Feedback from Nature: Simple Randomised Distributed Algorithms for Maximal Independent Set Selection and Greedy Colouring

Peter Jeavons · **Alex Scott** · **Lei Xu**

the date of receipt and acceptance should be inserted later

Abstract We propose distributed algorithms for two well-established problems that operate efficiently under extremely harsh conditions. Our algorithms achieve state-of-the-art performance in a simple and novel way.

Our algorithm for maximal independent set selection operates on a network of identical anonymous processors. The processor at each node has no prior information about the network. At each time step, each node can only broadcast a single bit to all its neighbours, or remain silent. Each node can detect whether one or more neighbours have broadcast, but cannot tell how many of its neighbours have broadcast, or which ones.

We build on recent work of Afek et al. which was inspired by studying the development of a network of cells in the fruit fly [2]. However we incorporate for the first time another important feature of the biological system: varying the probability value used at each node based on local feedback from neighbouring nodes. Given any n-node network, our algorithm achieves with high probability the optimal time complexity of $O(\log n)$ rounds and the optimal expected message complexity of $O(1)$ single-bit messages broadcast by each node. We also show that the previous approach, without feedback, cannot achieve better than $\Omega(\log^2 n)$ time complexity with high probability, whatever global scheme is used to choose the probabilities.

Our algorithm for distributed greedy colouring works under similar harsh conditions: each identical node has no prior information about the network, can only broadcast a single message to all neighbours at each time step representing a desired colour, and can only detect whether at least one neighbour has broadcast each colour value. We show that with high probability our algorithm has a time complexity of $O(\Delta + \log n)$, where Δ is the maximum degree of the network, and also has an expected message complexity of $O(1)$ messages broadcast by each node.

P. Jeavons, Lei Xu
Dept. of Computer Science, University of Oxford
E-mail: Peter.Jeavons@cs.ox.ac.uk

A. Scott
Mathematical Institute, University of Oxford
E-mail: scott@maths.ox.ac.uk

1 Introduction

One of the most fundamental problems in distributed computing is to distributively choose a set of local leaders in a network of connected processors so that every processor is either a leader or connected to a leader, and no two leaders are connected to each other. This problem is known as the distributed *maximal independent set* (MIS) selection problem and has been considered as a challenging problem for three decades [2]. It has many applications, especially in wireless networks [25,46] and has been extensively studied [36,35,3,27,29,41,40].

Another fundamental problem in distributed computing that is closely related to the distributed MIS selection problem is the $(\Delta+1)$-*colouring* problem. In this problem the aim is to colour the vertices of a graph which has maximum degree Δ using no more than $\Delta+1$ colours so that adjacent vertices are assigned different colours. Like the distributed MIS selection problem, the distributed $(\Delta+1)$-colouring problem also serves as a basic building block in many other distributed algorithms, and has many applications for resource assignment, in particular for frequency assignment in radio-communication networks [38,51,42,45]. Because of this, it has also been extensively studied [4,6,26,44,49, 22,43,20,19].

A more restricted variant of the colouring problem is called *greedy colouring* [17,13], where the aim is to obtain a colouring with the property that no individual vertex can

be recoloured with a smaller colour (in some fixed ordering of the colours). Computing a greedy colouring distributively is believed to be more difficult than computing an arbitrary $(\Delta+1)$-colouring distributively [13], but such colourings often use a much smaller number of colours [54].

1.1 Our Results

In this paper, we first propose a randomised distributed MIS selection algorithm that is able to operate under very harsh conditions. Our model of distributed computing assumes an identical anonymous processor at each node that has *no information about the network*. At each time step, each node can only *broadcast a single bit* to all its neighbours, or remain silent. Each node can detect whether one or more neighbours have broadcast, but cannot tell how many neighbours have broadcast, *or which ones*.

We prove that our algorithm is optimal in both time and bit complexity for such a model, by showing that it runs in expected $O(\log n)$ time, where n is the number of nodes, and the expected number of messages sent by each node is bounded by a constant, regardless of the network.

We then extend the approach to obtain an algorithm for the distributed greedy colouring problem. This algorithm also runs under very harsh conditions where the processors are anonymous and have no information about the network. For this problem we allow each node to broadcast only a single message to all neighbours at each time step representing a single desired colour value. Once again nodes can only detect whether at least one neighbour has broadcast a colour, and cannot tell how many neighbours have broadcast, or which ones.

The algorithm we obtain is remarkably simple and computes a greedy colouring in expected $O(\Delta+\log n)$ time, where n is the number of nodes and Δ is the maximum degree of the network. Once again the expected total number of messages sent by each processor is bounded by a constant. As well as matching the best known time complexity for obtaining a greedy colouring, our algorithm is the first proposed algorithm for this problem where the nodes require no prior knowledge of the network and cannot distinguish between their neighbours.

To obtain our results we introduce a new form of analysis to determine the time complexity. Most previous analysis has relied on a general technique, originally devised by Luby [36], which divides the computation into successive phases and shows that some fixed fraction of the network is expected to be eliminated in each phase, so that there are at most logarithmically many phases. Our algorithms do not have this property, and hence require a more flexible form of analysis, which we describe in detail below.

2 Preliminaries

Given an undirected graph $G=(V,E)$, the *neighbourhood* of each vertex $v \in V$ is defined to be the set $\Gamma(v)=\{u:\{u,v\}\in E\}$ and the *degree* of each vertex v is defined to be the number $\deg_G(v)=|\Gamma(v)|$. We define the maximum degree of the graph G to be the maximum value of the degree over all vertices of G, which is denoted by $\Delta=\max_{v\in V}\{\deg_G(v)\}$. The number of vertices of G is $|V|$ and will usually be denoted by n. We will say that an event on G occurs *with high probability* if it occurs with probability at least $1-O(1/n^c)$ for some $c\geq 1$. (Note that this is stronger than simply requiring that the probability tends to 1 as n tends to infinity.) We will write $\log a$ for the natural logarithm of a, and $\log_b a$ for the logarithm of a to the base b.

2.1 Maximal Independent Set Selection

Definition 1 (Maximal Independent Set) Given an undirected graph $G=(V,E)$, an *independent set* in G is a subset of vertices $U\subseteq V$, such that no two vertices in U are adjacent. An independent set U is called a *maximal independent set* (MIS) if no further vertex can be added to U without violating independence.

Different maximal independent sets for the same network can vary greatly in size. In contrast to the MIS selection problem, the related problem of finding a *maximum size* independent set (MaxIS) is notoriously hard. It is equivalent to finding a maximum clique in the complementary graph, and is therefore **NP**-hard [23]. However, computing an arbitrary MIS (which is not necessarily of the maximum possible size) in linear time using a centralised sequential algorithm is trivial: simply scan the nodes in arbitrary order. If a node u does not violate independence, add u to the MIS. If u violates independence, discard it. Hence the challenge is to compute such an MIS more efficiently in a distributed way with no centralised control.

2.2 Greedy Colouring

A proper colouring of a graph assigns a colour to each vertex such that no two adjacent vertices are assigned the same colour. The colouring is called a k-colouring if at most k different colours are used.

Definition 2 (Graph Colouring) For any undirected graph $G=(V,E)$, a *k-colouring* of G is a function f from the vertices V to a set of colours $\{c_1,c_2,\ldots,c_k\}$ such that $f(u)\neq f(v)$ for every edge $\{u,v\}\in E$. G is called *k-colourable* if and only if there exists a k-colouring of G.

For many practical applications it is desirable to minimize the number of colours used. The smallest possible positive integer k for which there exists a k-colouring of G is defined to be the *chromatic number* χ of G. It is known to be **NP**-hard to approximate the chromatic number χ within a factor of $|V|^{1-\varepsilon}$, for any $\varepsilon > 0$, even using a centralised algorithm with complete knowledge of the graph [55].

However, a number of heuristic approaches can be used to rapidly obtain colourings with a reasonably low number of colours on many graphs. For example, the following greedy approach produces a colouring in linear time using a centralised control.

Definition 3 (Greedy Colouring) Given an arbitrary ordering, $(v_1, v_2, \ldots, v_n)$, of the vertices of G, and an arbitrary ordering on the colour values, a *greedy colouring* algorithm considers each vertex from v_1 to v_n in turn, assigning each vertex the smallest possible colour value that is not already assigned to any of its neighbours.

Note that a colouring obtained in this way has the property that no individual vertex can be recoloured using a smaller colour. A colouring with this property is sometimes called a *Grundy colouring* [13,17]. Since every greedy colouring algorithm produces a Grundy colouring, and every Grundy colouring can be obtained using a greedy colouring algorithm (by choosing a suitable ordering on the vertices) [13], we will refer to any Grundy colouring as a greedy colouring, even if it is computed in some other way.

It is easy to see that a greedy colouring uses no more than $\Delta + 1$ colours, so we have that $\chi \leq \Delta + 1$ for any graph G. Brooks Theorem strengthens this observation by stating that Δ colours suffice for all graphs except odd cycles and complete graphs, which require $\Delta + 1$ colours.

2.3 Distributed Computation Model

In the widely-used Linial model [9,33,35,13], a distributed network is composed of a set V of processors and a set E of bidirectional communication links (channels) between pairs of processors. If there is a link (channel) between two processors, these two processors are said to be *neighbours*. A distributed network with n processors where each processor has no more than Δ neighbours corresponds to an undirected graph $G = (V, E)$ with n vertices and maximum degree Δ. A network is called *anonymous* if the processors cannot distinguish each other by unique identifiers. Linial's distributed computation model is a synchronous system and all processors operate in a lockstep fashion as we now describe. We assume that all processors wake up and start their computation at the same time step. During each time step, all processors act in parallel and carry out the following operations sequentially [37]:

1. Optionally send a message to each neighbouring node;
2. Receive any messages sent by neighbours;
3. Perform arbitrary local computation.

The computation is said to be complete only when the local computations at every vertex have terminated.

The distributed computation model we use is based on this model, but we impose the following severe additional conditions:

1. Each processor is anonymous and has no local or global information about the network;
2. At each time step, each processor either keeps silent or broadcasts one message to all its neighbours;
3. Each processor can tell whether at least one neighbour has broadcast a message, but cannot tell how many of them have done so, or which ones.

In our MIS algorithm we restrict the communication even further, so that each message contains only a single bit. In our greedy colouring algorithm we allow longer messages representing different colours.

Information about a network may be difficult to obtain, or subject to uncertainty and change, so it is desirable for some applications to find algorithms that can complete their task without using such information [18,47]. Moreover, using a small number of messages, each containing a single bit (or a small number of bits), allows an implementation to use less communication resources and less energy, and this may be crucial in some applications [32]. Because of the restrictions we impose, our algorithms can be implemented using very simple communication mechanisms such as radio waves, optical signals, or even chemical signals, as in biological intercellular signalling [8,10].

3 Related Results

3.1 Distributed MIS selection

The study of distributed MIS selection can be traced back to the 1980s. It was shown early on that the MIS selection problem is in the complexity class **NC** [24], and hence likely to be a good candidate for a parallel or distributed approach.

We review the current state-of-the-art here, focusing on the size of the messages (in bits) and the information about the network that is used at each node (see Table 1). In many cases it is possible to use estimates for the required graph parameters, and to iteratively refine these, at the cost of a more sophisticated algorithm and additional communication rounds, but we describe only the simplest versions of the algorithms, as originally presented.

A lower bound of $\Omega(\log^* n)$ time for distributed MIS selection on graphs with $\Delta \geq 2$ is given in [34]. The most well-known lower bound for distributed MIS selection on general

Table 1 Distributed MIS selection algorithms on general graphs with n nodes and maximal degree Δ

Type	Time Steps	Message size (bits)	Information about the graph and neighbourhood used at each node	Reference
Det.	$O(\Delta+\log^* n)$	$\Omega(\log n)$	Unique IDs, size and maximum degree of the graph, and distinguishable channels	[4]
				[26]
	$O(2^{O(\sqrt{\log n})})$			[44]
Rand.	$O(\log^2 n)$	$\Omega(\log\Delta)$	Maximum degree in 2-neighbourhood	[46]
		3	None	[12]
		1	Size of the graph	[2]
		1	None	[1]
	$O(\log\Delta\sqrt{\log n})$	$\Omega(\log n)$	Size, maximum degree of the graph, and distinguishable channels	[7]
	$O(\log n)$	$\Omega(\log n)$	Size of the graph	[37]
		$\Omega(\log\Delta)$	Degrees of neighbours	[52]
		1	Distinguishable channels	[40]
		1	**None**	**This paper**

graphs, $\Omega(\sqrt{\log n/(\log\log n)})$, is given in [28]. This was improved to $\Omega(\min\{\log\Delta,\sqrt{\log n}\})$ in [30] (see also [7]). All of these lower bounds have been shown to apply to both deterministic and randomised algorithms. It was observed in [40] that if only one-bit messages are allowed to be sent along each edge in any time step, then every distributed algorithm to select an MIS in a ring of size n requires at least $\Omega(\log n)$ time steps with high probability.

For deterministic distributed MIS selection on general graphs, the fastest known algorithms run in $O(\Delta+\log^* n)$ time [26,4] or $O(2^{O(\sqrt{\log n})})$ time [44]. These deterministic MIS algorithms rely on very sophisticated multi-phase techniques, use a considerable amount of global information about the graph at each node, including unique node IDs, and allow complex messages to be sent on specific channels between nodes. Note that any deterministic algorithm requires some information at each node (such as a unique node ID) in order to break the symmetry [21,46].

Using randomisation to break symmetry between nodes allows for simpler algorithms, often requiring a smaller number of time steps. A simple parallel randomised algorithm for distributed MIS selection in the PRAM model of computation was presented in 1986 by Luby [36] and independently by Alon et al. [3].

This algorithm has been adapted to the message-passing model of distributed computation in several slightly different ways. In the version presented by Lynch [37] each processor is assumed to know the total size, n, of the graph, and chooses a random integer in the range 1 to n^4 at each time step. These integers are then broadcast as messages to all neighbouring nodes, so the messages sent between processors contain $\Omega(\log n)$ bits. Using these messages the nodes are able to compute an MIS by selecting the nodes that choose the largest random values in their neighbourhood, removing those nodes and their neighbours, and iterating this process. Using the analysis from [36], this process is shown to terminate in $O(\log n)$ time on average and with high probability.

In the version presented by Wattenhofer [52] the nodes choose a probability value based on their degree in the graph, and use this value, together with the degree values of their neighbours to decide whether to join the MIS at each time step. In this variant the nodes exchange messages to determine the current degrees of their neighbours at each time step, and hence the messages sent between processors contain $\Omega(\log\Delta)$ bits. Once again, using the analysis from [36], this process is shown to terminate in $O(\log n)$ time on average and with high probability.

In the version presented by Peleg [46] the probability value at each node is chosen based on the maximum degree of the nodes at distance 1 or 2 away from it in the graph, and this value is then used to decide whether to join the MIS at each time step. Peleg shows with a simpler analysis that this algorithm halts in $O(\log^2 n)$ time on average and with high probability. Once again the nodes exchange messages to determine the current degrees of their neighbours at each time step, and hence the messages sent between processors contain $\Omega(\log\Delta)$ bits.

These distributed randomised algorithms, all based on a similar approach and generally known as Luby's algorithm, remained the state-of-the-art for many years, but there have recently been some new developments.

A new randomised MIS algorithm with time complexity $O(\log\Delta\sqrt{\log n})$ was proposed in [7]. This algorithm improves on the $O(\log n)$ algorithms when $\log\Delta<\sqrt{\log n}$. On the other hand, this algorithm assumes that each processor knows the size and maximal degree of the graph and can distinguish between channels so that it can send different messages along different edges. Since it relies on exchanging information about specific nodes, using node identities, the messages exchanged in this algorithm contain $\Omega(\log n)$ bits.

Algorithms for MIS selection on special graphs such as sparse graphs and growth-bounded graphs have also been studied [15,5,48].

3.2 MIS Selection with Limited Communication

There has recently been considerable interest in finding efficient distributed MIS selection algorithms that can work in more restricted computational models, such as wireless network models [41,12,11,2,1,40,53].

For example, the approach proposed in [40] splits the randomly generated values at each node into single bits, and communicates them one by one. When these bits are broadcast to all neighbours, this approach achieves a time complexity of $O(\log^2 n)$. By distinguishing between different neighbours, and having separate, overlapping, exchanges of messages with each neighbour, the overall time complexity is brought down to $O(\log n)$ time on average and with high probability. This is shown to be the optimal time complexity that can be achieved with one-bit messages [40]. However, to achieve this optimal performance requires that each vertex can distinguish between its neighbours by locally known channel names, so that different messages can be sent along different edges at the same time step.

A more radical approach is the novel distributed MIS selection algorithm inspired by the neurological development of the fruit fly which is given in [2,1].

During development, certain cells in the pre-neural clusters of the fruit fly specialise to become sensory organ precursor (SOP) cells, which later develop into cells attached to small bristles (microchaetes) on the fly that are used to sense the environment. During the first stage of this developmental process each cell either becomes an SOP or a neighbour of an SOP, and no two SOPs are neighbours. These observed conditions are identical to the formal requirements in the maximal independent set selection problem.

However, Afek *et al.* pointed out [2] that the method used by the fly to select the SOPs appears to be rather different from the standard algorithms for choosing an MIS described above. The cells of the fly appear to solve the problem using only simple local interactions between certain membrane-bound proteins, notably the proteins Notch and Delta [8,10]. Moreover, they require very little knowledge about connectivity. Based on their study of this developmental process, Afek *et al.* proposed an algorithm that works in a distributed model where each node can only broadcast to all its neighbours or remain silent. Moreover, each node can only detect whether at least one neighbour has broadcast a signal. This model of communication is sometimes referred to as a "beeping" model with collision detection [1].

In their proposed algorithm, each node broadcasts at each time step with a certain probability, which changes over time, and then checks whether any of its neighbours has broadcast at the same time. As originally presented [2], the algorithm uses a sequence of gradually increasing global probability values calculated from the total number of nodes of the graph n and its maximum degree Δ. The algorithm was further refined by Afek *et al.* in a later paper [1]. In the later version the probability values are chosen according to a fixed pattern, so that the individual nodes require no information at all about the graph. However, in both versions the expected number of time steps required was shown to be $O(\log^2 n)$ (see Table 1).

Another approach to distributed computing with very restricted communication and processing capabilities is the networked finite state machine model introduced in [12]. This only allows a fixed finite number of distinct messages, and very limited computation at each node, based on the notion of a randomised finite state machine, with no information about the network. It is shown in [12] that a MIS can be computed in this very restricted model in $O(\log^2 n)$ time, using only 7 states and 7 corresponding messages.

3.3 Distributed Colouring

The problem of $(\Delta+1)$-colouring is closely related to MIS selection [31]. Hence it can be shown that the lower bounds for distributed MIS selection mentioned above also apply for distributed $(\Delta+1)$-colouring. Similarly, some state-of-the-art distributed $(\Delta+1)$-colouring algorithms are closely related to the algorithms for distributed MIS selection described earlier (see Table 2).

A randomised distributed $(\Delta+1)$-colouring algorithm requiring $O(\log\Delta+\sqrt{\log n})$ time is proposed in [49]. In this algorithm the messages represent randomised preference levels for each of the possible colours. This algorithm needs to know an upper bound of the size of the graph and requires each processor to be able to send different messages along different channels. Since the messages exchanged represent colours, the message size is at least $\Omega(\log\Delta)$.

A randomised algorithm wwhich achieves a time complexity of $O(\log\Delta+2^{O(\sqrt{\log\log n})})$ with high probability is given in [7]. However, this algorithm relies on a deterministic algorithm to complete a partial colouring, and hence requires unique node IDs, and messages with $\Omega(\log n)$ bits.

Johansson proposed and analysed a simple randomised distributed $(\Delta+1)$-colouring algorithm requiring $O(\log n)$ time [22]. The algorithm of Johansson requires that each vertex knows the maximum degree of the graph. Each message corresponds to a potential colour choice, so the messages in this algorithm contain $\Omega(\log\Delta)$ bits.

These algorithms do not attempt to obtain *greedy* colourings, and hence tend to use the maximum number, $\Delta+1$, of colours. For many classes of graphs, a greedy colouring will

Table 2 Distributed $(\Delta+1)$-colouring algorithms on general graphs with n nodes and maximal degree Δ

Greediness	Type	Time Steps	Message size (bits)	Information used at each node	Reference
Non-greedy	Det.	$O(\Delta+\log^* n)$	$\Omega(\log n)$	Unique IDs, size and maximum degree of the graph, and distinguishable channels	[4]
					[26]
		$O(2^{O(\sqrt{\log n})})$			[44]
	Rand.	$O(\log\Delta+\sqrt{\log n})$	$\Omega(\log\Delta)$	Upper bound on the size of the graph and distinguishable channels	[49]
		$O(\log\Delta+2^{O(\sqrt{\log\log n})})$	$\Omega(\log n)$	Unique IDs, size and maximum degree of the graph, and distinguishable channels	[7]
		$O(\log n)$	$\Omega(\log n)$	Size of the graph	[37]
			$\Omega(\log\Delta)$	Degrees of neighbours	[52]
				Maximum degree of the graph	[22]
Greedy	Det.	$O(\Delta^2+\log^* n)$	$\Omega(\log n)$	Unique IDs, maximum degree of the graph, own degree and distinguishable channels	[43]
		$O(\Delta+\log^* n)$			[4] + [13]
	Rand.	$O(\Delta^2\log n)$	$\Omega(\log\Delta)$	Own degree and degrees of neighbours	[19]
		$O(\Delta+\log n)$	$\Omega(\log\Delta)$	Maximum degree of the graph	[13]
			$O(\log\mu)$ (where $\mu=$ max colour used)	Distinguishable channels	[39]
				None	**This paper**

often use considerably fewer colours. In fact, we showed in an earlier paper that a simple greedy colouring algorithm can produce colourings that use an optimal, or near-optimal number of colours on many standard graph colouring benchmarks [54]. However, computing a greedy colouring with a distributed algorithm is a more challenging problem. In fact, the problem of computing a greedy colouring for a given ordering of the vertices is known to be **P**-complete [16, 14].

Panconesi and Rizzi proposed a deterministic algorithm for graph colouring that attempts to use a small number of colours [43]. This algorithm was not originally designed to construct a greedy colouring. However, it can be easily modified to become a distributed greedy colouring algorithm by always choosing the first available colour when assigning a colour. In view of this it is described in [13] as the first distributed approach to greedy colouring. The number of time steps taken by this algorithm is $O(\Delta^2+\log^* n)$ [43]. The algorithm is divided into distinct phases, where the first is a preprocessing phase which produces a forest decomposition of the graph. The algorithm assumes that each vertex has a unique identifier, and it also requires that each vertex knows its own degree and the maximum degree of the whole graph. This algorithm also requires the ability to send different messages to different neighbours simultaneously. Because identifiers are exchanged the message size of the algorithm is $\Omega(\log n)$.

Hansen et al. proposed a randomised distributed algorithm for graph colouring in [19]. Even though the algorithm is not explicitly described in the original paper as a greedy colouring algorithm, it is pointed out in [13] that the colourings it produces are actually greedy colourings. The expected number of time steps taken by this algorithm to produce a colouring is $O(\Delta^2\log n)$ [19]. However, this algorithm assumes that each vertex knows its degree in the graph, and the messages exchanged include these numerical degree values as well as the colour values. Hence the size of each message sent is at least $\Omega(\log\Delta)$ bits.

Gavoille et al. give a detailed theoretical study of distributed greedy colouring [13]. They establish a lower bound for this problem of $\Omega(\log n/\log\log n)$ time steps. Moreover, they note that an arbitrary k-colouring can be converted to a greedy colouring by a simple distributed algorithm in $O(k)$ time steps. However, the conversion algorithm in [13] requires each node to know the value of k. Combining this approach with the most efficient randomised $(\Delta+1)$-colouring algorithms described earlier gives a two-stage algorithm with an overall expected time complexity of $O(\Delta+\log n)$. Similarly, combining with the most efficient deterministic $(\Delta+1)$-colouring algorithm described earlier gives a two-stage algorithm with an overall time complexity of $O(\Delta+\log n)$.

Métivier et al. proposed a simple randomised distributed $(\Delta+1)$-colouring algorithm requiring $O(\Delta+\log n)$ time [39]. The algorithm proposed by Métivier et al., does not assume any global knowledge of the network, but requires each processor to know from which channel it receives each message. The algorithm consists of two stages: it first uses randomisation to break the symmetry and obtains a colouring in $O(\log n)$ time with an unbounded number of colours; it then reduces the number of colours used to at most $\Delta+1$ in $O(\Delta+\log n)$ time. The colours in this second stage are cho-

ALGORITHM 1: The algorithm for distributed MIS selection at each node

Global constants: p_0 : lower bound on initial probability value;
f_1, f_2 : lower and upper bounds on change factor for probability value.

Local variables: p : local probability value, initialised to some value in $[p_0, 1]$;
f : change factor for probability value, chosen arbitrarily in $[f_1, f_2]$;
TRYING : Boolean flag, initialised to FALSE.

1. **while** active, in each round **do**
2. *FIRST EXCHANGE*
3. With probability p, set TRYING $\leftarrow$ TRUE and **send** signal to all neighbours;
4. **Receive** any signals sent by neighbours;
5. Set f to some arbitrary value in the interval $[f_1, f_2]$;
6. **if** any signal was received **then**
7. TRYING $\leftarrow$ FALSE and $p \leftarrow p/f$ (decrease p)
8. **else**
9. $p \leftarrow \min\{fp, 1\}$ (increase p)
10. *SECOND EXCHANGE*
11. **if** TRYING **then**
12. **Send** signal to all neighbours;
13. Join the MIS and terminate (become inactive).
14. **Receive** any signals sent by neighbours;
15. **if** any signal was received **then**
16. Terminate (become inactive)

sen to be the smallest available, so the resulting colouring is a greedy colouring (although this is not made explicit). Each message exchanged in the first stage contains only one bit, but each message in the second stage represents a final colour choice. Since the total number of colours used may be much lower than Δ in some classes of graphs, we give an upper bound on the message size of $(\log \mu)$ bits for this algorithm, where μ is the maximum colour number used.

4 Algorithm for MIS Selection

The distributed algorithm for MIS selection proposed by Afek *et al.* is remarkably simple [2]. At each step, each node may choose, with a certain probability p (that varies over time), to signal to all its neighbours that it wishes to join the independent set. If a node chooses to issue this signal, and none of its neighbours choose to do so in the same time step, then it successfully joins the independent set, and becomes inactive, along with all its immediate neighbours. However if any of these neighbouring nodes issue the same signal at the same time step, then the node does not succeed in joining the independent set at that step. This process is repeated until all nodes become inactive.

Our new algorithm uses a similar basic scheme, but with a different approach to the way that the probability value p varies over time (see Algorithm 1). Inspired by the positive feedback mechanisms that control cellular processes [8, 10], we give each node an independently updated probability value. These probabilities are initialised to arbitrary values (above some fixed threshold value, $p_0 > 0$). They are decreased whenever one or more neighbouring nodes signal that they wish to join the independent set, and are increased whenever no neighbouring node issues such a signal. We allow each increase or decrease to be by some arbitrary factor f, which may vary at each step, but is bounded by the global parameters f_1 and f_2 (with $1 < f_1 \leq f_2$).

Our main result below shows that varying the probabilities in this way, using a simple local feedback mechanism, gives an algorithm whose expected time to compute a maximal independent set is $O(\log n)$ (see Corollary 1, below). We also show that the expected number of signals sent by each node is bounded by a constant (see Theorem 3, below).

Note that Algorithm 1 consists of two successive message exchanges. We shall refer to each such pair of message exchanges as a *round* of the algorithm. Hence each round occupies two consecutive time steps.

To investigate the performance of our new algorithm in practice we constructed an implementation with the probability p at each node varying as follows: p is initially set to $1/2$. In any round where a signal is received from at least one neighbouring cell the value of p is halved. In all other rounds it is doubled (up to a maximum of 1). We then compared this algorithm with the algorithm in [1] by running both of them on random networks with different numbers of nodes, where each edge is present with probability $1/2$ [50].

We found that the mean number of rounds required in our experiments to complete our algorithm and choose a maximal independent set in these networks was approximately $2.5 \log_2 n$, for all values of n between 20 and 200. However, the mean number of rounds required by the algo-

rithm in [1] to select a maximal independent set was close to the exact value of $\log_2^2 n$. Afek et al. do not discuss the expected number of signals broadcast at each node in their algorithm. We found that the mean number of signals sent by each node in our algorithm was less than 2, regardless of the size of the network. However, our experiments indicated that the mean number of signals sent by each node when running the algorithm described in [1] increased with the size of the network.

Before we analyse the performance of this algorithm we first demonstrate in Section 4.1 that the use of a feedback mechanism to adjust the probability values, as described in lines 5-9 of Algorithm 1, is crucial to achieving the efficiency.

4.1 Lower Bound for Globally Chosen Probability Values

In this section we consider a class of algorithms similar to the one described in [2] where each node runs through the same fixed preset sequence of probability values, and does not adjust these to take into account the behaviour of other nodes. In other words, we consider a simplified version of Algorithm 1, where the probability values at all nodes are initialised to the same value p_0, and the probability updates described in lines 5-9 are replaced by a simple update rule that changes p to the next value in some fixed sequence $p_1, p_2, \ldots$. We refer to this modified algorithm as *MIS selection with global probability values*.

Our first result constructs an explicit family of graphs with $O(n)$ vertices, for which *any* such algorithm takes at least $\Omega(\log^2 n)$ rounds, no matter what sequence of probability values is used. (Note that we generally omit floors and ceilings for clarity, and the graphs we construct in this result have $O(n)$ vertices rather than exactly n vertices, to simplify their description.)

Theorem 1 *There is a constant $\kappa > 0$ such that the following holds. Let G be the graph consisting of $n^{1/3}$ disjoint copies of the complete graph K_d, for each $d = 1, \ldots, n^{1/3}$. Then with high probability, any MIS selection algorithm with global probability values running on G does not terminate within $\kappa \log^2 n$ rounds.*

Proof Let $p_0, p_1, p_2, \ldots$, be the sequence of probability values used by the algorithm. Fix d, and consider a copy K of K_d. The probability that some vertex of K is added to the independent set at the ith round is the probability that exactly one vertex of K beeps, and so equals

$$dp_i(1-p_i)^{d-1} \leq dp_i \exp(-(d-1)p_i). \tag{1}$$

Note that the function xe^{-x} is bounded on $[0,\infty)$, and has maximum $1/e$ (at $x = 1$). So for $d > 2$,

$$dp_i \exp(-(d-1)p_i) = \frac{d}{d-1} \cdot (d-1)p_i \exp(-(d-1)p_i) \leq \frac{3}{2e}.$$

Also, for $x \in [0, 3/2e]$, we have $1 - x \geq \exp(-2x)$. So, by inequality (1), the probability that all the vertices of K are still active after T rounds is at least

$$\begin{aligned}\prod_{i=1}^{T} \left(1 - dp_i e^{-(d-1)p_i}\right) &\geq \prod_{i=1}^{T} \exp(-2dp_i e^{-(d-1)p_i}) \\ &= \exp(-\sum_{i=1}^{T} 2dp_i e^{-(d-1)p_i}) \\ &\geq \exp(-\sum_{i=1}^{T} 6dp_i e^{-dp_i}).\end{aligned}$$

The last inequality follows from the fact that $e^{p_i} \leq e \leq 3$.

Hence if $\sum_{i=1}^{T} 6dp_i e^{-dp_i} < \frac{1}{4}\log n$ then the nodes of K remain active with probability at least $n^{-1/4}$. In that case the probability that the nodes in all the copies of K_d become inactive in T rounds is at most

$$(1 - n^{-1/4})^{n^{1/3}} \leq \exp(-n^{1/12}),$$

and so with high probability the algorithm fails to terminate in T rounds.

It follows that we need only consider the case when

$$\sum_{i=1}^{T} 6dp_i e^{-dp_i} > \frac{1}{4}\log n$$

for every choice of $d \geq 3$. We will show that this implies $T = \Omega(\log^2 n)$.

Let us choose d at random. We define a probability distribution for d by

$$\mathbb{P}[d = j] = \frac{c}{j \log n},$$

for $j = 3, \ldots, n^{1/3}$ (where c is a normalizing constant: note that $c = \Theta(1)$, as $\sum_{i=1}^{n^{1/3}} 1/j = \Theta(\log n)$). Then, for any $p \in [0,1]$,

$$\mathbb{E}[dpe^{-dp}] = \sum_{j=3}^{n^{1/3}} \frac{c}{j \log n} jpe^{-jp} \leq \frac{c}{\log n} \sum_{j=0}^{\infty} pe^{-jp}.$$

But $\sum_{j=0}^{\infty} pe^{-jp} = p/(1-e^{-p}) < 2$, as $p \in [0,1]$; so we have $\mathbb{E}[dpe^{-dp}] < 2c/\log n$. By linearity of expectation, choosing a random d, we have

$$\mathbb{E}\left[\sum_{i=1}^{T} 6dp_i e^{-dp_i}\right] < 12cT/\log n.$$

Hence there is some value of d for which

$$\sum_{i=1}^{T} 6dp_i e^{-dp_i} < 12cT/\log n.$$

By the argument above, this quantity must be at least $\frac{1}{4}\log n$, and so we must have $T = \Omega(\log^2 n)$. □

4.2 Time Complexity with Locally Chosen Probability Values and Feedback

In this section we analyse the running time of our new algorithm for distributed MIS selection (Algorithm 1, where the probability values at each node are locally varied in each round based on feedback from neighbouring nodes.

It follows from the analysis of [2] that if this algorithm terminates (i.e., all nodes become inactive) then it correctly identifies an MIS. The only question is the number of rounds required.

Note that, unlike Luby's algorithm [3,36], it is not true that in every round we can expect at least some constant fraction of the edges to be incident to nodes that become inactive in that round. For example, in a complete graph nodes will only become inactive when exactly one node signals. If all nodes are initialised with the same probability value and with the same (fixed) increase and decrease factor f, then all nodes will always possess the same probability value p_t. Whenever more than one node signals, all nodes will decrease their probabilities by f; if no node signals, all nodes will increase their probabilities by f. The probability of exactly one node signalling is thus $np_t(1-p_t)^{n-1}$ at each round t. Hence, for complete graphs, with high probability all nodes will remain active for any fixed constant number of rounds. It follows that we must carry out a more detailed analysis over a sequence of rounds whose length increases with n.

Theorem 2 *For any fixed values of $p_0 > 0$, and $1 < f_1 \leq f_2$, there is a constant K_0 such that the following holds: For any graph G with n vertices, and any $k \geq 1$, Algorithm 1 terminates in at most $K_0(k+1)\log n$ rounds, with probability at least $1 - O(1/n^k)$.*

Before beginning the proof of Theorem 2, it will be useful to define some notation and record a few simple facts. We will frequently use the well-known inequality

$$(1-\delta) \leq \exp(-\delta). \tag{2}$$

We will also use the following inequality, which holds for any $\lambda > 0$ and any $\delta \in [0, 1-e^{-\lambda}]$ (it holds with equality at the ends of this interval, and so holds at all points in between, by convexity).

$$(1-\delta) \geq \exp(-\delta\lambda/(1-e^{-\lambda})). \tag{3}$$

Finally, we will also need the following Chernoff-type inequality: if X is a sum of Bernoulli random variables, with expected value $\mathbb{E}X = m$, then for every $\delta > 0$,

$$\mathbb{P}[X > m+\delta] \leq \exp(-\delta^2/(2m+2\delta/3)).$$

In particular,

$$\mathbb{P}[X > 2m] \leq \exp(-m/3). \tag{4}$$

We refer to sending a signal in the first exchange of Algorithm 1 as "beeping", and receiving such a signal from a neighbour as "hearing a beep".

For any vertex v, we define $\mu_t(v)$, which we call the "weight" of v, to be the probability that v beeps in round t. (By convention, we set $\mu_t(v) = 0$ if v is inactive at time t; this simplifies notation, while allowing us to ignore the contribution of inactive vertices.) For any $W \subseteq V$ we write $\mu_t(W)$ for $\sum_{v\in W} \mu_t(v)$. Note that μ_t is a random measure on V, as it depends on the beeps of other vertices during the first $t-1$ rounds.

Recall that the set of vertices adjacent to a given vertex v is called the set of *neighbours* of v, and denoted by $\Gamma(v)$.

Definition 4 For any $\lambda > 0$, a vertex v will be called λ-*light* in round t if $\mu_t(\Gamma(v)) \leq \lambda$ and every neighbour of v has weight at most $1-\exp(-\lambda)$; otherwise, vertex v is called λ-*heavy*.

For any vertex that is λ-*light*, the weight of each of its neighbours individually is bounded by $1-\exp(-\lambda)$ and the sum of all its neighbours' weights is not too large (and so the vertex is not too likely to hear a beep at time t). Note that a fixed vertex may move back and forth between being λ-heavy and λ-light over time.

Our first result establishes a lower bound on the probability that at least one vertex in a set of λ-light vertices will be added to the independent set in the current round.

Lemma 1 *Let W be a set of vertices that are λ-light at round t. The probability that at least one vertex in W is added to the independent set in round t is at least $e^{-\phi\lambda}(1-e^{-\mu_t(W)})$ where $\phi = \lambda/(1-\exp(-\lambda))$.*

Proof Let us order the vertices of W as $w_1, \ldots, w_m$, where $m = |W|$. The probability that some vertex of W is added to the independent set in round t is at least the probability that the smallest vertex of W that beeps in round t is added to the independent set. For $i = 1, \ldots, m$, define events E_i and F_i by

$E_i = (w_i$ beeps; $w_1, \ldots, w_{i-1}$ do not beep)

$F_i = ($no neighbour of w_i beeps$)$.

The events $E_i \cap F_i$ are pairwise disjoint, so using the definition of conditional probability, we have that the probability that the smallest of W that beeps is added to the independent set is

$$\mathbb{P}\left[\bigcup_{i=1}^{m}(E_i \cap F_i)\right] = \sum_{i=1}^{m}\mathbb{P}[E_i \cap F_i] = \sum_{i=1}^{m}\mathbb{P}[E_i]\mathbb{P}[F_i|E_i].$$

It is easily seen that $\mathbb{P}[F_i|E_i] \geq \mathbb{P}[F_i]$ since $\mathbb{P}[F_i|E_i]$ is conditioned on the event that w_i beeps and $w_1, \ldots, w_{i-1}$ do not. Hence we have

$$\mathbb{P}[F_i|E_i] \geq \mathbb{P}[F_i] = \prod_{v\in\Gamma(w_i)}(1-\mu_t(v))$$

Since w_i is λ-light, we may apply Inequality (3), to conclude that

$$\begin{aligned}\prod_{v\in\Gamma(w_i)}(1-\mu_t(v)) &\geq \prod_{v\in\Gamma(w_i)}\exp(-\phi\mu_t(v))\\ &= \exp(-\phi\mu_t(\Gamma(w_i)))\\ &\geq \exp(-\phi\lambda)\end{aligned}$$

where $\phi = \lambda/(1-\exp(-\lambda))$. Hence we have

$$\sum_{i=1}^{m}\mathbb{P}[E_i]\mathbb{P}[F_i|E_i] \geq \exp(-\phi\lambda)\sum_{i=1}^{m}\mathbb{P}[E_i].$$

But $\sum_{i=1}^{m}\mathbb{P}[E_i]$ is simply the probability that some vertex in W beeps, which is given by $1-\prod_{v\in W}(1-\mu_t(v))$. Using Inequality (2) this value is at least $1-\exp(-\mu_t(W))$.

Thus the probability that some vertex of W is added to the independent set in round t is at least

$$\exp(-\phi\lambda)\sum_{i=1}^{m}\mathbb{P}[E_i] \geq e^{-\phi\lambda}(1-e^{-\mu_t(W)}).$$

□

Proof (of Theorem 2) Fix an arbitrary vertex v. We shall show that, with failure probability $O(1/n^{k+1})$, v becomes inactive within $K_0(k+1)\log n$ rounds, for a suitable choice of constant K_0. Taking a union bound over all n choices of v, it follows that with failure probability $O(1/n^k)$ every vertex becomes inactive and the algorithm terminates within $K_0(k+1)\log n$ rounds, which proves the theorem.

At each time step $t \geq 1$, we partition the neighbourhood of v into λ-light and λ-heavy vertices, for a suitable fixed choice of λ

$$\begin{aligned}L_t = L_t(v) &= \{x \in \Gamma(v) \mid x \text{ is } \lambda\text{-light at step } t\}\\ H_t = H_t(v) &= \{x \in \Gamma(v) \mid x \text{ is } \lambda\text{-heavy at step } t\}.\end{aligned}$$

We will follow the behaviour of $\mu_t(L_t)$ and $\mu_t(H_t)$ over time.

The idea of the argument is roughly as follows: if $\mu_t(L_t)$ is large at many rounds, then by Lemma 1 it is very likely that some neighbour of v will be added to the independent set on one of these occasions, leading to v becoming inactive. If this does not happen, then $\mu_t(L_t)$ must be small most of the time. Now consider H_t. Vertices that are λ-heavy at time t are likely to hear beeps and so drop in weight (as their signalling probability is reduced); it will follow that with high probability $\mu_{t+1}(H_t)$ is a constant factor smaller than $\mu_t(H_t)$ most of the time. Now we look at the evolution of $\mu_t(\Gamma(v))$, the weight of the whole neighbourhood of v. It may be large and increasing for some small fraction of the time, but mostly it is either shrinking or else it is already small. It will follow that, for at least some fixed fraction of the time, $\mu_t(\Gamma(v))$ is small. But this implies that, for at least some fixed fraction of the time, v will not hear any beeps, and hence $\mu_t(v)$ will be large for some fixed fraction of the time. This implies that it is very likely that at some point in the sequence of rounds we are considering v will beep and not hear any beeps, and so get added to the independent set.

To make this argument precise, we now define the following constants:

$$\begin{aligned}r &= 1+(\log f_2/\log f_1);\\ \lambda &= \log(32r(r+2)(f_2-f_1^{-1})/(f_2^{-1/r}-f_1^{-1}));\\ \phi &= \lambda/(1-\exp(-\lambda));\\ \beta &= 1/(4(r+2)\phi f_2);\\ \alpha &= (\beta/2)(f_2^{-1/r}-f_1^{-1})/(f_2-f_1^{-1});\\ K_0 &= (8r(r+2))\max\{6, 1/p_0, 1/\log f_2, 1/(e^{-\phi\lambda}(1-e^{-\alpha}))\};\end{aligned}$$

The values of these constants depend only on the fixed parameters f_1 and f_2 which bound the probability update factor f used in the algorithm, and on the initial minimum probability threshold p_0 (see Algorithm 1). Note that $1 < f_1 \leq f_2$, so $r \geq 2$ and $\lambda > \log 256 > 5$.

To simplify the presentation, we also define

$$K = K_0(k+1).$$

At each round t, we consider the following four possible events:

(E1) $\mu_t(L_t) \geq \alpha$
['$\Gamma(v)$ has a significant weight of light neighbours']

(E2) $\mu_t(L_t) < \alpha$ and $\mu_t(\Gamma(v)) \leq \beta$
['v is very light']

(E3) $\mu_t(L_t) < \alpha$, $\mu_t(\Gamma(v)) > \beta$ and $\mu_{t+1}(\Gamma(v)) \leq f_2^{-1/r}\mu_t(\Gamma(v))$
['the neighbourhood of v shrinks significantly in weight during round t']

(E4) $\mu_t(L_t) < \alpha$, $\mu_t(\Gamma(v)) > \beta$ and $\mu_{t+1}(\Gamma(v)) > f_2^{-1/r}\mu_t(\Gamma(v))$
['the neighbourhood of v does not shrink significantly in weight during round t (and may grow)']

Exactly one of these events must occur in each round.

We organize the rest of the proof as a series of claims.

Claim 1 With failure probability $O(1/n^{k+1})$, (E1) occurs at most $(K\log n)/(8r(r+2))$ times in the first $K\log n$ rounds.

Each time that (E1) occurs, it follows from Lemma 1 that with probability at least $e^{-\phi\lambda}(1-e^{-\mu_t(L_t)}) \geq e^{-\phi\lambda}(1-e^{-\alpha})$ some vertex of L_t is added to the independent set (and so v becomes inactive and the process at v terminates). Let $\phi_1 = e^{-\phi\lambda}(1-e^{-\alpha})$: the probability that (E1) occurs at least $(K\log n)/(8r(r+2))$ times without v becoming inactive is at most $(1-\phi_1)^{(K\log n)/(8r(r+2))}$ which is at most

$$\exp(-(\phi_1 K_0/(8r(r+2)))(k+1)\log n).$$

By our choice of K_0, we have $K_0 \geq (8r(r+2))/\phi_1$, so this probability is at most $\exp(-(k+1)\log n) = n^{-(k+1)}$. This proves Claim 1.

The bad event for us will be (E4), so let us bound the probability that (E4) occurs.

Claim 2 At each round t, the probability that (E4) occurs is at most $1/(16r(r+2))$.

If (E4) can occur, then we must have that $\mu_t(L_t) < \alpha$ and $\mu_t(\Gamma(v)) > \beta$. For any $x \in H_t$, there are two cases to consider - the first is that the total weight of all its neighbouring vertices is greater than λ; the second is that at least one of its neighbouring vertices individually has weight more than $1-\exp(-\lambda)$.

In the first case, using Inequality (2), the probability that no neighbour of x beeps in round t is at most $\exp(-\mu_t(\Gamma(x)))$ which is bounded by $\exp(-\lambda)$. In the second case, the probability that no neighbour of x beeps in round t is still at most $\exp(-\lambda)$. Thus, for any $x \in H_t$ we have shown that the probability that no neighbour of x beeps in round t is at most $\exp(-\lambda)$.

Let H_t^0 be the set of vertices in H_t that do not hear a beep in round t, and let $H_t^1 = H_t \setminus H_t^0$ be the remaining vertices in H_t that do hear a beep. Then $\mathbb{E}[\mu_t(H_t^0)] \leq \exp(-\lambda)\mu_t(H_t)$, and so by Markov's inequality

$$\mathbb{P}\left[\mu_t(H_t^0) \geq 16r(r+2)\exp(-\lambda)\mu_t(H_t)\right] \leq 1/(16r(r+2)). \tag{5}$$

Now all vertices in H_t^1 decrease their weight by a factor of at least f_1, while vertices in L_t and H_t^0 may either decrease or increase their weight (additionally, some weights may get set to 0 if vertices become inactive). So

$$\begin{aligned}\mu_{t+1}(\Gamma(v)) &\leq \frac{1}{f_1}\mu_t(H_t^1) + f_2\mu_t(H_t^0) + f_2\mu_t(L_t)\\ &= \frac{1}{f_1}\mu_t(\Gamma(v))\\ &\quad + (f_2 - \frac{1}{f_1})\mu_t(H_t^0) + (f_2 - \frac{1}{f_1})\mu_t(L_t)\end{aligned}$$

It follows from Inequality (5) that, with probability at least $1-1/(16r(r+2))$,

$$\begin{aligned}\mu_{t+1}(\Gamma(v)) &\leq \frac{1}{f_1}\mu_t(\Gamma(v)) + (f_2 - \frac{1}{f_1})16r(r+2)e^{-\lambda}\mu_t(H_t)\\ &\quad + (f_2 - \frac{1}{f_1})\mu_t(L_t)\\ &\leq f_2^{-1/r}\mu_t(\Gamma(v)),\end{aligned}$$

where the final inequality follows from our choice of λ, which gives

$$\begin{aligned}&(f_2 - f_1^{-1})16r(r+2)e^{-\lambda}\mu_t(H_t)\\ &\quad \leq (f_2 - f_1^{-1})16r(r+2)e^{-\lambda}\mu_t(\Gamma(v))\\ &\quad \leq 1/2(f_2^{-1/r} - f_1^{-1})\mu_t(\Gamma(v)),\end{aligned}$$

and our choice of α and β, because we are assuming that $\mu_t(L_t) < \alpha$ and $\mu_t(\Gamma(v)) > \beta$, so we have

$$\begin{aligned}(f_2 - f_1^{-1})\mu_t(L_t) &< (f_2 - f_1^{-1})\alpha\\ &= 1/2(f_2^{-1/r} - f_1^{-1})\beta\\ &< 1/2(f_2^{-1/r} - f_1^{-1})\mu_t(\Gamma(v)).\end{aligned}$$

Thus the probability that (E4) does occur is bounded above by $1/(16r(r+2))$. This proves Claim 2.

Claim 3 With failure probability $O(1/n^{k+1})$, (E4) occurs at most $(K\log n)/(8r(r+2))$ times in the first $K\log n$ rounds.

At each round, the probability of (E4) depends on the past history of the process. However, by Claim 2, it is always at most $1/(16r(r+2))$, and so we can couple occurrences of (E4) with a sequence of independent events each occurring with probability $1/(16r(r+2))$. It follows that the number of occurrences of (E4) in the first $K\log n$ rounds is stochastically dominated by a binomial random variable X with parameters $K\log n$ and $1/(16r(r+2))$. The probability that (E4) occurs more than $(K\log n)/(8r(r+2))$ times is therefore, by (4), at most

$$\mathbb{P}[X > 2\mathbb{E}X] \leq \exp(-\mathbb{E}X/3) \leq \exp(-(K\log n)/(48r(r+2))).$$

By our choice of K_0, we have $K_0 \geq 48r(r+2)$, so this probability is $O(n^{-(k+1)})$, which proves Claim 3.

From Claim 1 and Claim 3, we conclude that with failure probability $O(n^{-(k+1)})$, (E1) and (E4) altogether occur at most $K(\log n)/(4r(r+2))$ times in the first $K\log n$ rounds. We next show that, with small failure probability, $\mu_t(\Gamma(v))$ is small most of the time.

Claim 4 With failure probability $O(1/n^{k+1})$, $\mu_t(\Gamma(v)) > f_2\beta$ at most $(K\log n)/(2(r+2))$ times in the first $K\log n$ rounds.

Let T be the set of rounds $t \geq 1$ at which $\mu_t(\Gamma(v)) > f_2\beta$. We decompose T into (maximal) intervals of integers, say as $T_1 \cup \cdots \cup T_m$. Let $T_i = [s_i, t_i]$ be one of these intervals. We colour each integer $t \in T_i$ *red* if (E1) or (E4) occurred at the previous round, and *blue* if (E3) occurred (note that (E2) cannot occur, as $\mu_{t-1}(\Gamma(v)) \geq \mu_t(\Gamma(v))/f_2 > \beta$). By the definition of (E3), we have $\mu_t(\Gamma(v)) \leq f_2^{-1/r}\mu_{t-1}(\Gamma(v))$ at blue rounds, and we have $\mu_t(\Gamma(v)) \leq f_2\mu_{t-1}(\Gamma(v))$ otherwise. Let r_i be the number of red elements in T_i and b_i the number of blue elements. It follows that

$$\mu_{t_i}(\Gamma(v)) \leq \mu_{s_i-1}(\Gamma(v)) \cdot {f_2}^{r_i - b_i/r}.$$

Since $\mu_{t_i}(\Gamma(v)) > f_2\beta$ it follows that

$$r_i > \frac{1}{r}b_i + \log_{f_2} f_2\beta - \log_{f_2}(\mu_{s_i-1}(\Gamma(v))).$$

However, $\mu_{s_i-1}(\Gamma(v)) \leq f_2\beta$ in all cases where $s_i > 1$, and $\mu_0(\Gamma(v)) < n$. Summing over i, we see that

$$\sum_{i=1}^{m} r_i > \frac{1}{r}\sum_{i=1}^{m} b_i + \log_{f_2} f_2\beta - \log_{f_2} n$$

But red rounds correspond to events (E1) and (E4), and we have already shown in Claims 1 and 3 that these occur at most $(K\log n)/(4r(r+2))$ times altogether in the first $K\log n$ rounds. Hence the total number of rounds in T, both red and blue, is less than $(K\log n)/(4r(r+2)) + (K\log n)/(4(r+2)) + r\log n/\log f_2$. By our choice of K_0, this is less than $(K\log n)/(2(r+2))$. This proves Claim 4.

Claim 5 With failure probability $O(1/n^{k+1})$, v hears a beep at most $(K\log n)/(r+2)$ times in the first $K\log n$ rounds.

By our choice of β, we have that $f_2\beta = (1-e^{-\lambda})/(4(r+2)\lambda) < (1-e^{-\lambda})$. Hence we may apply Inequality (3), to show that when $\mu_t(\Gamma(v)) \leq f_2\beta$ the probability that v hears no beep is

$$\begin{aligned}\prod_{x\in\Gamma(v)}(1-\mu_t(x)) &\geq \prod_{x\in\Gamma(v)}\exp(-\phi\mu_t(x))\\ &\geq \exp(-\phi\mu_t(\Gamma(v)))\\ &\geq \exp(-\phi f_2\beta) \geq 1-\phi f_2\beta,\end{aligned}$$

and so v hears a beep with probability at most $\phi f_2\beta$ which equals $1/(4(r+2))$. Using (4), this implies that with failure probability $O(n^{-(k+1)})$ there are at most $(K\log n)/(2(r+2))$ rounds among the first $K\log n$ at which $\mu_t(\Gamma(v)) \leq f_2\beta$ and v hears a beep. By Claim 4, with the same failure probability, there are also at most $(K\log n)/(2(r+2))$ rounds at which $\mu_t(\Gamma(v)) > f_2\beta$ (and v might hear a beep at any of these steps). It follows that, with failure probability $O(n^{-(k+1)})$, v hears a beep at most $(K\log n)/(r+2)$ times in the first $K\log n$ rounds. This proves Claim 5.

Claim 6 With failure probability $O(1/n^{k+1})$, v becomes inactive during the first $K\log n$ rounds.

From the previous claim, we may assume that v hears a beep on at most $K\log n/(r+2)$ occasions during the first $K\log n$ rounds. On these occasions it decreases its local probability value p by a factor of at most f_2. We shall refer to these as red steps.

Hence there are at least $\frac{(r+1)}{(r+2)}K\log n$ rounds during the first $K\log n$ rounds where v does not hear a beep, so it either terminates, or increases its local probability value p by a factor of at least f_1, or else increases p to 1. We shall refer to these as blue rounds. Note that if v beeps in a blue round then it will terminate in that round. Hence a blue round where the value of p increases to 1 must be immediately followed by a red round, or a blue round where v terminates.

Now, by our choice of r, $f_1^r > f_2$. This means that there must be at least $\frac{1}{(r+2)}K\log n$ blue rounds during the first $K\log n$ rounds where either v has terminated, or else the local probability value p at v is at least as high as the initial value, p_0.

The probability that v will terminate at each of these blue rounds is at least p_0, so the probability that v remains active throughout all these blue rounds is at most $(1-p_0)^{K\log n/(r+2)}$. Using Inequality (2), this means that the probability that v remains active throughout these rounds is at most

$$\exp(-p_0 K\log n/(r+2)).$$

By our choice of K_0, we have $K_0 > (r+2)/p_0$, so this is $O(n^{-(k+1)})$. Hence v terminates with a failure probability that is $O(n^{-(k+1)})$.

This proves Claim 6, and completes the proof of Theorem 2. □

Corollary 1 *The expected number of rounds taken by Algorithm 1 on any graph with n nodes is* $O(\log n)$.

Proof Let T be the total number of rounds taken by the algorithm and let $T' = \lceil T/(K_0\log n)\rceil$, where K_0 is the constant identified in Theorem 2.

By Theorem 2, we have that, for any $k \geq 1$,

$$\mathbb{P}[T' > k+1] \leq c'/n^k$$

for some constant c'. Hence $\mathbb{E}[T'] = \sum_{k\geq 1}\mathbb{P}[T' \geq k] = O(1)$. □

4.3 Expected Number of Signals

In this section we will show that the expected number of times that each node signals is bounded by a constant. Hence the expected bit complexity per node for this algorithm *does not increase at all* with the number of nodes.

Theorem 3 *The expected total number of signals broadcast by any node executing Algorithm 1 is* $O(1)$.

Proof Let v be a node executing Algorithm 1, and consider the whole sequence of rounds until v becomes inactive.

Once again we will refer to sending a signal in the first exchange of Algorithm 1 as "beeping", and receiving such a signal from a neighbour as "hearing a beep".

During each round, one of the following 3 things happens in the first exchange:

Case 1 v hears a beep from its neighbours, and so decreases its probability of beeping by a factor of f (from p_t to $\frac{1}{f}p_t$). We will call these *red* rounds.

Case 2 v hears no beep from its neighbours, and so increases its probability of beeping by a factor of f (from p_t to fp_t). We will call these *blue* rounds.

Case 3 v hears no beep from its neighbours, and so increases its probability of beeping to 1. We will call these *dark blue* rounds.

If v beeps in any blue or dark blue round then it joins the MIS and becomes inactive, so the total number of beeps at such rounds is at most one, at the final round in the sequence. Hence we only need to consider the expected number of beeps at red rounds.

Consider first those red rounds (if any) where the value of p_t is at its lowest point in the sequence so far. At each such round the lowest probability value seen so far decreases by a factor of at least f_1. Hence, the expected number of times that v beeps during this subsequence of rounds is bounded by $p_0 + \frac{p_0}{f_1} + \frac{p_0}{f_1^2} + \frac{p_0}{f_1^3} \cdots \leq \frac{f_1}{f_1-1} p_0 \leq \frac{f_1}{f_1-1}$.

At all of the remaining red rounds the value of p_t is not at its lowest point so far, so it was lower at some previous blue round. Hence each of these red rounds can be associated with a corresponding earlier blue round: the most recent blue round where the value of p_t was lower. We now define the constant $r = \lceil \log f_2 / \log f_1 \rceil$, where f_1 and f_2 denote respectively the lower bound and upper bound of f as given in Algorithm 1. Note that $f_1^r \geq f_2$. Since blue rounds increase the value of p_t by at most a factor of f_2, and red rounds decrease the value of p_t by at least a factor of f_1, it follows that each blue round will be associated with at most r red rounds.

Hence we have partitioned the remaining red rounds into groups of at most r red rounds, each associated with a single (earlier) blue round. We now consider these groups of at most $r+1$ rounds, ordered by the position of the initial blue round. For any such group, if the probability of beeping at the blue round is p, then the probability of beeping at any of the associated red rounds is at most $f_2 p$. Hence the conditional probability of beeping at the initial blue round, given that a beep occurs somewhere in this group of rounds, is at least $1/(1+rf_2)$. Hence if we consider the subsequence of groups where at least one beep occurs, the expected number of such groups before a beep occurs at a blue round is at most rf_2 (expected number of failures before the first success in a geometric distribution). Since each group can contribute at most r beeps, the expected number of beeps added in these groups before terminating is at most $r(rf_2)$.

We have shown that the expected number of times that v beeps is at most $1 + \frac{f_1}{f_1-1} + (\lceil \log f_2 / \log f_1 \rceil)^2 f_2$, which proves the result. □

5 Algorithm for Distributed Greedy Colouring

Our new algorithm for distributed greedy colouring (Algorithm 2) is similar to our new distributed MIS selection algorithm. At each round, each node may choose, with a certain probability p, to broadcast its first available colour to all its neighbours, indicating that it wishes to use that colour. If two neighbouring nodes broadcast the same colour in the same round, then they will both abandon choosing that colour in that round. On the other hand, if a node broadcasts a colour and none of its neighbouring nodes broadcast the same colour in that round, then it is successfully coloured, and will notify all its neighbouring nodes that they are forbidden to use that colour.

As in our MIS selection algorithm, the way that the probability values p are chosen is inspired by the positive feedback mechanisms that control cellular processes. The value of p is initialised to some arbitrary value at each node (above some strictly positive fixed threshold value, p_0). These values are then independently updated at each node in each round using feedback from neighbouring nodes. The value of p is decreased at a node whenever one or more neighbouring nodes broadcast the same colour, and is increased whenever no neighbouring node broadcasts the same colour. As in our MIS selection algorithm, we allow each increase or decrease to be by some arbitrary factor f, which may vary at each round, but is always bounded by the global parameters f_1 and f_2 (with $1 < f_1 \leq f_2$).

The correctness of Algorithm 2 follows easily from the two facts below:

Fact 1 No two nodes that are assigned the same colour in the same round are adjacent.

Fact 2 The colour assigned to any node is the smallest colour that is different from all colours previously assigned to neighbouring nodes.

Thus, if Algorithm 2 is run on the nodes of any graph G, and all nodes become inactive, then the colour assigned to each of the nodes defines a greedy colouring of G.

Our analysis of the distributed greedy colouring algorithm (Algorithm 2) is very similar to the analysis for the MIS selection algorithm given in Section 4.2.

Theorem 4 *For any fixed values of $p_0 > 0$, and $1 < f_1 \leq f_2$, there is a constant K_0 and a constant r such that the following holds: For any graph G with n vertices and maximum degree Δ, and any $k \geq 1$, Algorithm 2 terminates in at most $8r(r+2)\Delta + K_0(k+1)\log n$ rounds, with probability at least $1 - O(1/n^k)$.*

As in Section 4.2, we refer to broadcasting any colour c in the first exchange as "beeping", and receiving *the same* colour c from a neighbour in that exchange as "hearing a beep". As before, for any vertex v, at any time step t, we define the measure $\mu_t(v)$, called the "weight" of v, to be the probability that v beeps in round t.

The set of neighbours of a vertex v which are competing for the same colour as v in any round will be called the *homogeneous* neighbours of v, and will be denoted by $\Gamma^{(h)}(v)$. We adapt Definition 4 to refer to homogeneous neighbours only, as follows.

ALGORITHM 2: The algorithm for distributed greedy colouring at each node

Global constants: p_0 : lower bound on initial probability value;
f_1, f_2 : lower and upper bounds on change factor for probability value.

Local variables: p : local probability value, initialised to some value in $[p_0, 1]$;
f : change factor for probability value, chosen arbitrarily in $[f_1, f_2]$;
TRYING : Boolean flag, initialised to FALSE;
S : Set of forbidden colours, i.e., those taken by neighbours, initialised to $\emptyset$.

1. **while** active, at each time step **do**
2. *FIRST EXCHANGE*
3. Choose the smallest available colour c that is not in S;
4. With probability p, set TRYING $\leftarrow$ TRUE and **send** c to all neighbours;
5. **Receive** any colour signals sent by neighbours;
6. Set f to some arbitrary value in the interval $[f_1, f_2]$;
7. **if** any neighbour sent colour c **then**
8. TRYING $\leftarrow$ FALSE and $p \leftarrow p/f$ (decrease p)
9. **else**
10. $p \leftarrow \min\{fp, 1\}$ (increase p)
11. *SECOND EXCHANGE*
12. **if** TRYING **then**
13. **Send** c to all neighbours;
14. Assign colour c to this node and terminate (become inactive).
15. **Receive** any colour signals sent by neighbours and add all distinct colours received to S.

Definition 5 For any $\lambda > 0$, a vertex v will be called λ-*light*$^{(h)}$ at round t if $\mu_t(\Gamma^{(h)}(v)) \leq \lambda$ and every homogeneous neighbour of v has weight at most $1 - \exp(-\lambda)$; otherwise, vertex v is called λ-*heavy*$^{(h)}$.

As in Section 4.2, we first establish a lower bound on the probability that at least one vertex in a set of λ-light$^{(h)}$ vertices will be coloured at each round.

Lemma 2 *Let W be a set of vertices that are λ-light$^{(h)}$ at round t. The probability that at least one vertex in W gets coloured in round t is at least $e^{-\phi\lambda}(1 - e^{-\mu_t(W)})$ where $\phi = \lambda/(1 - \exp(-\lambda))$.*

Proof Identical to the proof of Lemma 1. □

Proof (of Theorem 4) Fix an arbitrary vertex v. We use essentially the same argument as in the proof of Theorem 2, partitioning the neighbourhood of v into λ-light$^{(h)}$ and λ-heavy$^{(h)}$ vertices, and following the progress of these sets over time. Note that we consider the entire neighbourhood of v, not just the homogeneous neighbours. We show that the weight of this entire neighbourhood is small for at least a fixed fraction of the time, and hence v fails to receive any colour signals for at least a fixed fraction of the time.

We define the same constants, and the same events $(E1)$ to $(E4)$ (see page 14). However, in this proof we will need to allow for the possibility that one or more neighbours of v are successfully coloured at any time step, which can happen up to Δ times, and does not immediately force v to become inactive.

Each time that (E1) occurs, it follows from Lemma 2 that with probability at least $e^{-\phi\lambda}(1 - e^{-\alpha})$ some λ-light$^{(h)}$ neighbour of v will be coloured. Let $\phi_1 = e^{-\phi\lambda}(1 - e^{-\alpha})$. As in the proof of Claim 1, the probability that there are $(K \log n)/(8r(r+2))$ occurrences of $(E1)$ where no neighbour of v is coloured is at most $(1 - \phi_1)^{(K \log n)/(8r(r+2))} \leq \exp(-(\phi_1 K_0/(8r(r+2)))(k+1)\log n)$. By our choice of K_0, we have $K_0 \geq (8r(r+2))/\phi_1$, so this probability is at most $\exp(-(k+1)\log n) = n^{-(k+1)}$.

Hence with failure probability $O(1/n^{k+1})$, the *total* number of occurrences of $(E1)$ is at most $\Delta + (K \log n)/(8r(r+2))$.

Since we have a weaker upper bound on the number of occurrences of $(E1)$, we will need to consider a longer sequence of rounds overall. In fact, we will consider the first $8r(r+2)\Delta + K \log n$ rounds.

Following exactly the same arguments as in the proof of Claims 2 and 3, we obtain that with failure probability $O(1/n^{k+1})$, (E4) occurs at most $\Delta + (K \log n)/(8r(r+2))$ times in the first $8r(r+2)\Delta + K \log n$ rounds.

Next, following the argument used to prove Claim 4, but using the weaker bound of $2\Delta + (K \log n)/(4r(r+2))$ for the total number of red rounds, we obtain that with failure probability $O(1/n^{k+1})$, $\mu_t(\Gamma(v)) > f_2\beta$ at most $4r\Delta + (K \log n)/(2(r+2))$ times in the first $8r(r+2)\Delta + K \log n$ rounds.

Now, using the same argument as in Claim 5 we can show that with failure probability $O(1/n^{k+1})$, the fraction of rounds where v hears a beep (or in fact receives any colour signal in the first exchange) during the first $8r(r+2)\Delta + K \log n$ rounds is at most $1/(r+2)$.

Finally, using the same argument as in Claim 6, this implies that with failure probability $O(1/n^{k+1})$, v becomes inactive during the first $8r(r+2)\Delta + K\log n$ rounds.

Taking a union bound over all possible choices of v, as before, gives the result. □

Corollary 2 *The expected number of rounds taken by Algorithm 2 on any graph with n nodes is* $O(\Delta + \log n)$.

Finally, the proof of Theorem 3 considers only an individual node and does not take into account whether the neighbours of a node become inactive or not, so this proof applies equally well to our distributed colouring algorithm, giving the following result.

Theorem 5 *The expected total number of beeps broadcast by any node executing Algorithm 2 is* $O(1)$.

6 Conclusions

We have described a new, very straightforward, randomised distributed MIS selection algorithm that is able to operate on a network of identical anonymous processors and requires no global information about the network at any node.

By introducing a new analytical technique we have shown that on any network with n nodes our algorithm runs in $O(\log n)$ time (both in expectation and with high probability) and the expected number of messages sent by each node is bounded by a constant.

We have also shown that a very similar approach yields a simple and effective algorithm for the distributed greedy colouring problem. A very similar analysis shows that on any network with n nodes of maximum degree Δ this algorithm computes a greedy colouring in $O(\Delta + \log n)$ time (both in expectation and with high probability) and the total number of messages sent by each processor is again bounded by a constant.

The communication model that we use for our MIS algorithm is a version of the "beeping model" introduced in [11], known as "beeping with simultaneous wake-up and sender-side collision detection" [1]. At each time step, each node can either broadcast a single bit to all its neighbours, or remain silent, and each node can detect whether one or more neighbours have broadcast, but cannot tell how many neighbours have broadcast, or which ones. It is an interesting question whether similar algorithms can be developed for harsher versions of the beeping model, in particular when there is no collision detection, so nodes can either transmit or detect transmission from their neighbours, but not both.

It would also be interesting to see whether the simple idea of adjusting probabilities at each node using local feedback can be extended to other fundamental problems in distributed computing.

References

1. Afek, Y., Alon, N., Bar-Joseph, Z., Cornejo, A., Haeupler, B., Kuhn, F.: Beeping a maximal independent set. In: Proceedings of the 25th International Conference on Distributed Computing, DISC'11, pp. 32–50. Springer-Verlag (2011)
2. Afek, Y., Alon, N., Barad, O., Hornstein, E., Barkai, N., Bar-Joseph, Z.: A biological solution to a fundamental distributed computing problem. Science **331**(6014), 183–185 (2011)
3. Alon, N., Babai, L., Itai, A.: A fast and simple randomized parallel algorithm for the maximal independent set problem. Journal of Algorithms **7**, 567–583 (1986)
4. Barenboim, L., Elkin, M.: Distributed $(\delta+1)$-coloring in linear (in δ) time. In: Proceedings of the 41st Annual ACM Symposium on Theory of Computing, STOC '09, pp. 111–120. ACM, New York, NY, USA (2009)
5. Barenboim, L., Elkin, M.: Sublogarithmic distributed MIS algorithm for sparse graphs using Nash-Williams decomposition. Distributed Computing **22**(5-6), 363–379 (2010)
6. Barenboim, L., Elkin, M.: Deterministic distributed vertex coloring in polylogarithmic time. Journal of the ACM **58**(5), 23:1–23:25 (2011)
7. Barenboim, L., Elkin, M., Pettie, S., Schneider, J.: The locality of distributed symmetry breaking. In: Proceedings of the 2012 IEEE 53rd Annual Symposium on Foundations of Computer Science, FOCS '12, pp. 321–330. IEEE Computer Society, Washington, DC, USA (2012)
8. Bray, S.J.: Notch signalling: a simple pathway becomes complex. Nature reviews Molecular cell biology **7**(9), 678–689 (2006)
9. Chaudhuri, P.: Algorithms for some graph problems on a distributed computational model. Information Sciences **43**(3), 205–228 (1987)
10. Collier, J.R., Monk, N.A., Maini, P.K., Lewis, J.H.: Pattern formation by lateral inhibition with feedback: a mathematical model of delta-notch intercellular signalling. Journal of Theoretical Biology **183**(4), 429–446 (1996)
11. Cornejo, A., Kuhn, F.: Deploying wireless networks with beeps. In: Proceedings of the 24th International Conference on Distributed Computing, DISC'10, pp. 148–162. Springer-Verlag, Berlin, Heidelberg (2010)
12. Emek, Y., Wattenhofer, R.: Stone age distributed computing. In: Proceedings of the 2013 ACM Symposium on Principles of Distributed Computing, PODC '13, pp. 137–146. ACM, New York, NY, USA (2013)
13. Gavoille, C., Klasing, R., Kosowski, A., Kuszner, Ł., Navarra, A.: On the complexity of distributed graph coloring with local minimality constraints. Networks **54**(1), 12–19 (2009)
14. Gebremedhin, A.H., Lassous, I.G., Gustedt, J., Telle, J.A.: Graph coloring on coarse grained multicomputers. Discrete Applied Mathematics **131**(1), 179–198 (2003)
15. Goldberg, A.V., Plotkin, S.A., Shannon, G.E.: Parallel symmetry-breaking in sparse graphs. SIAM Journal on Discrete Mathematics **1**(4), 434–446 (1988)
16. Greenlaw, R., Hoover, H.J., Ruzzo, W.L. (eds.): Limits to Parallel Computation: P-completeness Theory. Oxford University Press, Inc., New York, NY, USA (1995)
17. Grundy, P.: Mathematics and games. Eureka **2**, 6–8 (1939)
18. Halpern, J.Y., Moses, Y.: Knowledge and common knowledge in a distributed environment. Journal of the ACM **37**(3), 549–587 (1990)
19. Hansen, J., Kubale, M., Kuszner, Ł., Nadolski, A.: Distributed largest-first algorithm for graph coloring. In: M. Danelutto, M. Vanneschi, D. Laforenza (eds.) Euro-Par 2004 Parallel Processing, *Lecture Notes in Computer Science*, vol. 3149, pp. 804–811. Springer Berlin Heidelberg (2004)

20. Hedetniemi, S.T., Jacobs, D.P., Srimani, P.K.: Linear time self-stabilizing colorings. Information Processing Letters **87**(5), 251–255 (2003)
21. Itai, A., Rodeh, M.: Symmetry breaking in distributed networks. Information and Computation **88**(1), 60–87 (1990)
22. Johansson, O.: Simple distributed $\Delta+1$-coloring of graphs. Information Processing Letters **70**(5), 229 – 232 (1999)
23. Karp, R.M.: Reducibility among combinatorial problems. In: R.E. Miller, J.W. Thatcher (eds.) Complexity of Computer Computations, The IBM Research Symposia Series, pp. 85–103. Plenum Press, New York (1972)
24. Karp, R.M., Wigderson, A.: A fast parallel algorithm for the maximal independent set problem. Journal of the ACM **32**(4), 762–773 (1985)
25. Kroeker, K.L.: Biology-inspired networking. Communications of the ACM **54**, 11–13 (2011)
26. Kuhn, F.: Weak graph colorings: distributed algorithms and applications. In: Proceedings of the 21st Annual Symposium on Parallelism in Algorithms and Architectures, SPAA '09, pp. 138–144. ACM, New York, NY, USA (2009)
27. Kuhn, F., Moscibroda, T., Nieberg, T., Wattenhofer, R.: Fast deterministic distributed maximal independent set computation on growth-bounded graphs. In: P. Fraigniaud (ed.) Distributed Computing: 19th International Conference, DISC 2005, *Lecture Notes in Computer Science*, vol. 3724, pp. 273–283. Springer (2005)
28. Kuhn, F., Moscibroda, T., Wattenhofer, R.: What cannot be computed locally! In: Proceedings of the 23rd Annual ACM Symposium on Principles of Distributed Computing, PODC '04, pp. 300–309. ACM, New York, NY, USA (2004)
29. Kuhn, F., Moscibroda, T., Wattenhofer, R.: The price of being near-sighted. In: Proceedings of the Seventeenth Annual ACM-SIAM Symposium on Discrete Algorithm, SODA '06, pp. 980–989. New York, NY, USA (2006)
30. Kuhn, F., Moscibroda, T., Wattenhofer, R.: Local computation: Lower and upper bounds. CoRR **abs/1011.5470** (2010). URL http://arxiv.org/abs/1011.5470
31. Kuhn, F., Wattenhofer, R.: On the complexity of distributed graph coloring. In: Proceedings of the 25th Annual ACM Symposium on Principles of Distributed Computing, PODC '06, pp. 7–15. New York, NY, USA (2006)
32. Lenzen, C., Wattenhofer, R.: Distributed algorithms for sensor networks. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **370**(1958), 11–26 (2012)
33. Linial, N.: Legal coloring of graphs. Combinatorica **6**(1), 49–54 (1986)
34. Linial, N.: Distributive graph algorithms – global solutions from local data. In: Proceedings of the 28th Annual Symposium on Foundations of Computer Science, SFCS '87, pp. 331–335. IEEE Computer Society, Washington, DC, USA (1987)
35. Linial, N.: Locality in distributed graph algorithms. SIAM Journal on Computing **21**(1), 193–201 (1992)
36. Luby, M.: A simple parallel algorithm for the maximal independent set problem. SIAM Journal on Computing **15**(4), 1036–1053 (1986)
37. Lynch, N.A.: Distributed Algorithms. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA (1996)
38. Maan, V., Purohit, G.N.: A distributed approach for frequency allocation using graph coloring in mobile networks. International Journal of Computer Applications **58**(6), 9–13 (2012). Published by Foundation of Computer Science, New York, USA
39. Métivier, Y., Robson, J.M., Saheb-Djahromi, N., Zemmari, A.: About randomised distributed graph colouring and graph partition algorithms. Information and Computation **208**(11), 1296–1304 (2010)
40. Métivier, Y., Robson, J.M., Saheb-Djahromi, N., Zemmari, A.: An optimal bit complexity randomized distributed MIS algorithm. Distributed Computing **23**(5-6), 331–340 (2011)
41. Moscibroda, T., Wattenhofer, R.: Maximal independent sets in radio networks. In: Proceedings of the 24th Annual ACM Symposium on Principles of Distributed Computing, PODC '05, pp. 148–157. ACM, New York, NY, USA (2005)
42. Ni, J., Srikant, R., Wu, X.: Coloring spatial point processes with applications to peer discovery in large wireless networks. IEEE/ACM Transactions on Networking **19**(2), 575 –588 (2011)
43. Panconesi, A., Rizzi, R.: Some simple distributed algorithms for sparse networks. Distributed Computing **14**(2), 97–100 (2001)
44. Panconesi, A., Srinivasan, A.: On the complexity of distributed network decomposition. Journal of Algorithms **20**(2), 356–374 (1996)
45. Park, T., Lee, C.Y.: Application of the graph coloring algorithm to the frequency assignment problem. Journal of the Operations Research Society of Japan-Keiei Kagaku **39**(2), 258–265 (1996)
46. Peleg, D.: Distributed computing: a locality-sensitive approach. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA (2000)
47. Prakash, R., Raynal, M., Singhal, M.: An adaptive causal ordering algorithm suited to mobile computing environments. Journal of Parallel and Distributed Computing **41**(2), 190–204 (1997)
48. Schneider, J., Wattenhofer, R.: A log-star distributed maximal independent set algorithm for growth-bounded graphs. In: Proceedings of the 27th ACM Symposium on Principles of Distributed Computing, PODC '08, pp. 35–44. ACM, New York, NY, USA (2008)
49. Schneider, J., Wattenhofer, R.: A new technique for distributed symmetry breaking. In: Proceedings of the 29th ACM SIGACT-SIGOPS Symposium on Principles of Distributed Computing, PODC '10, pp. 257–266. ACM, New York, NY, USA (2010)
50. Scott, A., Jeavons, P., Xu, L.: Feedback from nature: an optimal distributed algorithm for maximal independent set selection. In: Proceedings of the 2013 ACM Symposium on Principles of Distributed Computing, PODC '13, pp. 147–156. ACM, New York, NY, USA (2013)
51. Waters, R.J.: Graph colouring and frequency assignment. Ph.D. thesis (2005)
52. Wattenhofer, R.: http://dcg.ethz.ch/lectures/fs08/distcomp/lecture/chapter4.pdf (2007)
53. Xu, L., Jeavons, P.: Simple neural-like P systems for maximal independent set selection. Neural Computation **25**(6), 1642–1659 (2013)
54. Xu, L., Jeavons, P.: Patterns from nature: Distributed greedy colouring with simple messages and minimal graph knowledge. Information Sciences **316**, 550–566 (2015)
55. Zuckerman, D.: Linear degree extractors and the inapproximability of max clique and chromatic number. In: Proceedings of the 38th Annual ACM Symposium on Theory of Computing, STOC '06, pp. 681–690. ACM, New York, NY, USA (2006)