( m \). Thus, in total, we have

\[
\sum_{m \neq \lambda} \mathbb{P}(B_m | A_\lambda) \leq \sum_m C_m \leq C_\lambda + C_\lambda + \sum_{m=\lambda}^{\lambda} C_m = O(\pi_t).
\]

Choosing \( \pi_t \) small enough but still constant, then this value is less than 1. On the other hand, since we choose \( \pi_t \) to be a constant the event \( A_\lambda \) 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 \( H \) of the same or lower age as the being heralded are eliminated. By the definition of \( A \) 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 \( 2^{F} \) while the latter is an absolute increase from 0 to \( 2^{F} \). However, there are at most \( N \) nodes that can switch from the waiting state to the competing state, so that contribution is at most \( \frac{1}{2^{l/2}} \), 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 \( 2^{F} \) and an additive value of \( \frac{1}{2^{l/2}} \). 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 \( A \) and since the probability of listening on channel \( H \) is at least \( \frac{1}{2} \)). 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. □

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.
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., 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 $\Theta \left( \frac{F + \log N}{t} \right)$ rounds to guarantee a probability mass reduction of order $O(2^F)$ 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 $A$ on one of the channels in block $H$ in the following round.

**Proof.** Following the lines of the proof for Lemma 5.2 we get that there is a block $\lambda$ for which $P_A \in [2t, 4t]$. Let $\text{Dis}_{x} \subset \{1, 2, \ldots, 2t\}$ be the subset of disrupted sub-channels in block $\lambda$ and let $t_x := |\text{Dis}_{x}| \leq t$. Then define $q := 1 - \frac{t}{2t} \in (0, 1]$, $k := 2t - t_x$, $c := \frac{d_P}{2} = \frac{2}{3k} \geq 1$ 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 $\lambda$. Note that $\sum_{i=1}^{n} w_i = P_A$. We now apply Lemma 2.3, where the bins are the set of $k := 2t - t_x \geq t$ undisrupted sub-channels in block $\lambda$, i.e., the channels $[2t] \setminus \text{Dis}_{x}$. W.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 $\lambda$ is in $[1/3, 4/k]$. 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 $\lambda$ 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 the case $t = 0$ 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.$(t)$ $\Theta(t)$ heralds are created.

Due to lack of space we omit the details of this lengthy and technical proof. Here we provide a proof sketch and refer to [3] for the full proof.

**Proof Sketch.** In the case $t = 0$ we could show that on channels $i \neq \lambda$ no other herald is created with at least constant probability. We can not proceed analogously here for several reasons. First, for $t = 0$ the probability mass between consecutive channels did differ by a factor exactly $2$. Here this is still true for blocks, but it is certainly not true for different channels on the same block. Second, the probability to have no herald created outside block $\lambda$ is not a constant anymore. Third, we do not merely need a constant probability to bound the number of additional heralds, we need a very high probability that no more than $O(t)$ heralds are created.

We define the node-range $R_t := [2^{t-1} + 1, 2^t]$ and say that a channel is in node-range $i$ iff the number of nodes on it is in $R_t$. We also let $H_t$ be the number of heralds created on all channels in range $i$ and set $\gamma := \ln(1/\pi t)$. The proof is now carried out as follows:

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

(II) 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 $2^b t$ nodes.

(III) W.h.p.$(t)$, for every $i \geq 1$, there are at most $\nu t$ channels in node-range $R_t$ for a large constant $\nu$.

(IV) W.h.p.$(t)$, channels in node-ranges larger than $\log t$ do not create any heralds at all.

(V) Next we show that $\mathbb{P}(H_t \geq 2^e 1 - 2^{e+i-1} \nu t + x + 2^t) \leq e^{-x/4}$. Essentially, $x$ measures how much $H_t$ 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.h.p.$(t)$ on channels in node-range between $1$ and $\log t$, in total only $O(t)$ heralds are created.

The last step concludes the claim.

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 $P_f < 2^{F/2}$.

**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 $2^{F/2}$, $P_f$ therefore has to be at least $2t$ for $t$ 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 $P_f \in [2t, 2^F t]$ and assume that there are $t \in \{2t, 4t\}$ channels on block $H$ that are undisrupted. Let $X_1, X_2, \ldots, X_t$ be the random variables counting the number of heralds on each of the $t$ undisrupted channels in block $H$. By Lemma 2.1 the random variables $X_1, \ldots, X_t$ 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 $i \in \{t\}$ and $\alpha \geq 0$. We define $X_{\xi, \leq \alpha} := \sum_{i=1}^{t} X_{i, \leq \alpha}$ and let $p_{\leq \alpha}$ denote the probability that at most $\alpha$ heralds are on a specific single channel. Let $2\eta t$ be the number of heralds on block $H$ for this round for some $\eta = \Theta(t)$.

We then have

$$p_0 := \mathbb{P}(X_{\xi, \leq \alpha} = 0) \geq 1 - \frac{1}{2t} \left(1 - \frac{1}{2^t}\right)^{2\eta t} \geq e^{-\frac{\eta}{2}};$$

$$p_{\leq \alpha} = \mathbb{P}(X_{\xi, \leq \alpha} \leq \alpha \eta) \geq p_0 \left(1 + \frac{2\eta t}{2t - 1}\right) \geq p_0 \left(1 + \eta\right).$$

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

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

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 $\delta$ the $X_{\leq \alpha}$ are close to their expected values w.v.h.p.$(t)$.

Therefore if we choose $\delta$ small enough ($\delta < \frac{\eta}{2}$)
\[
\eta/(2\eta + 4) \), w.h.p. (t), we obtain
\[
X_{l \leq 1} - X_{l \leq 0} \geq (1 - \delta)\mu_{l \leq 1} - (1 + \delta)\mu_{l \leq 0}
\geq ((1 - \delta)(1 + \eta) - (1 + \delta)) \cdot p_0 l
\geq (\eta - \delta(2 + \eta)) \cdot p_0 l = \Theta(t).
\]

Therefore, w.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.h.p. (t), there are \( \Theta(t) \) heralds and a constant fraction of these heralds has age at least \( A \). 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 \( \gamma > 0 \) and \( \delta > 1 \), with probability \( p := 1 - e^{-\gamma t} \), a \( 1/\delta \)-fraction of the total probability mass is eliminated and let us call a round successful if an \( 1/\delta \)-fraction of the total probability mass is eliminated. In order to get from some \( P_F < tl^{2t} \) to \( P_F < 2t \) w.h.p., we need \( \Theta(F) \) successful rounds in a time span of \( t \) rounds w.h.p. (N).

If \( 1 - p = e^{-\gamma t} \) is more than \( e^{-3} \), then \( t \) is less than \( 3/\gamma = O(1) \), having a phase lasting \( \Omega(\log N) \) rounds, each being successful with a constant probability. A standard Chernoff argument gives us that \( \Theta(t) \) 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 \( t \) rounds are not successful. By choosing \( \delta = e^{1/\gamma - 1} \geq e^2 \) and letting \( X \) count the number of unsuccessful rounds we get:

\[
P(X \geq l/2) < \mathbb{P}(X \geq e^{-e^{-3t}}l) = \mathbb{P}(X \geq (1 + \delta)\mu) < e^{-\mu((\delta + 1)\ln(\delta + 1) - \delta)} \leq e^{-\mu e^{-\gamma t}} \leq e^{-t(e^{-1}(\gamma t - 1) - e^{-1} + e^{-\gamma t})} \leq e^{-\frac{\log N}{N^2}} \leq \frac{N^{-\frac{\gamma t}{2}}}{N} \leq N^{-e^{-\frac{\gamma t}{2}}}.
\]

That is, w.h.p. (N), a constant fraction of \( l \) rounds are successful. If \( e \) 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 \( 1/l^2 \).

**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 \( 1 - N^{-\delta} \), where the constant \( \delta \) 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 \( N^{-\delta} \).

We first show that between round \( r' = r_0 - 1/2 \) 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 \( r \), \( t \in [t, 2t] \) 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 \( q = t/(2t) \), \( k = \bar{k} \), and \( w_1 = 1 \) for \( i \in [n] \) which implies \( c = \frac{\bar{k}}{2t} \in [4, 12] \) we get that w.v.h.p. (t), on at least \( \bar{k}/4 \) channels, there are between \( [4/3]t = 24t \) 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 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.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 \( 1/2 \) 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 \( 16t \).

**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 \( \Theta(\log N) \) phases of length \( \Theta(F + \log N) \) each and is a candidate for \( \Theta(t \log N) \) rounds. In total that gives a running time of

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

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 \( (L, s) \) 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 \( t \) 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 \( \log \log \text{factors} \)) for \( t > F/6 \). Our strategy is to modify the trapdoor protocol of [5] 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 \( \frac{24}{N} \). 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 [5], the first $\log N - 1$ phases were of length $\Theta\left(\frac{\log N}{2^{\frac{3}{2}} - 1} \log N\right)$ and the final phase had length $O\left(\frac{\log N}{2^{\frac{3}{2}}} \log N\right)$. Here, we consider a more efficient variant where each of the first $\log N - 1$ phases is reduced to length $t = O\left(\frac{\log(N)^2}{\log(N)}\right)$. 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., $p(v) = 2^{-d} / N$ 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 $t$ rounds, $P$ can only increase by a factor of 2 by and 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 $t$ consecutive rounds before this point.

Consider some such round $r$, during which $P \in [2\mathcal{F}, 4\mathcal{F} + 1]$. Let $t \geq \mathcal{F} - t$ be the number of non-disrupted channels in round $r$. We can apply Lemma 2.3 to show that, w.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 $k = r, q = t/\mathcal{F}, w_v = p(v)$ 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.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 [5], 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 elect exactly one leader within $O\left(\frac{\log N + \mathcal{F}^2}{2^{\frac{3}{2}} - 1} \log N\right)$ rounds after the first node wakes up.

8. REFERENCES


