Properties of graphs with large girth

by

Carlos Hoppen

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement of the degree of
Doctor of Philosophy
in
Combinatorics and Optimization

Waterloo, Ontario, Canada, 2008

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

This thesis is devoted to the analysis of a class of iterative probabilistic algorithms in regular graphs, called *locally greedy algorithms*, which will provide bounds for graph functions in regular graphs with large girth. This class is useful because, by conveniently setting the parameters associated with it, we may derive algorithms for some well-known graph problems, such as algorithms to find a large independent set, a large induced forest, or even a small dominating set in an input graph G. The name "locally greedy" comes from the fact that, in an algorithm of this class, the probability associated with the random selection of a vertex v is determined by the current state of the vertices within some fixed distance of v.

Given $r \geq 3$ and an r-regular graph G, we determine the expected performance of a locally greedy algorithm in G, depending on the girth g of the input and on the degree r of its vertices. When the girth of the graph is sufficiently large, this analysis leads to new lower bounds on the independence number of G and on the maximum number of vertices in an induced forest in G, which, in both cases, improve the bounds previously known. It also implies bounds on the same functions in graphs with large girth and maximum degree r and in random regular graphs. As a matter of fact, the asymptotic lower bounds on the cardinality of a maximum induced forest in a random regular graph improve earlier bounds, while, for independent sets, our bounds coincide with asymptotic lower bounds first obtained by Wormald. Our result provides an alternative proof of these bounds which avoids sharp concentration arguments.

The main contribution of this work lies in the method presented rather than in these particular new bounds. This method allows us, in some sense, to directly analyse prioritised algorithms in regular graphs, so that the class of locally greedy algorithms, or slight modifications thereof, may be applied to a wider range of problems in regular graphs with large girth.

Acknowledgements

I thank faculty, staff and fellow graduate students of the Department of Combinatorics and Optimization for affording me very enjoyable years in Waterloo. I am especially grateful to my supervisor, Nick Wormald, whose support and guidance have been fundamental in the accomplishment of this thesis.

I would also like to express deep gratitude to my family, who, despite the distance, have always been cheerful supporters of my academic career.

Contents

Chapter 1

Introduction

In this thesis, we shall analyse a class of iterative probabilistic algorithms in regular graphs, which we call *locally greedy algorithms*. This will provide bounds for graph functions in regular graphs with large girth, since, by conveniently setting the parameters associated with locally greedy algorithms, we could derive algorithms for some well-known graph problems, such as algorithms to find a large independent set and a large induced forest in an input graph G. The name "locally greedy" comes from the fact that, in an algorithm of this class, the random selection of a vertex v occurs with a probability determined by the current state of the vertices within some fixed distance of v. A more precise name for this class would be "random locally greedy", but we omit the reference to randomness for concision.

The output produced by one of these algorithms is a subset of the vertices of the input graph. If this set satisfies a particular property P, the expected performance of the algorithm naturally leads to bounds on the size of an extremal set in G with property P. Indeed, this expectation is a lower bound on the cardinality of a maximum set with this property, as one of the possible outcomes of the algorithm is a set containing at least as many elements as the expected size of the set produced in a given application of the algorithm. Since another possible outcome is a set with at most as many elements, this expectation is also an upper bound on the cardinality of a minimum set satisfying property P.

The above observation motivates the following approach. We determine the expected performance of a locally greedy algorithm in any fixed regular graph, which, as we shall see, depends on two parameters, the girth g of the input and the degree r of its vertices. This is then used to give lower bounds, in the case of graphs with large girth, on the independence number of G and on the maximum number of vertices in an induced forest in G. In both cases, the bounds obtained here are a significant improvement in the bounds previously known. We would like to emphasise, however, that the main contribution of this work lies in the method presented rather than in these particular new bounds. This method allows us, in some sense, to directly analyse algorithms in regular graphs for which some operations are applied with

higher priority then others. Such algorithms are called *prioritised*. As a consequence, the class of locally greedy algorithms, or slight modifications thereof, may be applied to a wider range of problems in regular graphs with large girth.

1.1 Preliminaries

Before discussing our work in more detail, we give a few basic definitions. For the terminology not explicitly defined in this thesis, the reader is referred to Diestel [21].

In this thesis, a *graph* is a pair $G = (V, E)$ of sets, called the sets of *vertices* and *edges*, respectively, such that

$$E \subseteq \{\{u, v\} \in V^2 : u \neq v\}.$$

In particular, all the graphs we consider have no loops or multiple edges. The *degree* of a vertex v is defined to be the number of elements of E containing v, and a graph is *r-regular* if all its vertices have degree equal to r.

A *cycle* in a graph G is a sequence $v_0e_1v_1e_2 \dots e_tv_t$ in which the elements v_i are vertices, all of which are distinct with the exception of $v_0 = v_t$, and the elements e_i are the edges $\{v_{i-1}, v_i\}$ in the edge set of the graph. The number t of edges in a cycle is called the *length* of the cycle, while the *girth* of a graph is the length of a shortest cycle in the graph. By convention, the girth is infinite if the graph is acyclic. The Petersen graph, a 3-regular graph with girth 5 is given in Figure 1.1.1.

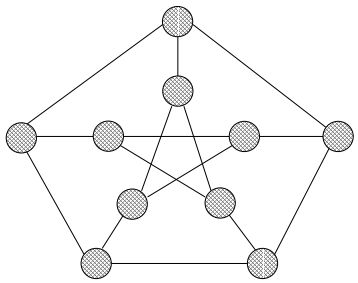

Figure 1.1.1: A 3-regular graph with girth 5.

The results in this thesis concern graphs with large girth. The interplay between the girth of a graph and other graph parameters has been an active subject of study in graph theory for a long time. An example of this is a seminal probabilistic result, proved by Erdős [25] in the late fifties, which establishes that, for any given positive integers k and g, there exists a graph with girth at least g and chromatic number at least k, evincing the global character of the chromatic number of a graph.

A series of structural results about this class of graphs have also been established. To cite a few, Thomassen [52] proved that graphs with large girth and minimum degree at least three share many properties with graphs of large minimum degree. For instance, graphs in

these two classes have a contraction containing a large clique, they contain cycles of all even lengths modulo a prescribed natural number and they have many disjoint cycles of the same length. Moreover, in terms of topological minors, Kühn and Osthus [38] showed that, if H is a graph whose maximum degree $\Delta(H)$ is at least 2, then every graph G of minimum degree at least $\max\{\Delta(H), 3\}$ and girth at least

$$\frac{166 \log |V(H)|}{\log \Delta(H)}$$

contains a subdivision of H.

Another driving force in the study of graphs with large girth is the interest in expander graphs, which play an important role in the theory of error-correcting codes, in the derandomisation of random algorithms and in the design of robust computer networks, to name a few of its applications. One way of obtaining an expander family is to find a sequence $\{G_1, G_2, \ldots\}$ of graphs with bounded degree satisfying

$$\lim_{n \to \infty} g(G_n) = \infty,$$

where $g(G_n)$ denotes the girth of the n-th graph in the sequence. This leads to the problem of explicitly constructing families of graphs containing elements with arbitrarily large girth. Several constructions, involving algebraic and number-theoretical tools, can be found in Davidoff et al. [19] and in Biggs [12]. For a comprehensive account of expander graphs, see Hoory, Linial and Wigderson [30].

1.2 Independent sets

A set of vertices in a graph is called *independent* if no two vertices in the set are joined by an edge. Such a set is also known under the name of *stable set*. A classical decision problem in graph theory is the *independent set problem*, which, for a given graph G and a fixed positive integer k, asks whether G contains an independent set of size at least k. The corresponding optimisation problem is the *maximum independent set problem*, which consists of determining the size $\alpha(G)$ of a largest independent set in a graph G. The interest in this problem has been partly fueled by its wide range of real world applications, which include problems such as information retrieval, shape and pattern recognition [16], signal transmission analysis and scheduling (see [27] and [5]). Independent sets also play an important role in coding and information theory [15].

The independent set problem is inherently difficult for general graphs. Indeed, it is among the early problems shown to be NP-complete in a seminal paper by Karp [36]. However, there exist efficient algorithms to obtain $\alpha(G)$ in some particular classes of graphs, such as line graphs, perfect graphs and their complements, claw-free graphs and bipartite graphs. For

references on these and other instances for which the independent set problem is polynomial-time solvable, the reader is referred to [48].

Since the independence number of a graph is in general hard to compute, much of the research in this problem has been devoted to obtaining bounds on $\alpha(G)$. A general lower bound in terms of the maximum degree $\Delta(G)$ has been provided by Berge (for a proof, see for instance [11]) through a simple greedy algorithm. It establishes that

$$\alpha(G) \geq \left\lceil \frac{|V(G)|}{\Delta(G)+1} \right\rceil.$$

This bound is tight for the complete graph K_n, the graph on n vertices for which every pair of vertices is adjacent. Clearly, the maximum degree is $\Delta = n-1$ and the largest independent set has size one. However, this bound is very weak in other cases. One such example is the complete bipartite graph $K_{n,n}$. On the one hand, since $\Delta = n$, Berge's bound is equal to 2. On the other hand, a largest independent set in the graph has size n, with an extremal set being given by each of the classes in the bipartition of the graph.

In terms of upper bounds, a spectral result stating that

$$\alpha(G) \leq \min\{|V(G)| - n^+(A), |V(G)| - n^-(A)\}, \tag{1.2.1}$$

where $n^+(A)$ and $n^-(A)$ denote the number of positive and negative eigenvalues in the adjacency matrix A of G, respectively, was found by Cvetković et al. [17]. The same examples of the previous case illustrate situations for which this bound is tight or behaves poorly. Indeed, the eigenvalues of the adjacency matrix of K_n are $n-1$, with multiplicity 1, and -1, with multiplicity $n-1$, so the above bound is tight. For $K_{n,n}$, however, the eigenvalues are n and $-n$, each with multiplicity 1, and 0, with multiplicity $2n-2$. As a consequence, the bound in (1.2.1) gives us

$$n = \alpha(G) \leq \min\{|V(G)| - n^+(A), |V(G)| - n^-(A)\} = 2n-1,$$

and therefore may be arbitrarily far from the actual size of the maximum independent set. The values of the eigenvalues of these graphs can be found in any book in algebraic graph theory, see for instance [28].

Again, more can be said for particular classes of graphs. One such class is the class of planar graphs. In their proof of the Four-Colour Theorem [46], Robertson et al. give a polynomial-time algorithm of $O(n^2)$ steps that finds a 4-colouring of a planar graph G, and hence provides an independent set of this graph containing at least $|V(G)|/4$ vertices, namely the largest colour class in this colouring.

Our work studies large independent sets in the class of r-regular graphs with large girth, where r is fixed. We only consider $r \geq 3$, since the cases $r = 1$ and $r = 2$ are very simple. To simplify the statement of bounds on the independence number of a graph, we introduce the

related concept of *independence ratio* $\alpha_R(G)$ of a graph G. This quantity is defined by

$$\alpha_R(G) = \frac{\alpha(G)}{|V(G)|},$$

and therefore gives the proportion of vertices in a maximum independent set.

The first result relating the girth of a graph and its independence ratio was provided by Ajtai, Komlós and Szemerédi, who proved in [1] that, if the average degree d of a triangle-free graph on n vertices is sufficiently large, then

$$\alpha_R(G) \geq \frac{\ln d}{100d}.$$

Shearer [49] improved this bound to

$$\alpha_R(G) \geq \frac{d \ln d - d + 1}{(d-1)^2},$$

which now holds for all values of d. Around the same time, Hopkins and Staton [31] showed that, for any $\delta > 0$, if G is a cubic graph on n vertices with girth sufficiently large, it contains an independent set with at least

$$\left(\frac{7}{18} - \delta\right) n$$

vertices. Note that, by a *cubic graph*, we mean a graph whose vertices all have degree three. The fraction in this bound was later improved to 125/302 by Shearer [50], who also gave bounds in the general r-regular graph case in terms of a recursive function $f(r)$. Lauer and Wormald [39] improved Shearer's bound for $r \geq 7$, by showing that, for any $\delta > 0$, r-regular graphs with girth large enough satisfy

$$\alpha_R(G) \geq \frac{1}{2}\left(1 - (r-1)^{-2/(r-2)}\right) - \delta.$$

Their proof method, which consists of analysing a simple algorithm to find an independent set in a graph, will be generalised here, and will therefore be presented in more detail as our contribution is discussed.

In this thesis, we give new lower bounds $\lambda(r)$ on the independence ratio of r-regular graphs with large girth. As with the bounds obtained in [39], the numbers $\lambda(r)$ are derived, for each $r \geq 3$, from the solution to a system of ordinary differential equations associated with r. However, these equations are not solved analytically, and the bounds are calculated separately for each r.

The numbers $\lambda(r)$ have been previously derived in the context of random r-regular graphs by Wormald in [56] (see also [57] for a detailed account, and [58] for an alternative approach). His result establishes that, if an r-regular graph on n vertices is chosen uniformly at random over all r-regular graphs on n vertices, then the probability that it contains an independent set of size at least $\lambda(r)n$ tends to one as n tends to infinity. This is commonly expressed,

in the language of random graphs, as a random r-regular graph on n a.a.s. containing an independent set of size $\lambda(r)n$. (For a sequence of probability spaces Ω_n, $n \geq 1$, an event A_n of Ω_n occurs asymptotically almost surely, or a.a.s. for brevity, if $\lim_{n\to\infty} \mathbf{P}(A_n) = 1$.)

Wormald proves this result through the analysis of a prioritised algorithm, which he calls *degree-greedy*, that finds a large independent set in a graph. This analysis cannot be directly carried over into our class of graphs, since we wish to analyse an algorithm in a fixed graph, and not in a random graph, yet Wormald's degree greedy algorithm, especially in its deprioritised version [58], plays an important role in motivating the choice of parameters to optimise the performance of a locally greedy algorithm for independent sets.

1.3 Induced forests

An *induced forest* in a graph G is an acyclic induced subgraph of G. The term "induced" comes from the notion of an *induced subgraph* of a graph $G = (V, E)$, which is a graph with vertex set $S \subseteq V$ and with edge set consisting of the set of all edges in G with both endpoints in S. The subgraph of G induced by a set $S \subseteq V$ is denoted by $G[S]$. The problem of finding a large induced forest in a graph G has been a widely studied topic in graph theory, especially in its form known as the *decycling set problem* or the *feedback vertex set problem*. A decycling set of a graph is a subset of its vertices whose deletion yields an acyclic graph. From this definition, we deduce that a set $S \subseteq V$ is such that $G[S]$ is an induced forest of $G = (V, E)$ if and only if $V \setminus S$ is a decycling set of G. Therefore, finding a lower bound for $\tau(G)$, the maximum number of vertices in an induced forest of G, amounts to finding an upper bound for $\phi(G)$, the minimum cardinality of a decycling set of G.

The problem of finding a large induced forest or, equivalently, a small decycling set, also has numerous real world applications, as varied as deadlock prevention [53], chip design [35] and genome assembly [45]. Historically, the problem of obtaining an acyclic subgraph of a graph G by removing vertices was already considered by Kirchhoff in his work on spanning trees [37]. Erdős et. al. also worked on this problem stated in terms of maximum induced trees in a graph [26]. As with the independent set problem, however, finding a decycling set of a given size in a graph is among the NP-complete problems in [36]. As a matter of fact, there is no efficient algorithm for this problem, unless $P = NP$, even in special families of graphs such as bipartite graphs, planar graphs or perfect graphs.

On the other hand, there exist polynomial-time algorithms to solve instances of this problem in cubic graphs [40], permutation graphs [41] and interval graphs [42]. Also, tighter bounds or even the exact value of the decycling number have been determined for bipartite graphs in [3] and for graphs such as grids and cubes in [6] and [8].

However, given a graph G on n vertices, there are not many results that bound the maximum number of vertices in an induced forest. Alon et al. showed in [2] that, if the

average degree of G is at most $d \geq 2$, then the maximum cardinality $\tau(G)$ of a subset of vertices inducing an acyclic subgraph of G satisfies

$$\tau(G) \geq \frac{2n}{d+1}.$$

Whenever $d+1$ divides n, this bound is tight, as shown by the graph containing $n/(d+1)$ disjoint copies of the graph K_{d+1}. With Mubayi and Thomas [3], Alon also established that, in the case of triangle-free graphs with maximum degree d,

$$\frac{\tau(G)}{n} = \Omega\left(\frac{\log d}{d}\right),$$

as d tends to infinity, which is tight for some classes of triangle-free graphs. Another result in this paper gives a lower bound on $\tau(G)$ as a function of its order n, its maximum degree Δ and its independence number, namely

$$\tau(G) \geq \alpha(G) + \frac{n - \alpha(G)}{(\Delta - 1)^2},$$

provided that $\Delta \geq 3$.

Results of this type have also been obtained in the probability space of random regular graphs. Indeed, asymptotic upper and lower bounds on the size of a minimum decycling set were given by Bau, Wormald and Zhou in [7]. Recall that the meaning of this result is that, if an r-regular graph on n vertices is chosen uniformly at random over all r-regular graphs on n vertices, then the probability that the proportion of vertices in a largest induced forest in this graph is bounded between the values provided in [7] tends to one as n tends to infinity. As with several bounds for random r-regular graphs, the lower bounds are obtained through the analysis of a randomised algorithm over the probability space of random regular graphs. As we shall see, a simple algorithm for induced forests in regular graphs finds the same numbers to be bounds in the case of graphs with large girth. The class of locally greedy algorithms generalises this algorithm, and, for a particular choice of parameters, implies even better bounds.

We should also mention that, for graphs with girth sufficiently large, the cardinality of a maximum induced forest is connected with the concept of fragmentability in a graph. Given a real $\lambda > 0$ and an integer m, a graph is (λ, m)-*fragmentable* if there is a set $X \subseteq V(G)$ such that $|X| \leq \lambda |V(G)|$ and $Y = V(G) \setminus X$ is m-*fragmented*, that is, every component of $G[Y]$ has at most m vertices. For an integer r, define $\lambda(r)$ as the infimum of λ such that there is an m for which every graph with maximum degree at most r is (λ, m)-fragmentable. A remark in [29] relates, for a random r-regular graph G, the problem of finding the cardinality of a minimum decycling set in G and the problem of finding the infimum of λ such that G is (λ, m)-fragmentable for some m. This relation can be easily restated in the case when G is a graph with large girth and maximum degree r: for any r, g and $\epsilon > 0$, there is an n_0 such

that, for any r-regular graph G of order $n \geq n_0$ and girth larger than g, the size $\phi(G)$ of a minimum decycling set in G satisfies

$$|\phi(G) - n\lambda(r)| < n\epsilon.$$

Thus, the problem of finding bounds on $\lambda(r)$ and on the size of $\phi(G)$ of a minimum decycling set (and hence on the cardinality $\tau(G)$ of a largest induced forest) are equivalent. The best upper bounds on the fragmentability of r-regular graphs with girth sufficiently large,

$$\lambda(r) \leq \frac{r-2}{r+1},$$

have been provided by Edwards and Farr [24]. These bounds give the best possible result for $r = 3$. Note that, for induced forests, this implies the bound $\tau(G)/n \geq 0.75$, where G is a graph on n vertices. However, for all values of $r \geq 4$ determined through our approach, the bounds provided in this thesis are superior.

1.4 Random regular graphs

Based on our discussion about independent sets and induced forests, there seems to be a connection between bounds for r-regular graphs with large girth and asymptotic bounds for random r-regular graphs. We now argue that this is indeed the case. On the one hand, if we determine a lower bound on the proportion of vertices satisfying some property in a graph with girth sufficiently large, then, structural results about random regular graphs proved by Bollobás [13] and Wormald [55] imply that this is also an asymptotic lower bound for random regular graphs, provided that this property is in a quite general class of properties introduced in Chapter 2. In particular, by showing that the property of being an independent set in a graph and the property of inducing a forest in a graph are in this class, we deduce that the bounds obtained in this work for r-regular graphs with large girth hold asymptotically for random r-regular graphs. In particular, this thesis improves the asymptotic bounds on the size of a maximum induced forest in a random r-regular graph, while, in the case of independent sets, it provides a new proof for the bounds first obtained in [56]. We note that this new proof is essentially different from the previous proofs, since it does not rely on sharp concentration arguments.

Less can be said about the converse. It is true that every result establishing that a random r-regular graph a.a.s. satisfies some property implies that, for every $g > 0$, there exists an r-regular graph with girth at least g satisfying this property, since the results in [13] and [55] also imply that the class of r-regular graphs on n vertices with girth at least g are an asymptotically positive proportion of all r-regular graphs on n vertices.

Furthermore, we shall see in this thesis that, although the algorithms used to prove asymptotic lower bounds for random r-regular graph, such as the algorithms in [56] and in [58], in

the case of independent sets, and the algorithm in [7], in the case of induced forests, cannot be analysed when the input graph is fixed, some of their properties will help us determine parameters for which the performance of a locally greedy algorithm is optimal. We should also point out that Wormald and the author of this thesis have work in progress concerning the derivation of a class of results in r-regular graphs with large girth assisted by asymptotic results in the probability space of random r-regular graphs. One such result would be the extension of the lower bounds on the independence number of random r-regular graphs established in [58] to r-regular graphs with large girth, providing an alternative way to derive the bounds obtained in this thesis. Several other bounds in random regular graphs would also be extended to r-regular graphs with large girth through this method, including the bounds on the independent domination number in [22], on small k-dominating sets in [23] and on large k-independent sets [9].

Nevertheless, it is not true that asymptotic properties of random regular graphs always hold in r-regular graphs with girth sufficiently large. For instance, Robinson and Wormald showed in [47] that random r-regular graphs are a.a.s. Hamiltonian, which is obviously not the case for all graphs with girth sufficiently large, since some of these graphs might not even be connected. Furthermore, in joint work with Benjamini, Ofek, Pralat and Wormald [10], the author of this thesis investigated the "shape" of random regular graphs. Among other things, we showed that, if two vertices in a random r-regular graph are chosen uniformly at random, then they are a.a.s. connected to each other through r internally disjoint paths, which obviously does not hold for all r-regular graphs with girth sufficiently large.

1.5 The structure of this thesis

We now describe the organisation of this thesis. The points of original contribution will be emphasised, as will the thesis' connection with previous work in this area.

In Chapter 2, we introduce a simple randomised algorithm to grow an induced forest in an r-regular graph, where $r \geq 3$ is a fixed integer. By estimating the expected size of the set produced by this algorithm, we obtain a lower bound $\xi_1(r)$ on the proportion of vertices in a maximum induced forest in an r-regular graph when the girth of the input graph is sufficiently large. The numbers $\xi_1(r)$ coincide with the numbers obtained in [7] as asymptotic lower bounds on the proportion of vertices in a maximum induced forest in a random r-regular graph. Our approach extends the method initiated by Lauer and Wormald in [39], and the ideas in the analysis of this simple algorithm are useful to illustrate the more general approach in terms of locally greedy algorithms. It should be mentioned that the construction and analysis of this algorithm consists of joint work with Wormald [32].

In Chapter 3, the class of locally greedy algorithms is presented. Recall that the origin of this name is in the fact that, in an algorithm of this class, the random selection of a ver-

tex v occurs with a probability determined by the current state of the vertices within some fixed distance of v. On the one hand, the assignment of particular parameters to this class yields algorithms producing sets with various graph properties, including the algorithm for independent sets analysed in [39] and the algorithm for induced forests analysed in Chapter 2. However, the importance of this class of algorithms goes beyond extending the application of the method from Chapter 2 to other problems. Indeed, it allows us to devise algorithms for which, in a given iteration, the probability that a random choice is made depends, in some sense, on the way that the vertex was affected by the algorithm up to this iteration. In particular, this will lead to an algorithm for independent sets in r-regular graphs whose expected performance in any fixed graph with sufficiently large girth approximates the expected performance of prioritised algorithms in random regular graphs, such as the ones analysed in [56] and [58]. We shall also devise an algorithm for induced forests in r-regular graphs for which the random choices are prioritised, whose analysis leads to an improvement of the lower bounds on the size of a maximum induced forest in an r-regular graph with large girth obtained in Chapter 2.

Despite the general framework, the analysis of a locally greedy algorithm resembles the analysis of the simple algorithm in Chapter 2. Nevertheless, the quality of the bounds obtained depends on the assignment of probabilities to each type of operation performed by the algorithm, and a lot of work is necessary to assign the probabilities so as to get good lower bounds in the case of independent sets and induced forests. In Chapter 4, we define the probabilities associated with a locally greedy algorithm in terms of a general set of functions. Moreover, we impose conditions upon this set so as to ensure that the expected performance of the algorithm can be estimated through the solutions of a system of ordinary differential equations, as was the case in Chapter 2.

In Chapters 5 and 6, we use the work of the previous chapters to provide instances of the locally greedy algorithm that provide the new bounds for independent sets and induced forests. Each of these bounds depends on the solutions of a system of ordinary differential equations that resembles the system obtained in [58] in the context of deprioritised algorithms in random regular graphs, and a good deal of the work in these two chapters consists of verifying that the conditions established in Chapter 4 are satisfied by our particular assignments of probabilities. We observe that, to obtain the numerical bounds given in Table 5.3.1, for independent sets, and in Table 6.3.1, for induced forests, the numerical solutions to these systems of differential equations need to be obtained.

Chapter 2

A simple algorithm in regular graphs

In this chapter, we shall introduce a simple algorithm to find induced forests in a regular graph. We shall extend the method initiated by Lauer and Wormald in [39] to find lower bounds on the size of largest independent sets [39] in such graphs. The proof involves analysing the performance of an iterative randomised algorithm that generates an independent set in a graph. Although their algorithm is applicable to any graph, the number of iterations allowed is bounded by a function that increases with the girth, and, because of this, better bounds can be obtained as the girth increases. We shall use a similar approach to obtain bounds on the size of an induced forest in a graph whose girth is large.

The algorithm that we analyse here is a particular instance of the class of locally greedy algorithms, which will be defined in Chapter 3. This is a simplified version, and its analysis is consequently simpler. Nevertheless, several ingredients in the analysis of the general algorithm will also be presented here. The results in this chapter constitute joint work with Wormald [32].

2.1 Lower bounds on induced forests

For each $r \geq 3$, we shall determine lower bounds on the size $\tau(G)$ of a largest induced forest in an r-regular graph with girth sufficiently large. The bounds are derived from the solution of a system of differential equations that will naturally arise in the analysis of an algorithm for this problem. More precisely, the following theorem will be proved.

Theorem 2.1.1 *Let $\delta > 0$ and $r \geq 3$. Then there exists $g > 0$ such that every r-regular graph G on n vertices with girth greater than or equal to g satisfies $\tau(G) \geq (\xi_1(r) - \delta)n$, where the constants*

$$\xi_1(r) = \sup_{p_0 \in (0,1]} \xi_1(r, p_0)$$

r	3	4	5	6	7	8	9	10
$\xi_1(r)$	0.7268	0.6045	0.5269	0.4711	0.4283	0.3940	0.3658	0.3419

Table 2.1.1: Lower bounds on $\tau(G)/n$, where G is an n-vertex, r-regular graph with sufficiently large girth.

are defined in (2.6.3) with respect to the solution to the initial value problem defined in (2.5.2).

Numerical lower bounds on $\xi_1(r)$ are given in Table 2.1.1 for some values of r.

Before discussing the proof of this theorem in more detail, we shall see that bounds for r-regular graphs with large girth can sometimes be translated into bounds for graphs with large girth and maximum degree r, and even into asymptotic bounds for random r-regular graphs. This is true if the graph property P under consideration is a property defined over subsets of the vertex sets of graphs such that, if $S \subseteq V(G) \cap V(H)$ satisfies property P in G and $G[S] = H[S]$, then S satisfies P in H. Moreover, if S satisfies P in G, then every subset of S satisfies P in G. Such a property P will be called a *vertex monotone property*. Note that the property of being an induced forest in a graph is vertex monotone. Indeed, if S lies in the intersection of the vertex sets of graphs G and H, and if we know that S induces an acyclic subgraph of G and $G[S] = H[S]$, then the fact that S induces a forest in H is trivial. Moreover, it is also clear that, if S induces a forest in G, then every subset of S must do the same.

The following result shows the connection between bounds for r-regular graphs with large girth and bounds for graphs with large girth and maximum degree r in this case.

Lemma 2.1.2 *Let $r \geq 3$ be a positive integer and let $\gamma > 0$. Consider a vertex monotone property P, and suppose that there exists $g > 0$ such that every r-regular graph G with girth at least g contains a set of vertices S with cardinality $|S| \geq \gamma|V(G)|$ satisfying property P. Then, for any graph H with maximum degree r and girth at least g, there exists $T \subseteq V(H)$ such that $|T|$ satisfies P and $|T| \geq \gamma|V(H)|$.*

Proof Let $g > 0$ be such that every r-regular graph G with girth at least g contains a set of vertices S with cardinality $|S| \geq \gamma|V(G)|$ satisfying property P, and let H be a graph with maximum degree r and with girth at least g. It is easy to see that we may create an r-regular graph H' by taking copies of H and joining some pairs of vertices from different copies so as to make the resulting graph H' r-regular. This can be done without decreasing the girth if sufficiently many copies of H are used.

Now, by hypothesis, the graph H' contains a set S' that satisfies property P and has cardinality at least $\gamma|V(H')|$. Let H_1 be one of the copies of H in H' with most vertices in S', and define $S_1 = S' \cap V(H_1)$. By our definition of S_1, it is clear that $|S_1| \geq \gamma|V(H)|$. On the other hand, S_1 satisfies property P in H', since it is contained in the set S', which

satisfies this property in H'. By construction, we also have $H[S] = H'[S]$, so that, using vertex monotonicity, we deduce that S is a subset of H with cardinality at least $\gamma|V(H)|$ satisfying property P, as required. □

In particular, the bounds $\xi_1(r)$ given in the statement of Theorem 2.1.1 also hold for graphs with maximum degree r and girth sufficiently large. A similar result can be obtained for random regular graphs. It relies on the following result in the theory of random regular graphs. For a proof, see Wormald [55] or Bollobás [13].

Lemma 2.1.3 *Let $r \geq 3$ and g be fixed positive integers. Then a random r-regular graph G on n vertices a.a.s. contains $o(n)$ cycles of length at most g.*

We can now find a connection between bounds for r-regular graphs with large girth and asymptotic bounds for random regular graphs.

Lemma 2.1.4 *Let $r \geq 3$ be a positive integer and let $\gamma > 0$. Consider a vertex monotone property P and suppose that there exists $g > 0$ such that every r-regular graph G with girth at least g contains a set of vertices S with cardinality at least $\gamma|V(G)|$ satisfying property P. Then a random r-regular graph on n vertices a.a.s contains a set of vertices T satisfying P such that $|T| \geq \gamma\,(n - o(n))$.*

Proof Consider a vertex monotone property P and suppose that there exists $g > 0$ such that every r-regular graph G with girth at least g contains a set of vertices S that satisfies property P and has cardinality at least $\gamma|V(G)|$.

By Lemma 2.1.3, G a.a.s. can be turned into a graph G' with maximum degree r and girth at least g by deleting $o(n)$ of its vertices. By Lemma 2.1.2, G' contains a set S satisfying P and with cardinality at least $\gamma\,(n - o(n)) \geq \gamma(n - \epsilon)$ vertices. Since G' was obtained from G through the deletion of vertices, we must have $G'[S] = G[S]$, and, therefore, S satisfies property P in G, establishing the lemma. □

Our main objective in this chapter is to establish Theorem 2.1.1, whose proof is structured as follows. We first introduce a randomised greedy algorithm that finds an induced forest of a given graph. When this algorithm is applied to an r-regular graph G with sufficiently large girth, its expected performance leads to the bounds in the statement of the theorem, and hence guarantees the existence of an induced forest on the same proportion of vertices by the first moment principle. To estimate the expected performance of the algorithm, we shall establish preliminary lemmas that help us understand the behaviour of our algorithm, which will then be used to derive a system of recurrence equations involving the cardinality of the set of vertices in the induced forest. Finally, we shall approximate this system of recurrence equations by a system of ordinary differential equations whose solution provides us with the bounds in the statement of Theorem 2.1.1.

Our method also produces (weaker) bounds on $\tau(G)$ if a specific lower bound on the girth of G is given. However, we do not compute the precise constants for any particular bound on

girth.

2.2 An algorithm for induced forests

We now introduce an algorithm that will help us find a large induced forest in a graph $G = (V, E)$. At any given step of the algorithm, we shall associate colours with the vertices of the graph as follows. The colour purple is assigned to vertices in a set P, a subset of V that induces a subgraph of G with "a few" cycles only. A vertex is blue if it is not yet in P, but could join it in the next iteration, whereas yellow is assigned to vertices whose addition to P would yield cycles in $G[P]$. The remaining vertices are coloured white and are the vertices not adjacent to vertices of the forest.

Algorithm 2.2.1 ***(Simple algorithm)***

Input: A graph G, a positive integer N and a pair of probabilities (p_0, p).

1. *Start with all the vertices of the graph coloured white. In the first step, colour each vertex purple with probability p_0, at random, independently of all others. Non-purple vertices are coloured blue if they have exactly one purple neighbour and yellow if they have at least two purple neighbours.*

2. *At each step i, choose blue vertices randomly and independently with probability p and colour them purple. The sets of blue and yellow vertices are updated using the rule given in 1. We refer to the set of white vertices as W, to the set of blue vertices as B and to the set of purple vertices as P. Repeat this iteratively for N steps.*

3. *Create a set $\bar{P} \subseteq P$ by deleting any pair of adjacent vertices added to P in the same step.*

Output: The acyclic set $\bar{P}$ and the set of white vertices $\bar{W}$ lying in acyclic components of $G[W]$.

In the first phase, the roots of the induced trees are chosen and coloured purple, and vertices that could be added to the trees without creating cycles or connecting distinct components are coloured blue. In each step of the second phase, the forest is extended by choosing blue vertices and adding them to P, and at each step the colours associated with each vertex are updated so that the sets of white, blue and yellow vertices at the end of each step represent the vertices with 0, 1, and more than one, purple neighbours, respectively. Note that it would be possible to alter p at each step, and this would be useful if optimising the algorithm for the set of graphs with particular girth (as done in [39] for independent sets), but we do not do this here.

The graph $G[P]$ at the end of Phase 2 is not necessarily acyclic. As a matter of fact, it may happen that two neighbouring blue vertices are added to the forest in the same step and create a cycle. These cycles are broken in the third phase of the algorithm.

A drawback to the analysis of Algorithm 2.2.1 in its original version is that the random selection of vertices at a given step depends on the outcome of the previous steps. To avoid this, we introduce an equivalent model for which the random choices are uniform over the whole set of vertices. Indeed, with each vertex $v \in V$, we shall first associate a random sequence of labels $S(v) \subseteq \{0, 1, 2, \ldots\}$ so that label i is in $S(v)$ independently at random with probability p_0, if $i = 0$, or p, if $i \geq 1$. In other words, we choose sets of vertices at times 0, 1, ..., and assign to a vertex v the labels $\{i : \ v \text{ was chosen at time } i\}$. In the context of our algorithm, we shall then consider the set of vertices with label 0 to be the set of vertices selected in Phase 1 and use vertices with label i to recreate the set of vertices added to P at step i in Phase 2 of our algorithm. It is clear that some of the labels are ill-suited. For instance, a vertex with label 1 will not be selected to join P at step 1 if it also has label 0, in which case it already belongs to P, or if none of its neighbours has label 0, in which case it is not blue after the first phase of the algorithm. This motivates a classification of the labels as relevant or irrelevant, that is, as labels that represent an action of our algorithm or as labels that do not.

Definition 2.2.2 ***(Relevant and irrelevant labels)***
Let $G = (V, E)$ be a graph, and, for every $v \in V$, let $S(v) \subseteq \mathbb{N}$ be the set of labels associated with v. We define relevant labels inductively (labels that are not relevant are said to be irrelevant). A label i is relevant for v if:
I. $i = 0 \in S(v)$, or
II. $i \in S(v)$, j is irrelevant for v for all $j < i$, and there is a unique neighbour of v with a relevant label strictly smaller than i.

The sets of vertices with relevant label equal to i are denoted by R_i, while the ones with relevant label less than or equal to i are denoted by $R_{\leq i}$. We refer to the sequence $[S(v) : v \in V]$ as $\mathcal{S}$. Now, for each $\ell \in \mathbb{N}$, the sequence $\mathcal{S}$ may be used to construct a colouring of G with colours purple, blue, white and yellow.

Definition 2.2.3 ***(Colouring of G at time ℓ)***
Given a graph G and a sequence $\mathcal{S}$ as above, the colouring of G at time $\ell \in \mathbb{N}$ is the function assigning colours purple, blue, yellow and white to the vertices of G defined as follows. Given $u \in V$,

(a) u is white if $u \notin R_{\leq l}$ and $v \notin R_{\leq l}$, for all $v \in N(u)$, where $N(u)$ denotes the neighbourhood of u.

(b) u is blue if $u \notin R_{\leq l}$ and there is a unique $v \in N(u)$ such that $v \in R_{\leq l}$.

(c) u is yellow if $u \notin R_{\leq l}$ and there exist distinct $v, w \in N(u)$ with $v, w \in R_{\leq l}$.

(d) u is purple if $u \in R_{\leq l}$.

It is clear from this definition that the colouring of G at time ℓ is fully determined by the sequence $[S(v) \cap \{0, \ldots, l\} : v \in V]$. Moreover, this colouring coincides with the colouring of the graph induced by our algorithm if we assume the set P after k steps to be $R_{\leq k}$, as formalised by the next lemma.

Lemma 2.2.4 *Let $G = (V, E)$ be a graph, and consider a subgraph H of G and a colouring c of H with colours purple, blue, yellow and white. Then the following events have the same probability:*

(i) the colouring of G at time ℓ induced by the sequence $\mathcal{S} = [S(v) : v \in V(G)]$ restricted to H is equal to c, where $\mathcal{S}$ is obtained by adding each nonnegative integer i to $S(v)$ independently with probability p_0, if $i = 0$, or p, if $i \geq 1$, for all $v \in V$.

(ii) Algorithm 2.2.1 applied to G obtains c as the colouring of H after step ℓ.

Proof We create a new algorithm by modifying part 2 of Algorithm 2.2.1 as follows. At each step i, every vertex $v \in V(G)$ is chosen with probability p. However, the algorithm takes no action if a non-blue vertex is selected. So, the modified algorithm makes more random choices, but the random choices that do not correspond to a random choice in the original algorithm are irrelevant. So, these two algorithms clearly obtain c as the colouring of H after step ℓ with the same probability. On the other hand, a simple inductive argument implies that the event that, after step ℓ, the new algorithm yields colouring c in a subgraph H of G corresponds to the event described in (i), establishing the lemma. ∎

Given this equivalence, we shall work in the probability space of the sequence $\mathcal{S}$ of sets of labels, which we call the *probability space of labellings*.

2.3 Independence lemmas

We prove results that allow us to compute the probability, using local information only, of a vertex of an r-regular graph G being assigned some given colour at time i. Henceforth, we shall fix an r-regular graph $G = (V, E)$ with girth g and consider a sequence of sets $\mathcal{S} = [S(v) : v \in V(G)]$, where $i \in \mathbb{N}$ is in $S(v)$ with probability p_0, if $i = 0$, or p, if $i \geq 1$, for all v.

Lemma 2.3.1 *Let $G = (V, E)$ be a graph and consider a sequence of sets of labels $\mathcal{S} = [S(v) : v \in V]$. Given $u \in V$, define a sequence of sets of labels $\mathcal{S}'$ by replacing, in $\mathcal{S}$, $S(u)$ by some*

set $S'(u)$. Let w be a vertex of G whose colour at time i with respect to $\mathcal{S}$ and $\mathcal{S}'$ differs, where i is a nonnegative integer.

Then there exists a path $\mathcal{P}$ from u to w for which every vertex except possibly w gained or lost a relevant label less than or equal to i when $\mathcal{S}$ was replaced by $\mathcal{S}'$. Moreover, the relevant labels gained or lost by each vertex along the path are in strictly increasing order when the path is considered from u to w.

Proof The proof is by induction on i. For $i = 0$, since the colour of w at time 0 has changed after replacing $S(u)$ by $S'(u)$, it must be that u has gained or lost relevant label 0 and that either $u = w$ or u and w are neighbours. In both cases, $\mathcal{P} = (u, w)$ satisfies the conditions in the statement of this lemma.

Now, let $i > 0$ and assume that this result holds earlier. If $u = w$ nothing needs to be done, so suppose that this is not the case. Since the colour of w changed at time i, there exists a neighbour w' of w that gained or lost a relevant label smaller than or equal to i. If $w' = u$, our result is clearly true, so suppose that they are distinct. Then the relevant label gained or lost by w' is not equal to 0 and, by the definition of relevant label, there is a neighbour w'' of w' that gained or lost a relevant label at a time j strictly smaller than the relevant label gained or lost by w'. In particular, the colour of w'' changed at time j, so, by induction, there is a path $\mathcal{P}''$ from u to w'' under the conditions of the lemma. Thus, the path $\mathcal{P}$ obtained by appending vertices w' and u to $\mathcal{P}''$ satisfies the required properties. ∎

Corollary 2.3.2 *Let $u \in V(G)$ and $i \in \mathbb{N}$. Then for any given colour c and any collection of subsets S'_v of $\mathbb{N}$, where v ranges over the vertices at distance at least $i + 2$ of u, the event that u has colour c at time i is independent of the event that $S(v) = S'_v$.*

Proof It is sufficient to show that, if $\hat{\mathcal{S}} = [\hat{S}(v) : v \in V(G)]$ is any given family of sets of labels and new sets $S'(v)$ are assigned to each vertex v satisfying $d(u, v) \geq i + 2$, then the colour of u at time i relative to $\hat{\mathcal{S}}$ is the same as the colour of u at time i relative to $\mathcal{S}'$, where $\mathcal{S}'$ is obtained by replacing each $\hat{S}(v)$ by $S'(v)$.

We now prove this sufficient condition. Suppose for a contradiction that the colours of u with respect to $\hat{\mathcal{S}}$ and $\mathcal{S}'$ differ, and order the vertices $v \in V$ satisfying $d(u, v) \geq i + 2$ as $v_1, v_2, \ldots, v_m$. Consider, for $l \in \{0, \ldots, m\}$, the sequences $\mathcal{S}_l$ obtained from $\hat{\mathcal{S}}$ by replacing $\hat{S}(v_1), \ldots, \hat{S}(v_l)$ by $S'(v_1), \ldots, S'(v_l)$. Our assumption implies the existence of j such that the colours of u with respect to $\mathcal{S}_j$ and $\mathcal{S}_{j+1}$ are distinct. By Lemma 2.3.1, there is a path $\mathcal{P}$ in G from v_{j+1} to u such that every vertex except possibly u gained or lost a relevant label less than or equal to i when $\mathcal{S}_j$ was replaced by $\mathcal{S}_{j+1}$. Also, the relevant labels gained or lost on each vertex along the path are in strictly increasing order when the path is considered from v_{j+1} to u. In particular, $\mathcal{P}$ contains at most $i + 2$ vertices, i.e., $d(u, v_{j+1}) \leq i + 1$, a contradiction. ∎

Let B_i and W_i denote the sets of vertices coloured blue and white at time i, respectively.

For events E_1 and E_2, let $E_1 \wedge E_2$ denote their intersection.

Corollary 2.3.3 *Let $u \in V$ and let v be one its neighbours. Then the probabilities $\mathbf{P}(u \in W_i)$, $\mathbf{P}(u \in B_i)$, $\mathbf{P}(u \in W_i \wedge v \in W_i)$, $\mathbf{P}(u \in B_i \wedge u \in W_i)$ and $\mathbf{P}(u \in B_i \wedge v \in B_i)$ are independent of u and v whenever $2i < g - 3$. Moreover, if we let w be a neighbour of u distinct from v, $\mathbf{P}(v \in B_i \wedge u \in B_i \wedge w \in R_{\leq i})$ does not depend on u, v or w.*

Proof We know from Corollary 2.3.2 that the colour of u at time i depends only on the sets of labels of vertices at distance at most $i+1$ from u. In other words, u is fully determined by the sets of labels in the subgraph $G_u = G[\{v : d(u,v) \leq i+1\}]$. But our restriction on i implies that, for every $u \in V$, the graphs G_u are isomorphic. Our first two claims immediately follow, since distinct vertices are assigned sets of labels independently with the same probability. It is clear that an analogous argument can be used to prove the remaining statements. ∎

Before stating the next result, a definition is necessary.

Definition 2.3.4 ***(Branches around vertex** u**)** *Let $u \in V$ and let $u_1, \ldots, u_r$ be its neighbours. For each $s \in \{1, \ldots, r\}$ and positive integer $m < g/2$, the component $T_{u,s,m}$ of $G[\{v : d(u,v) \leq m\} \setminus u]$ containing u_s is a tree, which we consider as a rooted tree with root u_s. We shall refer to these trees as branches around vertex u.*

In the next result, we shall use the concept of mutual independence of events. By a collection of events $H_1, ..., H_m$ being *mutually independent*, it is meant that, for any subset of the collection, the joint probability of all events occurring is equal to the product of the probabilities of the individual events.

Lemma 2.3.5 *Let $u \in V$ and let $v_1, \ldots, v_r$ be its neighbours. Fix $i \in \mathbb{N}$ such that $2(i+2) < g-2$ and consider, for each $j \in \{1, \ldots, r\}$, the tree $T_{u,j,2}$ rooted at v_j given by the component of $G[\{v : d(u,v) \leq 2\} - u]$ containing v_j. Then the following assertions hold.*
1. Let $X_1, \ldots, X_r$ be colourings of the tree isomorphic to the rooted trees $T_{u,j,2}$ (the isomorphism is a consequence of our restriction on i). Then, conditional upon $u \in W_i$, the events $E_1, \ldots, E_r$ are mutually independent, where E_j stands for the event that $T_{u,j,2}$ has colouring X_j at time i.
2. Conditional upon $u \in B_i$ and $v_l \in R_{\leq i}$ for some $l \in \{1, \ldots, r\}$, the same events E_j are mutually independent for all $j \neq l$.

Proof By Lemma 2.3.1, the colour of a vertex w at time i is altered when replacing $S_1(v) \times \cdots \times S_r(v)$ by $S'_1(v) \times \cdots \times S'_r(v)$ only if there is a path $\mathcal{P}$ from v to w such that all the vertices on $\mathcal{P}$ that are not purple at time i with respect to $\mathcal{S}$ have a different colour with respect to $\mathcal{S}'$. This is because, given any non-purple vertex w' at time i lying on $\mathcal{P}$, it either gains a relevant label, in which case it is purple at time i with respect to $\mathcal{S}'$, or it is equal to w, in which case its colour changes by assumption.

Now, if w and v are vertices in different branches with respect to u, our restriction on i implies by Corollary 2.3.2 that any path from v to w that is short enough for every interior vertex to gain or lose a relevant label passes through u. Hence, conditional upon u being white, changes in $S_1(v) \times \cdots \times S_r(v)$ do not affect the colour of w at time i.

Moreover, we also know by Corollary 2.3.2 that the colour at time i of vertices at distance at most two from u are not affected by changes in the set of labels of vertices whose distance to u is greater that $i+3$. Let $V_{u,i+3}$ be the set of vertices in G at distance at most $i+3$ from u, excluding vertex u.

So,

$$\mathbf{P}(E_1 \wedge E_2 \wedge \ldots \wedge E_r \mid u \in W_i) = \sum_{\star} \mathbf{P}(S(v) = S_v, \forall v \in V_{u,i+3} \mid u \in W_i),$$

where $\sum_\star$ denotes the sum over vectors $(S_v : v \in V_{u,i+3})$ such that the event $S(v) = S_v,\ \forall v \in V_{u,i+3}$, implies $E_1 \wedge E_2 \wedge \ldots \wedge E_r$. Now, observe that our restriction on i implies that the branches $T_{u,j,i+3}$ are all disjoint. In particular, we can first sum over sets of labels of vertices in $T_{u,1,i+3}$ (notation $\sum_{\star\star}$) and then over the remaining vertices (notation $\sum_{\star\star\star}$) to obtain

$$\mathbf{P}(E_1 \wedge E_2 \wedge \ldots \wedge E_r \mid u \in W_i) = \sum_{\star\star} \sum_{\star\star\star} \mathbf{P}(S(v) = S_v, \forall v \in V_{u,i+3} \mid u \in W_i).$$

Using conditional probability and rearranging the sum, this becomes

$$\begin{aligned}
&\sum_{\star\star} \mathbf{P}(S(v) = S_v, \forall v \in T_{u,1,i+3} \mid u \in W_i) \\
&\quad \times \sum_{\star\star\star} \mathbf{P}(S(v) = S_v, \forall v \notin T_{u,1,i+3} \mid (u \in W_i) \wedge (S(v) = S_v, \forall v \in T_{u,1,i+3})) \\
&= \sum_{\star\star} \mathbf{P}(S(v) = S_v, \forall v \in T_{u,1,i+3} \mid u \in W_i)\ \mathbf{P}(E_2 \wedge \ldots \wedge E_r \mid u \in W_i) \\
&= \mathbf{P}(E_1 \mid u \in W_i)\ \mathbf{P}(E_2 \wedge \ldots \wedge E_r \mid u \in W_i).
\end{aligned}$$

These manipulations can be done since, conditional upon u being white, changes in $S(v)$ do not affect the colours of other branches, for any $v \in T_{u,1,i+3}$.

Repeating this argument for the remaining branches, we obtain

$$\mathbf{P}(E_1 \wedge E_2 \wedge \ldots \wedge E_r \mid u \in W_i) = \prod_{j=1}^{r} \mathbf{P}(E_j \mid u \in W_i),$$

and our first claim is true.

For the second part, we proceed analogously by leaving both the blue vertex u and the branch of its neighbour with relevant label untouched, and then summing over all possibilities of labels for vertices in the other branches. ∎

2.4 Applications of the independence lemmas

In this section, the independence results of the previous section will be used to obtain recurrence equations relating the probabilities of events that are important in the analysis of Algorithm 2.2.1. We introduce some notation. Let u be a vertex of graph G. An arbitrary neighbour of u will be denoted by v, while we use $v_1, \ldots, v_r$ to refer to the set of neighbours of u. When u has a neighbour with relevant label, this will be referred as v_k and we shall assume that $v \neq v_k$.

Furthermore, for any $i \geq 0$, we know by Corollary 2.3.3 that the quantities $w_i = \mathbf{P}(u \in W_i)$, $b_i = \mathbf{P}(u \in B_i)$, $q_i = \mathbf{P}(u \in W_i \wedge v \in W_i)$, $s_i = \mathbf{P}(u \in B_i \wedge v \in W_i)$ and $t_i = \mathbf{P}(u \in B_i \wedge v \in B_i)$, or even $\mathbf{P}(u \in B_i \wedge v \in W_i \wedge v_k \in R_{\leq i})$ and $\mathbf{P}(u \in B_i \wedge v \in B_i \wedge v_k \in R_{\leq i})$, do not depend on u, v or k. We now let $1 \leq i < (g-3)/2$ and establish the following consequences of the previous independence lemmas.

Corollary 2.4.1 *The following equations hold in the probability space of labellings.*

(i) Let $J = \{j_1, \ldots, j_m\} \subseteq \{1, \ldots, r\}$. Then

$$\mathbf{P}(v_j \notin R_i,\ \forall j \in J \mid u \in W_{i-1}) = \prod_{j \in J} \mathbf{P}(v_j \notin R_i \mid u \in W_{i-1}).$$

(ii) Let $J \subseteq \{1, \ldots, r\} \setminus \{k\}$. Then

$$\mathbf{P}(v_j \notin R_i,\ \forall j \in J \mid u \in B_{i-1} \wedge v_k \in R_{\leq i-1}) = \prod_{j \in J} \mathbf{P}(v_j \notin R_i \mid u \in B_{i-1} \wedge v_k \in R_{\leq i-1}).$$

Proof We prove part (i) by induction on m. For $m = 1$, the result follows immediately, so let $m > 1$ and assume the result holds for any smaller set J.

First observe that, because a vertex receives relevant label $i \geq 1$ only if it is blue at time $i-1$, it is important to consider the set of blue neighbours of u at time $i-1$. In light of this, we associate a vector $\omega \in \mathbb{Z}_2^m$ with the set of neighbours v_{j_t} of u so that $\omega(t) = 1$ if and only if $v_{j_t} \in B_{i-1}$.

Note that, for a vertex not to become purple at time i, it either was not blue at the previous step or it was blue, but i is not contained in its set of labels. Thus,

$$\begin{aligned}&\mathbf{P}(v_j \notin R_i,\ \forall j \in J \mid u \in W_{i-1})\\ &= \sum_{\omega \in \mathbb{Z}_2^m} \mathbf{P}((v_{j_t} \in B_{i-1} \wedge i \notin S(v_{j_t}),\ \text{if } \omega(t) = 1) \wedge (v_{j_t} \notin B_{i-1},\ \text{if } \omega(t) = 0) \mid u \in W_{i-1}).\end{aligned}$$

The fact that $S(v)$ contains any nonnegative integer independently at random (and label i does not influence the colouring at time $i-1$), together with Lemma 2.3.5, ensures that

the events of the form $(v_{j_t} \in B_{i-1} \wedge i \notin S(v_{j_t}))$ and $v_{j_s} \notin B_{i-1}$ are mutually independent conditional upon u being white. So, the equation becomes

$$\begin{aligned}
&\mathbf{P}(v_j \notin R_i,\ \forall j \in J \mid u \in W_{i-1}) \\
&\quad = \sum_{\omega \in \mathbb{Z}_2^m} \prod_{\{j_t:\omega(t)=1\}} \mathbf{P}(v_{j_t} \in B_{i-1} \wedge i \notin S(v_{j_t}) \mid u \in W_{i-1}) \prod_{\{j_t:\omega(t)=0\}} \mathbf{P}(v_{j_t} \notin B_{i-1} \mid u \in W_{i-1}). \\
&\quad = \mathbf{P}(v_{j_m} \in B_{i-1} \wedge i \notin S(v_{j_m}) \mid u \in W_{i-1}) \sum_{\omega' \in \mathbb{Z}_2^{m-1}} \prod_{\{j_t:\omega'(t)=0\}} \mathbf{P}(v_{j_t} \notin B_{i-1} \mid u \in W_{i-1}) \\
&\qquad\qquad \times \prod_{\{j_t:\omega'(t)=1\}} \mathbf{P}(v_{j_t} \in B_{i-1} \wedge i \notin S(v_{j_t}) \mid u \in W_{i-1}) \\
&\qquad + \mathbf{P}(v_{j_m} \notin B_{i-1} \mid u \in W_{i-1}) \sum_{\omega' \in \mathbb{Z}_2^{m-1}} \prod_{\{j_t:\omega'(t)=0\}} \mathbf{P}(v_{j_t} \notin B_{i-1} \mid u \in W_{i-1}) \\
&\qquad\qquad \times \prod_{\{j_t:\omega'(t)=1\}} \mathbf{P}(v_{j_t} \in B_{i-1} \wedge i \notin S(v_{j_t}) \mid u \in W_{i-1}) \\
&\quad = \mathbf{P}(v_{j_m} \in B_{i-1} \wedge i \notin S(v_{j_m}) \mid u \in W_{i-1}) \mathbf{P}(v_j \notin R_i,\ \forall j \in J \setminus \{m\} \mid u \in W_{i-1}) \\
&\qquad + \mathbf{P}(v_{j_m} \notin B_{i-1} \mid u \in W_{i-1}) \mathbf{P}(v_j \notin R_i,\ \forall j \in J \setminus \{m\} \mid u \in W_{i-1})
\end{aligned}$$

By induction, this is equal to

$$\begin{aligned}
&\mathbf{P}(v_{j_m} \in B_{i-1} \wedge i \notin S(v_{j_m}) \mid u \in W_{i-1}) \prod_{j \in J \setminus \{m\}} \mathbf{P}(v_j \notin R_i \mid u \in W_{i-1}) \\
&\qquad + \mathbf{P}(v_{j_m} \notin B_{i-1} \mid u \in W_{i-1}) \prod_{j \in J \setminus \{m\}} \mathbf{P}(v_j \notin R_i \mid u \in W_{i-1}) \\
&\quad = \prod_{j \in J} \mathbf{P}(v_j \notin R_i \mid u \in W_{i-1}),
\end{aligned}$$

as required for (i).

An analogous argument gives (ii). ∎

Remark 2.4.1 *This corollary can also be extended to conditioning upon $u \in W_{i-1} \wedge v \in W_{i-1}$, where u, v are neighbours in G (or any other combination of restrictions on u, v being white or blue). As a matter of fact, if $u_1, \ldots, u_{r-1}$, $v_1, \ldots, v_{r-1}$ denote the neighbours of u, v distinct from u and v, and $J, K \subseteq \{1, \ldots, r-1\}$, then*

$$\begin{aligned}
&\mathbf{P}((u_j \notin R_i,\ \forall j \in J) \wedge (v_m \notin R_i,\ \forall m \in K) \mid u \in W_{i-1} \wedge v \in W_{i-1}) \\
&\quad = \prod_{j \in J} \mathbf{P}(u_j \notin R_i \mid u \in W_{i-1}) \prod_{m \in K} \mathbf{P}(v_m \notin R_i \mid v \in W_{i-1}).
\end{aligned}$$

This can be obtained by expanding the initial probability into a sum over vectors $\omega \in \mathbb{Z}_2^{|J|+|K|}$ and then using the fact that, for any event E, we have

$$\mathbf{P}(E \mid u \in W_{i-1} \wedge v \in W_{i-1}) = \frac{\mathbf{P}(E \wedge u \in W_{i-1} \mid v \in W_{i-1})}{\mathbf{P}(u \in W_{i-1} \mid v \in W_{i-1})},$$

so that Lemma 2.3.5 can be applied first with respect to $u \in W_{i-1}$ and then with respect to $v \in W_{i-1}$. It is clear that similar results can be stated by conditioning upon other combinations of u and v being white or blue.

Corollary 2.4.2 *The following formulae hold in the probability space of labellings.*

(i) $\mathbf{P}(u \in W_i \mid u \in W_{i-1}) = \left(1 - \frac{ps_{i-1}}{w_{i-1}}\right)^r$

(ii) $\mathbf{P}(u \in B_i \mid u \in W_{i-1}) = \frac{rps_{i-1}}{w_{i-1}} \left(1 - \frac{ps_{i-1}}{w_{i-1}}\right)^{r-1}$

Proof For (i), we just observe that, for u to cease to be white at time i, at least one of its neighbours has relevant neighbour i. Thus,

$$\mathbf{P}(u \in W_i \mid u \in W_{i-1}) = \mathbf{P}(v_j \notin R_i,\ \forall j \mid u \in W_{i-1}).$$

Now, by Corollary 2.4.1, part (i), this last expression is equal to

$$\prod_{j=1}^{r} \mathbf{P}(v_j \notin R_i \mid u \in W_{i-1}).$$

Finally, Corollary 2.3.3 guarantees that the probability of v_j having relevant label i is independent of v_j and equals the probability of the event that $i \in S(v_j)$ and v_j is coloured blue at time $i-1$. So,

$$\mathbf{P}(v_j \in R_i \mid u \in W_{i-1}) = \frac{ps_{i-1}}{w_{i-1}},$$

and

$$\mathbf{P}(u \in W_i \mid u \in W_{i-1}) = \left(1 - \frac{ps_{i-1}}{w_{i-1}}\right)^r$$

as a consequence.

Assertion (ii) may be proven using a similar approach. ∎

Corollary 2.4.3 $\mathbf{P}(u \in B_i \mid u \in B_{i-1}) = (1-p)\left(1 - \frac{rpt_{i-1}}{(r-1)b_{i-1}}\right)^{r-1}$

Proof The fact that u is blue at step $i-1$ implies that exactly one of its neighbours $v_1, \ldots, v_r$ has a relevant label less than or equal to $i-1$. Thus,

$$\mathbf{P}(u \in B_i \mid u \in B_{i-1}) = \sum_{m=1}^{r} \mathbf{P}(u \in B_i \mid u \in B_{i-1} \wedge v_m \in R_{\leq i-1}) \mathbf{P}(v_m \in R_{\leq i-1} \mid u \in B_{i-1}).$$

Moreover, u remains blue at time i if neither itself nor any of its neighbours gains a relevant label at time i, i.e.,

$$\mathbf{P}(u \in B_i \mid u \in B_{i-1} \wedge v_m \in R_{\leq i-1}) = (1-p)\mathbf{P}(v_j \notin R_i,\ \forall j \neq m \mid u \in B_{i-1} \wedge v_m \in R_{\leq i-1}).$$

By Corollary 2.4.1, part (ii), we obtain

$$\begin{aligned}&\mathbf{P}(v_j \notin R_i,\ \forall j \neq m \mid u \in B_{i-1} \wedge v_m \in R_{\leq i-1})\\ &\quad = \prod_{j \neq m} \mathbf{P}(v_j \notin R_i \mid u \in B_{i-1} \wedge v_m \in R_{\leq i-1})\\ &\quad = (1 - p\,\mathbf{P}(v_j \in B_{i-1} \mid u \in B_{i-1} \wedge v_m \in R_{\leq i-1}))^{r-1}.\end{aligned}$$

The last equality follows from the fact that $v \in R_i$ only if it has label i and was blue at time $i-1$.

Finally, we note that

$$\begin{aligned}\mathbf{P}(v_j \in B_{i-1} \mid u \in B_{i-1} \wedge v_m \in R_{\leq i-1}) &= \frac{\mathbf{P}(v_j \in B_{i-1} \wedge u \in B_{i-1} \wedge v_m \in R_{\leq i-1})}{\mathbf{P}(u \in B_{i-1} \wedge v_m \in R_{\leq i-1})}\\ = \frac{\mathbf{P}(v_j \in B_{i-1} \wedge u \in B_{i-1})\,\mathbf{P}(v_m \in R_{\leq i-1} \mid v_j \in B_{i-1} \wedge u \in B_{i-1})}{\mathbf{P}(u \in B_{i-1})\,\mathbf{P}(v_m \in R_{\leq i-1} \mid u \in B_{i-1})}&.\end{aligned}$$

By Corollary 2.3.3, we conclude that all the neighbours of u have the same probability of having a relevant label earlier than the other neighbours, since the probability of having relevant label i is equal to p_0, if $i = 0$, or pb_{i-1}, if $i \geq 1$, for any vertex. In particular, we must have $\mathbf{P}(v_m \in R_{\leq i-1} \mid u \in B_{i-1}) = \frac{1}{r}$ and $\mathbf{P}(v_m \in R_{\leq i-1} \mid v_j \in B_{i-1} \wedge u \in B_{i-1}) = \frac{1}{r-1}$. So,

$$\begin{aligned}\mathbf{P}(u \in B_i \mid u \in B_{i-1}) &= (1-p)\sum_{m=1}^{r} \mathbf{P}(v_m \in R_{\leq i-1} \mid u \in B_{i-1}) \left(1 - \frac{rpt_{i-1}}{(r-1)b_{i-1}}\right)^{r-1}\\ &= (1-p)\left(1 - \frac{rpt_{i-1}}{(r-1)b_{i-1}}\right)^{r-1},\end{aligned}$$

with the last equation following from $\sum_{m=1}^{r} \mathbf{P}(v_m \in R_{\leq i-1} \mid u \in B_{i-1}) = 1$. This concludes the proof. ∎

Corollary 2.4.4 *The following equations hold in the probability space of labellings.*

(i) $\mathbf{P}(u \in W_i \wedge v \in W_i \mid u \in W_{i-1} \wedge v \in W_{i-1}) = \left(1 - \frac{ps_{i-1}}{w_{i-1}}\right)^{2r-2}$,

(ii) $\mathbf{P}(u \in W_i \wedge v \in B_i \mid u \in W_{i-1} \wedge v \in W_{i-1}) = \frac{(r-1)ps_{i-1}}{w_{i-1}}\left(1 - \frac{ps_{i-1}}{w_{i-1}}\right)^{2r-3}$,

(iii) $\mathbf{P}(u \in B_i \wedge v \in B_i \mid u \in W_{i-1} \wedge v \in W_{i-1}) = \frac{(r-1)^2p^2s_{i-1}^2}{w_{i-1}^2}\left(1 - \frac{ps_{i-1}}{w_{i-1}}\right)^{2r-4}$,

(iv) $\mathbf{P}(u \in W_i \wedge v \in B_i \mid u \in W_{i-1} \wedge v \in B_{i-1})$
$= (1-p)\left(1 - \frac{ps_{i-1}}{w_{i-1}}\right)^{r-1}\left(1 - \frac{rpt_{i-1}}{(r-1)b_{i-1}}\right)^{r-2}$,

(v) $\mathbf{P}(u \in B_i \wedge v \in B_i \mid u \in W_{i-1} \wedge v \in B_{i-1})$
$$= \frac{(r-1)p(1-p)s_{i-1}}{w_{i-1}} \left(1 - \frac{ps_{i-1}}{w_{i-1}}\right)^{r-2} \left(1 - \frac{rpt_{i-1}}{(r-1)b_{i-1}}\right)^{r-2},$$

(vi) $\mathbf{P}(u \in B_i \wedge v \in B_i \mid u \in B_{i-1} \wedge v \in B_{i-1}) = (1-p)^2 \left(1 - \frac{rpt_{i-1}}{(r-1)b_{i-1}}\right)^{2r-4}.$

Proof Let $u_1, \ldots, u_{r-1}$ be the neighbours of u other than v and $v_1, \ldots, v_{r-1}$ be the neighbours of v distinct from u. Then

$$\begin{aligned}
&\mathbf{P}(u \in W_i \wedge v \in W_i \mid u \in W_{i-1} \wedge v \in W_{i-1}) \\
&= \mathbf{P}(u_1, \ldots, u_{r-1}, v_1, \ldots, v_{r-1} \notin R_i \mid u \in W_{i-1} \wedge v \in W_{i-1}) \\
&= \prod_{j=1}^{r-1} \mathbf{P}(u_j \notin R_i \mid u \in W_{i-1}) \prod_{j=1}^{r-1} \mathbf{P}(v_j \notin R_i \mid v \in W_{i-1}) \\
&= \left(1 - \frac{ps_{i-1}}{w_{i-1}}\right)^{2r-2}.
\end{aligned}$$

This is based on the remark after Corollary 2.4.1.

A similar strategy leads to the other formulae. ∎

2.5 Differential equations

Using the expressions calculated in the last section, we can now determine recursive formulae for the variables introduced for the analysis of our algorithm.

1. Formula for w_i:
$$\begin{aligned}
w_i &= \mathbf{P}(u \in W_i) = \mathbf{P}(u \in W_i \wedge u \in W_{i-1}) \\
&= \mathbf{P}(u \in W_{i-1})\mathbf{P}(u \in W_i \mid u \in W_{i-1}) = w_{i-1}\left(1 - \frac{ps_{i-1}}{w_{i-1}}\right)^r.
\end{aligned}$$

2. Formula for b_i:
$$\begin{aligned}
b_i &= \mathbf{P}(u \in B_i) = \mathbf{P}(u \in B_i \wedge u \in B_{i-1}) + \mathbf{P}(u \in B_i \wedge u \in W_{i-1}) \\
&= \mathbf{P}(u \in B_{i-1})\mathbf{P}(u \in B_i \mid u \in B_{i-1}) + \mathbf{P}(u \in W_{i-1})\mathbf{P}(u \in B_i \mid u \in W_{i-1}) \\
&= b_{i-1}(1-p)\left(1 - \frac{rpt_{i-1}}{(r-1)b_{i-1}}\right)^{r-1} + rps_{i-1}\left(1 - \frac{ps_{i-1}}{w_{i-1}}\right)^{r-1}.
\end{aligned}$$

3. Formula for q_i:
$$\begin{aligned}
q_i &= \mathbf{P}(u \in W_i \wedge v \in W_i) \\
&= \mathbf{P}(u \in W_{i-1} \wedge v \in W_{i-1})\mathbf{P}(u \in W_i \wedge v \in W_i \mid u \in W_{i-1} \wedge v \in W_{i-1}) \\
&= q_{i-1}\left(1 - \frac{ps_{i-1}}{w_{i-1}}\right)^{2r-2}.
\end{aligned}$$

4. Formula for s_i:

$$\begin{aligned}
s_i &= \mathbf{P}(u \in B_i \wedge v \in W_i) \\
&= \mathbf{P}(u \in B_{i-1} \wedge v \in W_{i-1})\mathbf{P}(u \in B_i \wedge v \in W_i \mid u \in B_{i-1} \wedge v \in W_{i-1}) \\
&+ \mathbf{P}(u \in W_{i-1} \wedge v \in W_{i-1})\mathbf{P}(u \in B_i \wedge v \in W_i \mid u \in W_{i-1} \wedge v \in W_{i-1}) \\
&= s_{i-1}(1-p)\left(1 - \frac{ps_{i-1}}{w_{i-1}}\right)^{r-1}\left(1 - \frac{rpt_{i-1}}{(r-1)b_{i-1}}\right)^{r-2} \\
&+ \frac{(r-1)pq_{i-1}s_{i-1}}{w_{i-1}}\left(1 - \frac{ps_{i-1}}{w_{i-1}}\right)^{2r-3}.
\end{aligned}$$

5. Formula for t_i:

$$\begin{aligned}
t_i &= \mathbf{P}(u \in B_i \wedge v \in B_i) \\
&= \mathbf{P}(u \in B_{i-1} \wedge v \in B_{i-1})\mathbf{P}(u \in B_i \wedge v \in B_i \mid u \in B_{i-1} \wedge v \in B_{i-1}) \\
&+ \mathbf{P}(u \in B_{i-1} \wedge v \in W_{i-1})\mathbf{P}(u \in B_i \wedge v \in B_i \mid u \in B_{i-1} \wedge v \in W_{i-1}) \\
&+ \mathbf{P}(v \in W_{i-1} \wedge u \in B_{i-1})\mathbf{P}(u \in B_i \wedge v \in B_i \mid u \in W_{i-1} \wedge v \in B_{i-1}) \\
&+ \mathbf{P}(v \in W_{i-1} \wedge u \in W_{i-1})\mathbf{P}(u \in B_i \wedge v \in B_i \mid u \in W_{i-1} \wedge v \in W_{i-1}) \\
&= t_{i-1}(1-p)^2\left(1 - \frac{rpt_{i-1}}{(r-1)b_{i-1}}\right)^{2r-4} \\
&+ 2s_{i-1}(1-p)\left(1 - \frac{rpt_{i-1}}{(r-1)b_{i-1}}\right)^{r-2}\frac{(r-1)ps_{i-1}}{w_{i-1}}\left(1 - \frac{ps_{i-1}}{w_{i-1}}\right)^{r-2} \\
&+ q_{i-1}\frac{(r-1)^2p^2s_{i-1}^2}{w_{i-1}^2}\left(1 - \frac{ps_{i-1}}{w_{i-1}}\right)^{2r-4}.
\end{aligned}$$

We need to evaluate w_0, b_0, q_0, s_0 and t_0 to have the necessary set of initial conditions for solving the system of recurrence equations found above. It is easy to see that $w_0 = \mathbf{P}(u \in W_0) = (1-p_0)^{r+1}$ and $b_0 = rp_0(1-p_0)^r$, since for the former neither u nor its neighbours can have relevant label 0, and for the latter u cannot have relevant label 0, but exactly one of its neighbours must have it.

Now,

$$q_0 = \mathbf{P}(u \in W_0 \wedge v \in W_0) = (1-p_0)^{2r},$$

since the event $v \in W_0 \wedge u \in W_0$ is equivalent to neither u, v nor any of their other neighbours being chosen in the first phase of the algorithm (and each vertex is chosen independently with probability p_0).

The equation for s_0 is given by

$$s_0 = \mathbf{P}(u \in B_0 \wedge u \in W_0) = (r-1)p_0(1-p_0)^{2r-1}$$

because $v \in B_0 \wedge u \in W_0$ occurs when u, v are not chosen, no neighbours of u are chosen and precisely one neighbour of v is chosen.

Finally, the equation for t_0 is

$$t_0 = \mathbf{P}(u \in B_0 \wedge v \in B_0) = (r-1)^2 p_0^2 (1-p_0)^{2r-2}$$

with similar justification.

The recurrence equation for w_i obtained at the beginning of this section can be seen as

$$w_i = w_{i-1} - prs_{i-1} + O(p^2).$$

For p small, the term $O(p^2)$ should only have a minor influence. Similarly, each of the other equations of the system of recurrence equations can be rewritten as a main term added to a term of the order of p^2. By ignoring the latter, we obtain the following auxiliary system of recurrence equations:

$$\begin{aligned}
w'_i &= w'_{i-1} - prs'_{i-1} \\
b'_i &= b'_{i-1} + p\left(-b'_{i-1} - rt'_{i-1} + rs'_{i-1}\right) \\
q'_i &= q'_{i-1} - p\frac{(2r-2)q'_{i-1}s'_{i-1}}{w'_{i-1}} \\
s'_i &= s'_{i-1} + p\left(-s'_{i-1} + \frac{(r-1)q'_{i-1}s'_{i-1}}{w'_{i-1}}\right. \\
&\qquad \left. - \frac{(r-1)\ {s'_{i-1}}^2}{w'_{i-1}} - \frac{r(r-2)s'_{i-1}t'_{i-1}}{(r-1)b'_{i-1}}\right) \\
t'_i &= t'_{i-1} + p\left(-2t'_{i-1} + \frac{2(r-1)\ {s'_{i-1}}^2}{w'_{i-1}} - \frac{2r(r-2)\ {t'_{i-1}}^2}{(r-1)b'_{i-1}}\right) \\
w'_0 &= (1-p_0)^{r+1},\ b'_0 = rp_0(1-p_0)^r,\ q'_0 = (1-p_0)^{2r}, \\
s'_0 &= (r-1)p_0(1-p_0)^{2r-1},\ t'_0 = (r-1)^2 p_0^2 (1-p_0)^{2r-2}
\end{aligned} \tag{2.5.1}$$

Note that the auxiliary system of recurrence equations (2.5.1) can be converted into a system of differential equations by means of first order approximations. Setting $p = \epsilon$ in the recurrence equation for w'_i obtained above implies

$$w'_i - w'_{i-1} = -\epsilon r s'_{i-1}.$$

This suggests that, for ϵ small, the solutions to this recurrence equation are approximated by the functions $\hat{w}, \hat{s}$ satisfying the differential equation

$$\frac{d\hat{w}}{dx} = -r\hat{s}.$$

Using the same idea with the other recurrence formulae in (2.5.1), the following system of differential equations arises. This system will be referred to as *the system of differential*

equations associated with (r, p_0).

$$\begin{aligned}
&\frac{d\hat{w}}{dx} = -r\hat{s}\\
&\frac{d\hat{b}}{dx} = -\hat{b} - r\hat{t} + r\hat{s}\\
&\frac{d\hat{q}}{dx} = -\frac{(2r-2)\hat{q}\hat{s}}{\hat{w}}\\
&\frac{d\hat{s}}{dx} = -\hat{s} + \frac{(r-1)\hat{q}\hat{s}}{\hat{w}} - \frac{(r-1)\hat{s}^2}{\hat{w}} - \frac{r(r-2)\hat{s}\hat{t}}{(r-1)\hat{b}}\\
&\frac{d\hat{t}}{dx} = -2\hat{t} + \frac{2(r-1)\hat{s}^2}{\hat{w}} - \frac{2r(r-2)\hat{t}^2}{(r-1)\hat{b}}\\
&\hat{w}(0) = (1-p_0)^{r+1},\ \hat{b}(0) = rp_0(1-p_0)^r,\ \hat{q}(0) = (1-p_0)^{2r},\\
&\hat{s}(0) = (r-1)p_0(1-p_0)^{2r-1},\ \hat{t}(0) = (r-1)^2p_0^2(1-p_0)^{2r-2}.
\end{aligned} \tag{2.5.2}$$

Given $p_0 \in (0,1)$, $T > 0$ and $\gamma > 0$, where $\gamma < \min\{w_0, b_0, q_0, s_0, t_0\}$, this system of differential equations has a solution in the domain $\Omega(\gamma, T) = \{(x, \hat{w}, \hat{b}, \hat{q}, \hat{s}, \hat{t}) \in (-\gamma, T) \times (\gamma, 1)^5\}$ which may be uniquely extended arbitrarily close to the boundary of the domain, by the following standard result in the theory of first order differential equations (see Hurewicz [33], Chapter 2, Theorem 11). The proof is omitted.

Lemma 2.5.1 *If a set of functions* $f_i : \mathbb{R}^{s+1} \to \mathbb{R}$ *is Lipschitz in a bounded region* Ω *and the point* $(x_0, y_1', \ldots, y_s')$ *lies in* Ω*, then the solution of*

$$\begin{aligned}
&\frac{dz_i}{dx} = f_i(x, z_1, \ldots, z_s),\ i = 1, \ldots, s,\\
&z_i(x_0) = y_i',\ i = 1, \ldots, s,
\end{aligned}$$

may be uniquely extended arbitrarily close to the boundary of Ω.

As expected, there is a connection between the original system of recurrence equations and the system of differential equations (2.5.2). This connection is summarised in the lemma below and follows from the solutions to the original system being well-approximated by the solutions of the modified system (2.5.1), as well as from the relation between the solutions of (2.5.1) and of (2.5.2) given by Euler's method. The proof is routine, so is omitted. However, the main ideas involved in the proof of this result can be found in the proof of Theorem 4.2.2 in Chapter 4.

Lemma 2.5.2 *Let* $r \geq 3$ *be an integer and* $p_0 \in (0,1)$*. Let* $k_0 > 0$ *such that the system of differential equations (2.5.2) with the initial conditions defined by* p_0 *has positive solutions in* Ω *defined at* $x = k_0$*. Then, given* $\xi > 0$*,*

(i) there exists $\epsilon_0 > 0$ *satisfying the following property. If* $0 < \epsilon \leq \epsilon_0$ *and the system of recurrence equations (2.5.1) is solved with* $p = \epsilon$*, then* $|w_i - \hat{w}(\epsilon i)| < \xi$*,* $|b_i - \hat{b}(\epsilon i)| < \xi$*,* $|q_i - \hat{q}(\epsilon i)| < \xi$*,* $|s_i - \hat{s}(\epsilon i)| < \xi$ *and* $|t_i - \hat{t}(\epsilon i)| < \xi$*, for* $i = 0, 1, \ldots, \lceil k_0/\epsilon \rceil$.

(ii) there exists $\epsilon_1 > 0$ *such that for* $0 \le \epsilon \le \epsilon_1$,

$$\left| \int_0^{k_0} \hat{b}(x)\, dx - \sum_{i=0}^{\lceil k_0/\epsilon \rceil - 1} \epsilon b_i \right| < \xi, \text{ for every } 0 < \epsilon \le \epsilon_1.$$

Using this lemma, we can now determine additional properties of the solutions to (2.5.2).

Lemma 2.5.3 *Given* $p_0 \in (0,1)$, *the system of differential equations (2.5.2) has unique solutions* $\hat{w}(x)$, $\hat{b}(x)$, $\hat{q}(x)$, $\hat{s}(x)$ *and* $\hat{t}(x)$ *defined over the entire nonnegative real line satisfying the following properties:*

(i) $\hat{w}(x)$, $\hat{b}(x)$, $\hat{q}(x)$, $\hat{s}(x)$ *and* $\hat{t}(x)$ *are positive,*

(ii) $\int_0^{\infty} \hat{b}(x)\, dx$ *converges.*

Proof As mentioned before, a standard result in the theory of first order differential equations ensures that, for $p_0 \in (0,1)$, $T > 0$ and $\gamma < \min\{w_0, b_0, q_0, s_0, t_0\}$, the system of differential equations has a solution in the domain $\Omega(\gamma, T) = \{(x, \hat{w}, \hat{b}, \hat{q}, \hat{s}, \hat{t}) \in (-\gamma, T) \times (\gamma, 1)^2 \times (\gamma, 1)^3\}$ which may be uniquely extended arbitrarily close to the boundary of the domain.

Given $p_0 \in (0,1)$ and $T > 0$, we show that there exists $\gamma = \gamma(T) > 0$ such that this system of differential equations in the domain $\Omega(\gamma, T)$ has a unique solution defined for x arbitrarily close to $x = T$. This implies that the solutions are defined over the nonnegative real line.

Suppose on the contrary that, for some $T > 0$, no $\gamma(T)$ with the above property exists. Let x_0 denote the infimum of such T. Let (x', w', b', q', s', t') be any point in the interior of a region $\Omega(\gamma_0, x_0)$ such that a solution to the system of differential equations exists for $0 \le x \le x'$ and $\hat{w}(x') = w'$, $\hat{b}(x') = b'$, $\hat{q}(x') = q'$, $\hat{s}(x') = s'$ and $\hat{t}(x') = t'$, where $\gamma_0 > 0$. By Lemma 2.5.2, given $\xi > 0$, there exists ϵ_0 such that for $0 < \epsilon \le \epsilon_0$ and $0 \le i \le \lceil x'/\epsilon \rceil$,

$$|w_i - \hat{w}(\epsilon i)| < \xi, |b_i - \hat{b}(\epsilon i)| < \xi, |q_i - \hat{q}(\epsilon i)| < \xi, |s_i - \hat{s}(\epsilon i)| < \xi, |t_i - \hat{t}(\epsilon i)| < \xi.$$

Since the quantities w_i, b_i, q_i, s_i and t_i represent probabilities of specific events after i steps of a randomised algorithm, we conclude that

$$q_i + s_i \le w_i, \; s_i + t_i \le b_i.$$

Using this and the fact that $w', b', q', s', t' > \gamma_0$ (which is independent of ξ), we have

$$\max\{\hat{q}(x'), \hat{s}(x')\} < \hat{w}(x'), \quad \max\{\hat{s}(x'), \hat{t}(x')\} < \hat{b}(x'). \tag{2.5.3}$$

Let m_0 be a positive integer such that

$$\frac{1}{m_0} < \min\{w_0, b_0, q_0, s_0, t_0\}.$$

The definition of x_0 implies that one of the functions $\hat{w}, \hat{b}, \hat{q}, \hat{s}, \hat{t}$ must get arbitrarily close to 0 in the neighbourhood of a point x_0, $0 < x_0 < T$. By (2.5.3), it must be one of $\hat{q}$, $\hat{s}$ or $\hat{t}$. (Note that this argument also applies in the case that $x = 0$.)

Suppose this is the case for $\hat{q}$. Let x' be such that the system of differential equations have a positive solution in $[0, x']$. Recall that

$$\frac{d\hat{q}}{dx} = -\frac{(2r-2)\hat{q}\hat{s}}{\hat{w}} = \hat{q}\left(\frac{-(2r-2)\hat{s}}{\hat{w}}\right),$$

and, by equation (2.5.3),

$$\frac{-(2r-2)\hat{s}(x)}{\hat{w}(x)} \geq -(2r-2)$$

for $0 \leq x \leq x'$. Now, if f is the solution for

$$\frac{df}{dx} = -(2r-2)f, \ f(0) = \hat{q}(0),$$

we must have $\hat{q}(x) \geq f(x)$ for every x in the interval $[0, x']$. However, $f(x) = f(0)e^{-(2r-2)x}$ is a strictly positive function in this interval bounded below by the constant $f(0)e^{-(2r-2)x'}$. So $\hat{q}(x)$ cannot approach 0 at x'. Similar arguments yield contradictions for the cases when $\hat{s}(x)$ or $\hat{t}(x)$ approach 0 in the neighbourhood of the point x_0, since

$$\begin{aligned}\frac{d\hat{s}}{dx} &= \left(-1 + \frac{(r-1)\hat{q}}{\hat{w}} - \frac{(r-1)\hat{s}}{\hat{w}} - \frac{r(r-2)\hat{t}}{(r-1)\hat{b}}\right)\hat{s} \\ &\geq \left(-1 - (r-1) - \frac{r(r-2)}{r-1}\right)\hat{s},\end{aligned}$$

and

$$\begin{aligned}\frac{d\hat{t}}{dx} &= -2\hat{t} + \frac{2(r-1)\hat{s}^2}{\hat{w}} - \frac{2r(r-2)\hat{t}^2}{(r-1)\hat{b}} \\ &\geq \left(-2 - \frac{2r(r-2)}{r-1}\right)\hat{t}.\end{aligned}$$

Thus, the solutions to the system of differential equations are indeed defined over the entire nonnegative real line. Furthermore, the previous argument ensures that they are positive, concluding the proof of part (i).

For part (ii), note that the differential equations for $\hat{w}$ and $\hat{b}$ in (2.5.2) imply

$$\frac{d(\hat{w} + \hat{b})}{dx} = -\hat{b} - r\hat{t},$$

so

$$\hat{b}(x) = -\frac{d(\hat{w} + \hat{b})}{dx}(x) - r\hat{t}(x) \leq -\frac{d(\hat{w} + \hat{b})}{dx}(x), \ \forall x.$$

As a consequence, for every $T > 0$,

$$\int_0^T \hat{b}(x)\ dx \leq \int_0^T -\frac{d(\hat{w} + \hat{b})}{dx}(x)\ dx = \hat{w}(0) + \hat{b}(0) - \hat{w}(T) - \hat{b}(T) \leq \hat{w}(0) + \hat{b}(0).$$

This proves part (ii). ∎

2.6 Proof of Theorem 2.1.1

We wish to obtain a lower bound on the cardinality of a largest vertex subset that induces a forest in an r-regular graph G not containing short cycles. Recall our definition of $\tau(G)$, given by

$$\tau(G) = \max \{|V(F)| : F \text{ is an induced forest in } G\}.$$

Let G be an r-regular graph on n vertices with girth g and consider the set P of purple vertices at the end of step 2 when Algorithm 2.2.1 is applied to G with $N < g/2 - 2$. It is clear that the induced graph $G[P]$ contains a cycle only if some vertex v with at least two purple neighbours has been added to P. By the description of our algorithm, this cannot happen unless v was selected in the same step as one of its neighbours. It follows that, if $\bar{P}$ is the set obtained from P by deleting any pairs of adjacent vertices added to P in the same step, the induced subgraph $G[\bar{P}]$ is acyclic.

Now, given a vertex in R_i, the probability that none of its neighbours is also selected is at least $(1-p)^r$, since a vertex has at most r neighbours that could be added to R_i. Therefore, the expected number of vertices added to P at time i that are not removed is at least $p(1-p)^r b_{i-1} n$ and

$$\mathbf{E}|\bar{P}| \geq p_0(1-p_0)^r n + \sum_{i=1}^{N} p(1-p)^r b_{i-1} n. \tag{2.6.1}$$

Part of the set W of white vertices produced at the end of the algorithm will also be added to the forest. By definition, these vertices have no purple neighbours, so that no cycle containing purple vertices is created by adding white vertices to $\bar{P}$. Thus $G[\bar{P} \cup \bar{W}]$ is still acyclic, where $\bar{W}$ denotes the set of vertices in acyclic components of $G[W]$.

Now, since G has girth g, no cycles appear if we add white vertices lying in components of $G[W]$ for which every pair of vertices are at distance at most $g-1$ from each other. Therefore, a lower bound on the size of $\bar{W}$ can be obtained by estimating the number of vertices in small components of $G[W]$. This will be done through a branching process argument.

To define the branching process, start with a white vertex v_0 and set the random variable $Y_0 = \{v_0\}$. In general, Y_i denotes the set of white vertices already exposed, but whose neighbours have not been considered yet. Define $U_0 = V(G) - \{v_0\}$ and let U_i be the random variable accounting for the set of vertices which have not been exposed by the branching process up to step i. After step i, either $|Y_i| = 0$, in which case the process has died out, or $|Y_i| > 0$, in which case we choose a white vertex v_i in Y_i, expose its white neighbours $N_W(v_i) \subseteq U_i$ and define $Y_{i+1} = Y_i \cup N_W(v_i) - \{v_i\}$, $U_{i+1} = U_i \setminus N(v_i)$. We are interested in estimating the probability that $|Y_{g/2-1}| > 0$, i.e., that the branching process has not died out after $g/2 - 1$ steps.

Proposition 2.6.1 *Let $\delta > 0$, fix an integer $r \geq 3$ and suppose the existence of $p_0 > 0$ such that the solutions to the system of differential equations associated with (r, p_0) satisfy*

$$\lim_{x\to\infty} \frac{(r-1)\hat{q}(x)}{\hat{w}(x)} < 1.$$

Then there exist $g > 0$, $0 < N < g/2 - 1$ and $0 < p < 1$ such that, if Algorithm 2.2.1 is applied to an r-regular graph G with girth at least g for N steps with probabilities (p_0, p), then

$$\mathbf{P}(|Y_{g/2-1}| > 0) < \delta.$$

Proof Let Z_i denote the random variable counting the number of neighbours of v_i in U_i. Note that Z_0 has binomial distribution $\mathbf{Bin}(r, q_N/w_N)$, since Corollary 2.3.3 and Lemma 2.3.5 ensure that, conditional upon v_0 being white, the events associated with each of its neighbours being white are mutually independent and have probability q_N/w_N. Furthermore, Z_i has distribution $\mathbf{Bin}(r-1, q_N/w_N)$ for every $i \geq 1$, since the condition $0 < i < g-1$ implies $|N(v_i) \cap U_i| = r - 1$, and Corollary 2.3.3 and Lemma 2.3.5 are applicable in the same way.

Let $k_0 > 0$ such that the solution to the system of differential equations (2.5.2) satisfies

$$\frac{(r-1)\hat{q}(k_0)}{\hat{w}(k_0)} < 1.$$

Let $\xi > 0$ be such that

$$\frac{(r-1)(\hat{q}(k_0)+\xi)}{\hat{w}(k_0)-\xi} < 1.$$

Fix ϵ_0 as in Lemma 2.5.2, part (i), and let $\epsilon < \epsilon_0$ such that $N = k_0/\epsilon$ is an integer. Now, apply Algorithm 2.2.1 for N steps with the given p_0 and $p = \epsilon$, for all $i \geq 1$, to a graph G with girth $g \geq 2N + 3$. Then

$$\frac{(r-1)q_N}{w_N} \leq \frac{(r-1)(\hat{q}(k_0)+\xi)}{\hat{w}(k_0)-\xi} < 1.$$

So, we have $(r-1)q_N/w_N < 1$, and a branching process argument as in [4] shows that, by choosing g sufficiently large,

$$\mathbf{P}(|Y_{g/2-1}| > 0) < \delta,$$

as required. ∎

By the above proposition, given $\delta > 0$ and $p_0 > 0$ such that the solutions to the system of differential equations associated with (r, p_0) satisfy

$$\lim_{x\to\infty} \frac{(r-1)\hat{q}(x)}{\hat{w}(x)} < 1,$$

we may fix g, N and p so as to have the property $\mathbf{P}(Y_{g/2-1} > 0) < \delta$, i.e., $\mathbf{P}(Y_{g/2-1} = 0) \geq 1 - \delta$. It follows that for such g the expected number of white vertices in acyclic components of $G[W]$ is bounded below by

$$(1-\delta)w_N n. \tag{2.6.2}$$

We are now ready to prove Theorem 2.1.1.

Proof of Theorem 2.1.1 Fix $r \in \mathbb{N}$ and $\delta > 0$. We show that, given $p_0 \in (0,1)$, the inequality $\tau(G) \geq (\xi(p_0) - \delta)\, n$ holds, where

$$\xi(p_0, r) = \begin{cases} p_0(1-p_0)^r + \displaystyle\int_0^\infty \hat{b}(x)\ dx + \lim_{x\to\infty} \hat{w}(x), & \text{if } \lim_{x\to\infty} \dfrac{(r-1)\hat{q}(x)}{\hat{w}(x)} < 1 \\ p_0(1-p_0)^r + \displaystyle\int_0^\infty \hat{b}(x)\ dx, & \text{otherwise.} \end{cases} \tag{2.6.3}$$

Here, $\hat{w}$, $\hat{b}$ and $\hat{q}$ are solutions to the system of differential equations associated with (r, p_0). By Lemma 2.5.3, this system has positive solutions $\hat{w}, \hat{b}, \hat{q}, \hat{s}, \hat{t}$ defined over the nonnegative real line such that $\displaystyle\int_0^\infty \hat{b}(x)\ dx$ converges.

Let $k_0 > 0$ be such that, for every $k > k_0$,

$$\left| \int_0^\infty \hat{b}(x)\ dx - \int_0^k \hat{b}(x)\ dx \right| < \frac{\delta}{6}. \tag{2.6.4}$$

Using Lemma 2.5.2, fix $\epsilon_0 > 0$ such that

$$|w_i - \hat{w}(\epsilon i)| < \frac{\delta}{6},\ i = 0, 1, \ldots, \left\lceil \frac{k_0}{\epsilon} \right\rceil,$$

and fix $\epsilon_1 > 0$ satisfying

$$\left| \int_0^{k_0} \hat{b}(x)\ dx - \sum_{i=0}^{\left\lceil \frac{k_0}{\epsilon} \right\rceil - 1} \epsilon b_i \right| < \frac{\delta}{6},\ \text{for every } 0 < \epsilon \leq \epsilon_1.$$

Let $\epsilon = \min\{\epsilon_0, \epsilon_1, 1 - (1-\delta/6)^{1/r}\}$ and $N = \lceil k_0/\epsilon \rceil$. Fix $g = 2N + 3$. Then, given an r-regular graph G with girth larger than or equal to g, we apply Algorithm 2.2.1 for N steps with probabilities $(p_0, p = \epsilon)$. The first moment principle leads to a lower bound for $\tau(G)$. As a matter of fact, our lower bound (2.6.1) on the cardinality of $\bar{P}$ implies

$$\begin{aligned} \mathbf{E}|\bar{P}| &\geq np_0(1-p_0)^r + n(1-\epsilon)^r \left(\sum_{i=1}^{N} \epsilon b_{i-1} \right) \\ &\geq np_0(1-p_0)^r + n(1-\epsilon)^r \left(\int_0^{k_0} \hat{b}(x)\ dx - \frac{\delta}{6} \right) \\ &\geq np_0(1-p_0)^r + n\left(1 - \frac{\delta}{6}\right) \left(\int_0^\infty \hat{b}(x)\ dx - \frac{2\delta}{6} \right) \\ &\geq n\left(p_0(1-p_0)^r + \int_0^\infty \hat{b}(x)\ dx \right) - \frac{\delta\ n}{2}, \end{aligned} \tag{2.6.5}$$

If, in addition, the solutions to the system of differential equations associated with (r, p_0) satisfy

$$\lim_{x\to\infty} \frac{(r-1)\hat{q}(x)}{\hat{w}(x)} < 1,$$

Proposition 2.6.1 establishes a lower bound (2.6.2) on the cardinality of the set $\bar{W}$ of white vertices that can be added to the forest. Clearly, k_0 in (2.6.4) may be chosen so that, for every $k > k_0$, we also have

$$\left| \lim_{x->\infty} \hat{w}(x) - \hat{w}(k) \right| < \frac{\delta}{6}$$

and

$$\frac{(r-1)\hat{q}(k)}{\hat{w}(k)} < 1.$$

The girth g can also be taken larger, if necessary, to ensure that the size of $\bar{W}$ is bounded below by $(1-\delta/6)w_N n$.

Thus,

$$\begin{aligned} \mathbf{E}|\bar{W}| &\geq n\left(1-\frac{\delta}{6}\right) w_N \geq n\left(1-\frac{\delta}{6}\right)^2 \hat{w}(\epsilon N) \\ &\geq n\left(1-\frac{\delta}{6}\right)^3 \lim_{x\to\infty} \hat{w}(x) \geq n \lim_{x\to\infty} \hat{w}(x) - \frac{\delta\, n}{2}, \end{aligned} \tag{2.6.6}$$

Now, given that $\tau(G) \geq \mathbf{E}|\bar{P} \cup \bar{W}|$ and using equations (2.6.5) and (2.6.6), we conclude that $\tau(G) \geq (\xi(p_0, r) - \delta)\, n$. In particular, if

$$\xi(r) = \sup_{p_0 \in (0,1]} \xi(p_0, r),$$

we have

$$\tau(G) \geq (\xi(r) - \delta)\, n,$$

as claimed. Numerical approximations of these quantities lead us to the bounds in Table 2.1.1. We note that, for every value of r tested, we were able to choose a constant p_0 such that the numerical solutions to the system of differential equations associated with (r, p_0) satisfy $\lim_{x\to\infty} \frac{(r-1)\hat{q}(x)}{\hat{w}(x)} < 1$. ∎

Chapter 3

Locally greedy algorithms

In this chapter, we introduce a class of probabilistic algorithms which can be applied to regular graphs. This class of algorithms is a generalisation of the algorithm for induced forests introduced in the previous chapter, which, in Chapters 5 and 6, will be used to derive new lower bounds on the independence number and on the number of vertices in a maximum induced forest in a regular graph with girth sufficiently large.

3.1 The class of algorithms

The class of locally greedy algorithms is given as follows. The parameters consist of positive integers r, d and ℓ, with $\ell \leq r$, and of an assignment of probabilities to the random choices made by a particular instance of the algorithm.

Algorithm 3.1.1 ***(Locally greedy algorithm)***

Input: An r-regular graph G, a positive integer N, an initial probability p_0 and vectors of probabilities $\mathbf{p}_i = (p_{i,j,k} : j, k \geq 0,\ j + k \leq r,\ j < l)$, $i = 1, \ldots, N$.

1. *Start with all the vertices of the graph coloured white. In the first step, colour each vertex purple with probability p_0, at random, independently of all others. Vertices are coloured purple if they are at distance at most $d - 1$ of one of the chosen vertices. Non-purple vertices are coloured yellow if they have at least ℓ purple neighbours.*

2. *At each step i, a set of white vertices is chosen, where a white vertex with j purple neighbours and k yellow neighbours is chosen randomly, independently of all others, with probability $p_{i,j,k}$. The chosen vertices and the vertices at distance at most $d - 1$ of them become purple. Non-purple vertices are coloured yellow if, at the end of this step, they have at least ℓ purple neighbours. Repeat this iteratively for N steps.*

3. *Create a set $\bar{P}$ by deleting, from the set of purple vertices, any pair of vertices within distance $2d-1$ that were chosen in the same step and all vertices at distance at most $d-1$ of one of these two vertices that turned purple because of this.*

Output: The set $\bar{P}$.

This algorithm is used to generate a set of vertices satisfying some graph property of interest. This set of vertices is often contained in the set of vertices that are chosen by the algorithm, as will be the case in the applications we shall present, but there are examples in which this may not be the case. Also, for the graph properties studied in this thesis, this algorithm will be applied with parameter $d = 1$, so the sets of chosen vertices and of purple vertices are equal. However, we shall consider the more general case for which d is a fixed positive integer, since this may lead to useful generalisations of our results without creating serious complications. Note that, if a vertex is purple at a given step of the algorithm, it cannot be chosen in the later steps. The same is true for yellow vertices, so the white vertices are the vertices that could still be chosen in the future.

The next proposition describes the output of the algorithm with two specific assignments of parameters.

Proposition 3.1.2 *The output produced by an instance of the locally greedy algorithm with parameters $d = \ell = 1$ is an independent set in the input graph. If the parameters are $\ell = 1$ and $d = 2$, the set obtained by the locally greedy algorithm induces a forest in the input graph.*

Proof In the case $\ell = d = 1$, we show that the output $\bar{P}$ produced by the algorithm is an independent set. Indeed, any two adjacent vertices in the set of purple vertices must have been chosen in the same step, since, once a white vertex is chosen, all its white neighbours that were not chosen in this step become yellow, which prevents them from being chosen in the future. But adjacent vertices chosen in the same step are removed when $\bar{P}$ is created at the third part of the algorithm, so $\bar{P}$ is independent as claimed.

Now, let $d = 1$ and $\ell = 2$ and suppose that, at the end of part 2 of the algorithm, a set of purple vertices induces a cycle C in the input graph G. Let v be a vertex in this cycle that became purple at the latest step. If the two vertices adjacent to v in this cycle had been chosen in earlier steps, v would have had two purple neighbours before being chosen, which would have forced it to turn yellow by the definition of the algorithm, a contradiction. This implies that vertex v is not in the set $\bar{P}$ produced by the algorithm, and, as a consequence, $\bar{P}$ induces an acyclic subgraph of G. ∎

Special cases of this class of algorithms have been used before. As a matter of fact, if $r \geq 3$ and $d = \ell = 1$, and if the probabilities are $p_0 = 0$ and $p_{i,j,k} = p_i$, for some fixed constants p_i, for every j and k, this algorithm coincides with the algorithm to generate independent sets given in [39].

In the case when $d = 1$, $\ell = 2$ and $p_{i,j,k} = \delta_{j,1}p$, where p is a constant, for every j and k, this algorithm is just the simple algorithm for induced forests introduced in Chapter 2. (Here $\delta_{m,n}$ denotes the Kronecker delta function.) In that algorithm, however, the white vertices with a purple neighbour were treated separately and called *blue* vertices.

Let $G = (V, E)$ be an r-regular graph with girth at least g. Consider a set $\mathcal{P}$ of probabilities $p_{i,j,k}$, $i \geq 0$, $(j, k) \in \mathcal{I}$, where

$$\mathcal{I} = \{(j, k) \in \mathbb{Z}^2 \,:\, j, k \geq 0,\ j < \ell,\ j + k \leq r\}. \tag{3.1.1}$$

As with the simple algorithm in Chapter 2, we shall analyse the performance of an algorithm in the class of locally greedy algorithms by looking at an equivalent model that makes the random choices uniform over the whole set of vertices. However, since the probability of a vertex v being chosen at a given step i may depend on the number of purple and yellow neighbours of v after step $i-1$ in the algorithm, we now associate several sequences of labels with each vertex, one for each type of random choice. Indeed, with each vertex $v \in V$, and each pair $(j, k) \in \mathcal{I}$, we associate random variables $S_{j,k}(v) \subseteq \{0, 1, 2, \ldots\}$, where label i is placed in $S_{j,k}(v)$ with probability $p_{i,j,k}$, and all these choices are made independently of each other and independently of the other sets $S_{j',k'}(v')$. The sequence $[(S_{j,k}(v))_{(j,k)\in\mathcal{I}} : v \in V]$ will be used to denote the random variables of sets of labels associated with each vertex in the graph, whereas the sequence $\mathcal{S} = [(S^v_{j,k})_{(j,k)\in\mathcal{I}} : v \in V]$ will denote an assignment of sets to these random variables. These assignments of sets to the vertices will also be called *labellings* of the graph.

As in Chapter 2, given a graph G, parameters d and ℓ, and a labelling $\mathcal{S}$ of the graph, we shall inductively define the concepts of *relevant label of a vertex* and of *colouring of G at time i*. As before, a vertex has relevant label i if and only if it is chosen by the algorithm at the i-th step when the random choices made by the algorithm are determined by the labelling $\mathcal{S}$. Moreover, the colouring of the graph at time i simply corresponds to the colours assigned to each vertex in the graph in an application of the algorithm for which the set of vertices chosen at time i are precisely the sets of vertices with relevant label i.

For a vertex $u \in V$, we let $N(u) = \{v \in V \,:\, uv \in E\}$ denote the set of vertices adjacent to u and we let $N_s(u) = \{v \in V : d(u, v) \leq s\}$, so that $N(u)$ denotes the set of vertices in the graph at distance at most s from u.

Definition 3.1.3 *A nonnegative integer i is the relevant label of a vertex u if $i = 0$ and $0 \in S_{0,0}(u)$, or if $i > 0$, $u \in W_{i-1}$, $|P_{i-1} \cap N(u)| = j$, $|Y_{i-1} \cap N(u)| = k$ and $i \in S_{j,k}(u)$, for some pair $(j, k) \in \mathcal{I}$, where the definition of the sets W_{i-1}, P_{i-1} and Y_{i-1} follows. The colouring of $G = (V, E)$ at time i is the partition of V into three colour classes P_i, W_i and Y_i, with P_i being the set of all vertices at distance at most $d - 1$ of a vertex with relevant label smaller than or equal to i. Of the vertices not in P_i, W_i is the set of vertices adjacent*

to fewer than ℓ elements of P_i and Y_i is the set of vertices adjacent to at least ℓ elements of P_i. If the relevant label of u is undefined for any $i \geq 0$, we define it at this point to be ∞.

The set of vertices with relevant label i is denoted by R_i. So, $P_i = \{u \in V : \bigcup_{t=0}^{i} R_t \cap N_{d-1}(u) \neq \emptyset\}$, $W_i = \{u \in V : \bigcup_{t=0}^{i} R_t \cap N_{d-1}(u) = \emptyset \wedge |P_i \cap N(u)| < \ell\}$ and $Y_i = \{u \in V : \bigcup_{t=0}^{i} R_t \cap N_{d-1}(u) = \emptyset \wedge |P_i \cap N(u)| \geq \ell\}$. The elements of each of these sets will be called the *purple*, the *yellow* and the *white* vertices at time i, respectively. The relevant label of v with respect to the labelling $\mathcal{S}$ is denoted by $L_{\mathcal{S}}(v)$.

It is clear from this definition that the colouring of G at time i is fully determined by the sequence $\mathcal{S} = [(S^v_{j,k} \cap \{0, \ldots, i\})_{(j,k)\in\mathcal{I}} : v \in V]$. Moreover, this colouring coincides with the colouring of the graph induced by Algorithm 3.1.1 in a natural way, which is formalised in the next proposition. Note the similarity with Lemma 2.2.4.

Proposition 3.1.4 *Let r, d, ℓ and N be positive integers with $r \geq 3$ and $\ell \leq r$, and fix an initial probability p_0 and vectors of probabilities $\mathbf{p}_i = (p_{i,j,k} : (j,k) \in \mathcal{I})$, $i = 1, \ldots, N$. Let $G = (V, E)$ be a graph, and consider a subgraph H of G and a colouring c of H with colours purple, yellow and white. Then the following events have the same probability:*

(i) the colouring of G at time i induced by the sequence $\mathcal{S} = [(S_{j,k}(v))_{(j,k)\in\mathcal{I}} : v \in V(G)]$ restricted to H is equal to c, where $\mathcal{S}$ is obtained by adding each nonnegative integer i to $S_{j,k}(v)$ independently with probability p_0, if $(j,k) = (0,0)$ and $i = 0$, or $p_{i,j,k}$, if $i \geq 1$, for all $v \in V$.

(ii) Algorithm 3.1.1 applied to G obtains c as the colouring of H after step i.

Proof We may modify part 2 of Algorithm 3.1.1 as follows. At each step i, every vertex $v \in V(G)$ is chosen with respect to each pair $(j,k) \in \mathcal{I}$ with probability $p_{i,j,k}$. However, the algorithm takes no action unless a white vertex with j purple neighbours and k yellow neighbours is chosen for the corresponding pair (j,k). So, the modified algorithm undertakes more random choices, but a random choice is irrelevant unless it corresponds to a choice that the original algorithm could have made.

By induction, it is easy to see that the event that, at step i, this new algorithm yields colouring c in a subgraph H of graph G corresponds to the event described in (i), establishing our result. ∎

In the remainder of this work, we shall work in the probability space of the sequence $\mathcal{S}$ of sets of labels. In particular, we shall analyse the performance of Algorithm 3.1.1 by calculating the probability of several events in this probability space, such as the probability of a fixed vertex having a given colour at a fixed time. To achieve this, a pair of independence results will be proved.

The first, which we call *independence of vertex labelling*, shows that the colouring produced by a given labelling in a small connected subgraph of G does not depend on the particular

vertices in the subgraph, where "small" is measured with respect to the girth of the graph. This type of independence is the equivalent for locally greedy algorithms of the independence given by Corollary 2.3.3 for the algorithm of Chapter 2. More precisely, if there is an isomorphism $\phi : V(H') \to V(H)$ between two connected subgraphs H' and H of G, then the event that H has colouring c at time i occurs with the same probability as the event that, at time i, each vertex $v \in V(H')$ has the colour assigned by c to $\phi(v)$. As a consequence, independence of vertex labelling allows us to consider the probability of an unlabelled subgraph of G having a given colouring.

Recall that, if u is a fixed vertex, the definition of the branches around u was given in Definition 2.3.4. Also recall that, by a collection of events $H_1, ..., H_m$ being mutually independent, it is meant that, for any subset of the collection, the joint probability of all events occurring is equal to the product of the probabilities of the individual events. Our second independence result, which we call *conditional independence of branches*, specifies conditions under which the colourings of a set of branches rooted at different neighbours of u are mutually independent. Note that independence of branches is the analog, in the case of locally greedy algorithms, of the independence given by Lemma 2.3.5 for the algorithm of Chapter 2.

We now concentrate on establishing independence of vertex labelling. The following result is a first step in this direction, as it deals with the way in which the influence of a small change in $\mathcal{S}$ propagates as more steps of the algorithm are performed.

Lemma 3.1.5 *Let $G = (V, E)$ be a graph, let $u \in V$ and consider labellings $\mathcal{S}$ and $\hat{\mathcal{S}}$ that only differ at u, that is, $(S^v_{j,k})_{(j,k)\in\mathcal{I}} = (\hat{S}^v_{j,k})_{(j,k)\in\mathcal{I}}$ for all $v \neq u$. Let w be a vertex of G such that $L_{\mathcal{S}}(w) \neq L_{\hat{\mathcal{S}}}(w)$ and let i be the minimum of the labels in the symmetric difference. Then there exists a sequence of vertices $u = w_0, w_1, \ldots, w_t = w$ such that $L_{\mathcal{S}}(w_s) \neq L_{\hat{\mathcal{S}}}(w_s)$ for all s, $d(w_{s-1}, w_s) \leq d + 1$ and $\min\{L_{\mathcal{S}}(w_{s-1}), L_{\hat{\mathcal{S}}}(w_{s-1})\} < \min\{L_{\mathcal{S}}(w_s), L_{\hat{\mathcal{S}}}(w_s)\}$ for all $1 \leq s \leq t$.*

Proof The proof is by induction on i. If $i = 0$, it must be that $u = w$, so the sequence $w_0 = u$ satisfies the properties in the statement of this lemma. Now, let $i > 0$ and assume that this result holds for all smaller values of i. If $u = w$, the result is immediate, so assume this is not the case. By definition of relevant label, one of the sets $\bigcup_{t=0}^{i-1} R_t \cap N_{d-1}(w)$, $P_{i-1} \cap N(w)$ or $Y_{i-1} \cap N(w)$ has different cardinalities with respect to $\mathcal{S}$ and to $\hat{\mathcal{S}}$. This implies that a vertex w' at distance at most $d + 1$ from w satisfies $L_{\mathcal{S}}(w') \neq L_{\hat{\mathcal{S}}}(w')$ with $i' = \min\{L_{\mathcal{S}}(w'), L_{\hat{\mathcal{S}}}(w')\} < i$, since, in the worst case, the sets of yellow neighbours of w with respect to the two labellings are distinct at time $i - 1$, which could be caused by the relevant label of a vertex within distance d from a neighbour of w. By induction, there is a sequence $u = w_0, w_1, \ldots, w_{t-1} = w'$ such that $d(w_{s-1}, w_s) \leq d + 1$, $L_{\mathcal{S}}(w_s) \neq L_{\hat{\mathcal{S}}}(w_s)$ and $\min\{L_{\mathcal{S}}(w_{s-1}), L_{\hat{\mathcal{S}}}(w_{s-1})\} < \min\{L_{\mathcal{S}}(w_s), L_{\hat{\mathcal{S}}}(w_s)\}$ for all $1 \leq s \leq t - 1$. So, we can append w to this sequence to obtain a sequence with the required properties. ∎

Using this, we can immediately bound the distance of vertices whose labels influence the colour of a vertex at time i.

Corollary 3.1.6 *Let $G = (V, E)$ be a graph, let $u \in V$ and consider labellings $\mathcal{S}$ and $\hat{\mathcal{S}}$ that only differ at u. If the colours of $w \in V$ with respect to $\mathcal{S}$ and $\hat{\mathcal{S}}$ are distinct at time i, then $d(u, w) \leq (d+1)i + d$.*

Proof If w has different colours at time i with respect to $\mathcal{S}$ and $\hat{\mathcal{S}}$, then, by definition of colouring, there is a vertex w' and a positive integer i' such that $d(w', w) \leq d$ and $L_{\mathcal{S}}(w') \neq L_{\hat{\mathcal{S}}}(w')$, $\min\{L_{\mathcal{S}}(w'), L_{\hat{\mathcal{S}}}(w')\} = i'$. Lemma 3.1.5 applies and, in particular, $d(u, w') \leq (d+1)i$. As a consequence, $d(u, w) \leq d(u, w') + d(w', w) \leq (d+1)i + d$, as required. ∎

We are now ready to prove independence of vertex labelling. Recall that, if u is a vertex with neighbours $u_1, \ldots, u_r$, and if $s \in \{1, \ldots, r\}$ and $0 < m < g/2$, the rooted tree $T_{u,s,m}$ is the component of $G[\{v : d(u, v) \leq m\} \setminus u]$ rooted at u_s.

Corollary 3.1.7 ***(Independence of vertex labelling)*** *Let $u \in V$ and fix nonnegative integers i and t. Consider $u_{s_1}, \ldots, u_{s_t} \in N(u)$ and positive integers $d'_1, \cdots, d'_t$ such that $2((d+1)i + d + \max\{d'_j\}) < g$. Then the probability that u has colour c and each T_{u,s_j,d'_j} has colouring C_{j,d'_j} at time i is independent of u and of the set $\{u_{s_1}, \ldots, u_{s_t}\}$.*

Proof Let $d' = \max_j\{d'_j\}$. We know from Corollary 3.1.6 that the colours of vertices at distance at most d' from u at time i depend only on the sets of labels of vertices at distance at most $(d+1)i + d + d'$ from u. In other words, the colouring of $\{u\} \cup \bigcup_j T_{u,s_j,d'_j}$ at time i is fully determined by the sets of labels in the subgraph $G_u = G[\{v : d(u, v) \leq (d+1)i + d + d'\}]$. Now, the probability that a vertex has given sets of labels associated with it is the same for every vertex and is independent of the sets of labels associated with any other vertex. Also, the graphs G_u are isomorphic trees for every $u \in V$ by our restriction on i, and the operation of interchanging two branches $T_{u,s,d'}$ and $T_{u,t,d'}$ in G_u induces an automorphism of this tree. The result follows. ∎

An immediate consequence of this result is that the probability of a vertex v being white and having purple neighbours $v_1, \ldots, v_j$ and yellow neighbours $v_{j+1}, \ldots, v_{j+k}$, with the remaining neighbours being white, does not depend on the choice of v or on the choice of the sets of neighbours $\{v_1, \ldots, v_j\}$ and $\{v_{j+1}, \ldots, v_{j+k}\}$. In other words, if $W_i^{j,k}$ is the set of white vertices with exactly j yellow neighbours and k purple neighbours at time i, then $\mathbf{P}(v \in W_i^{j,k})$ does not depend on v and can be denoted by $w_{i,j,k}$.

Before moving to conditional independence of branches, we use the definition of colouring at time i and the results leading to independence of vertex labelling to actually calculate the probability of an event in the probability space of labellings. This example is not fundamental for the comprehension of the results following it. We also note that the calculations are not complete, but are developed up to a point in which the interplay between the colours of the

vertices is understood. Rigorous calculations of the probability of this event are given in the appendix.

Example 3.1.8 *We consider a version of Algorithm 3.1.1 with parameters $r = 3$, $d = 1$ and $\ell = 2$. Let $N = 1$ and fix probabilities $p_0 > 0$ and $p_{1,j,k}$, where $p_{1,0,k} = 0$, for every k, $p_{1,1,0} = p_{1,0} > 0$, $p_{1,1,1} = p_{1,1} > 0$ and $p_{1,1,2} = 1$. Let G be a 3-regular graph with girth larger than 9. We shall calculate the probability of the partial colouring of G in Figure 3.1.1, at time 1. In the figure, the white vertices are represented by unfilled circles and the purple vertices are represented by black circles.*

Colouring 1

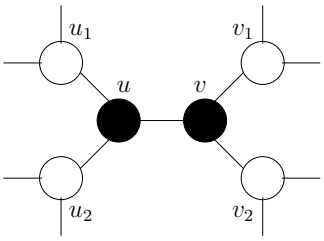

Figure 3.1.1: A partial colouring of G.

We wish to calculate the probability that the adjacent vertices u and v are both purple and that their remaining neighbours are all white at time 1, that is, that they are white after step 1 in an application of the locally greedy algorithm. First note that, because $p_{i,0,k} = 0$ for every $i \geq 1$ and $k \in \{0, \ldots, 3\}$, a vertex can only turn purple at some Step i, $i \geq 1$, if at least one of its neighbours is purple before step i. So, since we would like u_1, u_2, v_1 and v_2 to be white at time 1, at least one of u and v has to be purple at time 0, otherwise they would not become purple at time 1.

We first consider the event for which both u and v are purple at time 0. This happens with probability p_0^2, since each vertex is chosen with probability p_0, and the choices are independent. Now, we also need the remaining neighbours of u and v to be white at time 0. On the one hand, these vertices cannot be chosen at time 0, otherwise they would be purple. On the other hand, since each of them already has a purple neighbour, namely u or v, their other neighbours also cannot be chosen at time 0. Indeed, if one such vertex were chosen its neighbour among u_1, u_2, v_1 or v_2 would have two purple neighbours at time 0, so that it would turn yellow. In total, we are conditioning on 12 vertices not being chosen at time 0, which happens with probability $(1 - p_0)^{12}$.

Conditioning on all the previous events, which occur with a combined probability of $p_0^2(1 - p_0)^{12}$, the colouring of the graph at time 0 is given by Figure 3.1.2. The grey vertices in the figure corresponds to a vertex whose colour is unknown. We observe that the fact that

this figure represents the subgraph containing all vertices at distance at most four of u or v is due to our assumption on the girth.

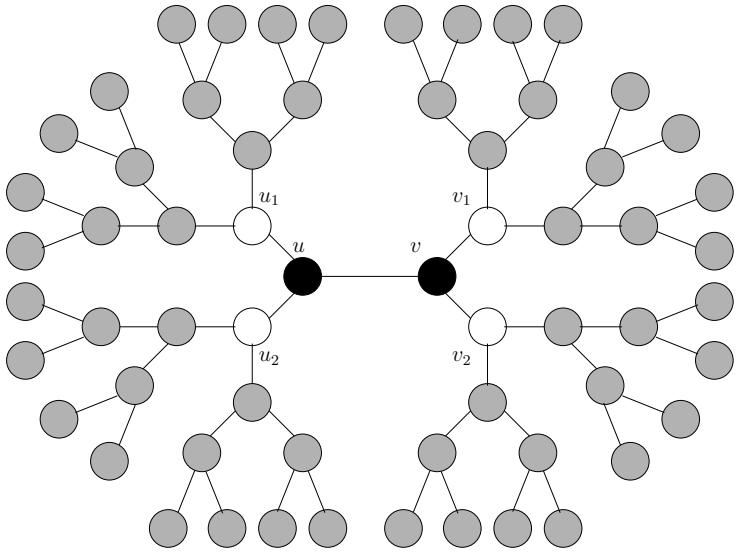

Figure 3.1.2: The vertices whose labels may influence Colouring 1.

Now, define the branches rooted at u_1, u_2, v_1 and v_2 are defined as the trees $T_{u,1,4}$, $T_{u,2,4}$, $T_{v,1,4}$ and $T_{v,2,4}$, respectively. It can be shown that the colourings of the different branches at time 1 are mutually independent. We omit details of this in this sketch. So, to ensure that u_1 is white at time 1, we only need to look at the branch rooted at u_1. Moreover, using independence and symmetry, the probability that u_1, u_2, v_1 and v_2 are all white at time 1 is the fourth power of the probability that u_1 is white at time 1. We look more closely at the branch rooted at u_1, which is given in Figure 3.1.3.

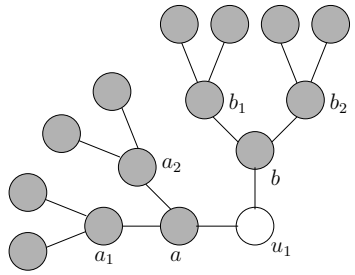

Figure 3.1.3: The vertices in the branch rooted at u_1.

For u_1 to remain white at time 1, it must be that u_1 is not chosen at time 1, and neither are its neighbours a and b. Now, the probability that u_1 is not chosen at time 1 depends on the colours of its neighbours at time 0, since we already know that, at time 0, u_1 has a purple neighbour u and two neighbours a and b. Clearly, the colours of a and b at time 0 depend on the number of purple neighbours that they have. For instance, if a_1, a_2 and b_1 are purple,

but b_2 is not, then a is yellow and b is white at time 0, so u_1 is not chosen with probability $1 - p_{1,1}$. Conditional on this, we know that a cannot become purple at time 1, since it is yellow at time 0. On the other hand, b has exactly one purple neighbour at time 0, so it could turn purple with a probability depending on the colour of b_2, which in turn depends on the number of purple neighbours it has at time 0. With probability p_0^2, both neighbours of b_2 other than b are purple at time 0, so that b_2 is yellow and b is not chosen with probability $1 - p_{1,1}$. With probability $1 - p_0^2$, however, b_2 is not yellow and hence b_2 is not chosen with probability $1 - p_{1,0}$.

In particular, the probability of u_1 being white at time 1 with the assumptions of the previous paragraph is equal to

$$(1 - p_0)^3 p_0 (1 - p_{1,1})(p_0^2(1 - p_{1,1}) + (1 - p_0^2)(1 - p_{1,0})).$$

The product $(1 - p_0)^3 p_0$ corresponds to the event that a_1, a_2 and b_1 are chosen, but b_2 is not chosen. The factor $(1 - p_{1,1})$ is the probability of u_1 given that a is yellow and b is white, while the probability that b is not chosen at time 1 under this circumstances is $(p_0^2(1 - p_{1,1}) + (1 - p_0^2)(1 - p_{1,0}))$. If we consider all the other possible cases in this way, we find the total probability of u_1 being white at time 1. This is done in the appendix. □

This example suggests that the calculation of the probability of an event in the probability space of random labellings may become quite involved if we rely only on the independence results presented so far. So, we consider an additional tool for computing probabilities, which we already introduced as conditional independence of branches. Recall that, with conditional independence of branches, our aim is to specify conditions under which the colourings of a set of branches rooted at different neighbours of u are mutually independent. Our next result is a first step in this direction, since it establishes that, if u and v are white adjacent vertices, then the colouring of branches around u is independent of the colouring of branches around v, where, by the branches around u, we mean the branches rooted at neighbours of u other than v, and vice-versa. To illustrate the importance of conditioning upon u and v being white, we show first that conditional independence of branches does not always hold.

Example 3.1.9 *Consider the instance of Algorithm 3.1.1 and the graph G introduced in Example 3.1.8 with $p_0 = 1/10$, $p_{1,0} = 1/2$ and $p_{1,1} = 3/4$. Let C_1 and C_2 be the events that, at time 1, a labelling of G induces the partial colourings χ_1 and χ_2 given in Figure 3.1.4. Recall that white vertices are unfilled in the picture, while the purple vertices are filled. We show that these events are not independent if we assume that u and v are purple at time 1.*

For C_1 and C_2 to be independent conditional upon u and v being purple at time 1, we must have

$$\mathbf{P}(C_2 \mid C_1) = \mathbf{P}(C_2 \mid u \in P_1 \wedge v \in P_1).$$

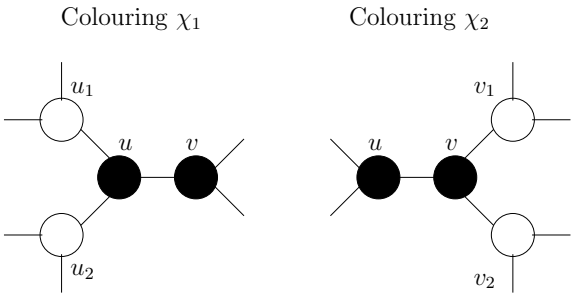

Figure 3.1.4: Two partial colourings of G.

This implies

$$\begin{aligned}&\mathbf{P}(C_1 \wedge C_2)\mathbf{P}(u \in P_1 \wedge v \in P_1)\\ &\quad = \mathbf{P}(C_1)\mathbf{P}(C_2 \mid C_1)\mathbf{P}(u \in P_1 \wedge v \in P_1)\\ &\quad = \mathbf{P}(C_1)\mathbf{P}(C_2).\end{aligned} \tag{3.1.2}$$

Now, the event $\{C_1 \wedge C_2\}$ is precisely the event that, at time 1, a labelling of G induces the partial colouring in Example 3.1.8. On the other hand, the events C_1 and $\{u \in P_1 \wedge v \in P_1\}$ give the colourings in Figure 3.1.5. For $p_0 = 1/10$, $p_{1,0} = 1/2$ and $p_{1,1} = 3/4$, the event

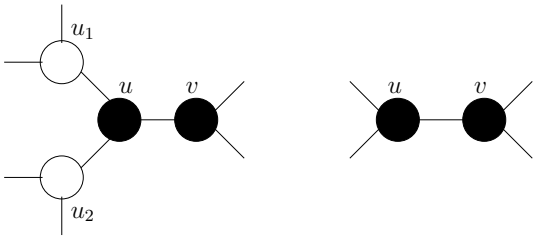

Figure 3.1.5: The colourings induced by $\{C_1 \wedge u \in P_1 \wedge v \in P_1\}$ and $\{u \in P_1 \wedge v \in P_1\}$.

$\{C_1 \wedge C_2\}$ occurs with probability approximately 0.0029, while the other two colourings have probability approximately 0.0205 and 0.0902, respectively. Note that these probabilities may be calculated as in Example 3.1.8.

Now, by independence of vertex labelling, the events C_1 and C_2 occur with the same probability, so the above values are a contradiction to (3.1.2). □

We now establish that, if u and v are white adjacent vertices, then the colouring of branches around u is independent of the colouring of branches around v.

Theorem 3.1.10 *Consider adjacent vertices $u, v \in V$ and fix $i, d' \in \mathbb{N}$ such that $2((d+1)i + d + d') + d + 1 < g$. The neighbours of u and v are denoted by $u_1, \ldots, u_r$ and $v_1, \ldots, v_r$, respectively. Let C_u and C_v be the events that $N_{u,d'} = \bigcup_{u_t \neq v} T_{u,t,d'}$ and $N_{v,d'} = \bigcup_{v_t \neq u} T_{v,t,d'}$ have colourings χ_u and χ_v. Then, conditional upon the event $\{u \in W_i \wedge v \in W_i\}$, the events C_u and C_v are independent.*

Proof We have to show that

$$\mathbf{P}\left(C_u \wedge C_v \mid E\right) = \mathbf{P}\left(C_u \mid E\right)\mathbf{P}\left(C_v \mid E\right),$$

where E is the event $\{u \in W_i \wedge v \in W_i\}$. The main ingredient of the proof is to show that, conditional upon E, the colouring of $N_{u,d'}$ is not influenced by the labelling of the vertices in $\{v\} \cup \bigcup_{v_t \neq u} T_{v,t,(d+1)i+d+d'}$.

We now analyse the colourings at the level of labellings. Given $i' \leq i$, let the i'-branches rooted at u and v be the sets $B_{u,i'} = \{u\} \cup \left(\bigcup_{u_t \neq v} T_{u,t,(d+1)(i-i')+d+d'}\right)$ and $B_{v,i'} = \{v\} \cup \left(\bigcup_{v_t \neq u} T_{v,t,(d+1)(i-i')+d+d'}\right)$, respectively.

Let $\mathcal{S} = [(S_{j,k}(w))_{(j,k)\in\mathcal{I}} : w \in V]$ and $\hat{\mathcal{S}} = [(\hat{S}_{j,k}(w))_{(j,k)\in\mathcal{I}} : w \in V]$ be assignments of labels such that $S_{j,k}(w) = \hat{S}_{j,k}(w)$ if $\min\{d(u,w), d(v,w)\} > (d+1)i + d + d'$, for every j and k. Further suppose that $\mathcal{S}$ and $\hat{\mathcal{S}}$ each imply the event $F = C_u \wedge C_v \wedge E$. Now, let $\mathcal{S}'$ be the assignment of labels defined by $S'_{j,k}(w) = \hat{S}_{j,k}(w)$ if $w \in B_{v,0}$ and $S'_{j,k}(w) = S_{j,k}(w)$ otherwise, for every j and k.

Claim 3.1.11 *For every $i' \leq i$,*

(i) The colourings of the vertices in $B_{u,i'}$ induced by $\mathcal{S}$ and by $\mathcal{S}'$ are the same at time i'.

(ii) The colourings of the vertices in $B_{v,i'}$ induced by $\hat{\mathcal{S}}$ and by $\mathcal{S}'$ are the same at time i'.

Proof of the Claim We use induction on i'.

For the base case $i' = 0$, note that the vertices with relevant label 0 in $B_{u,0}$ with respect to $\mathcal{S}$ and $\mathcal{S}'$ coincide. The vertices with relevant label 0 in $B_{v,0}$ with respect to $\hat{\mathcal{S}}$ and $\mathcal{S}'$ are also clearly the same, whereas all other vertices have relevant label 0 with respect to $\mathcal{S}'$ if and only if they have relevant label 0 with respect to both $\mathcal{S}$ and $\hat{\mathcal{S}}$.

Let $w \in B_{u,0}$ and suppose $w \in P_0$ with respect to $\mathcal{S}$. If $w \in R_0$, then w is also purple with respect to $\mathcal{S}'$ by the previous discussion. Now, if $w' \in R_0 \cap N_{d-1}(w)$, then $w' \in B_{u,0}$ or $(d+1)i + d + d' < d(u,w') < (d+1)(i+1) + d + d'$, otherwise we would have $d(u,w') < d(w,w') \leq d-1$, contradicting the fact that u is white at time i with respect to $\mathcal{S}$ and $\hat{\mathcal{S}}$. Since the restriction on the girth implies that all cycles have length larger than $2((d+1)i+d+d')+d$, we conclude that the vertex w' has the same labels with respect to $\mathcal{S}$ and $\mathcal{S}'$, hence w is also purple with respect to $\mathcal{S}'$ in this case. The same argument applies to the converse, and w is purple at time 0 with respect to $\mathcal{S}$ if $w \in P_0$ with respect to $\mathcal{S}'$. By symmetry, this shows that, if $w \in B_{v,0}$, then $w \in P_0$ with respect to $\mathcal{S}'$ if and only if it is purple at time 0 with respect to $\hat{\mathcal{S}}$. Clearly, the condition on the girth also implies that the vertices at distance exactly $(d+1)i + d + d' + 1$ from u or v are purple at time 0 with respect to $\mathcal{S}'$ if and only if they are purple with respect to $\mathcal{S}$ or $\hat{\mathcal{S}}$, according to the branch they are in.

Now, assume that w is not purple with respect to $\mathcal{S}'$. If $w \in B_{u,0} \cup B_{v,0} \setminus \{u,v\}$, then its neighbours are either in $B_{u,0} \cup N(B_{u,0}) \setminus \{v\}$ or in $B_{v,0} \cup N(B_{v,0}) \setminus \{u\}$. Supposing without

loss of generality that the former is the case, the neighbours are in P_0 with respect to $\mathcal{S}'$ if and only if they are in P_0 with respect to $\mathcal{S}$. Hence, w has the same colour at time 0 with respect to $\mathcal{S}'$ and $\hat{\mathcal{S}}$. If $w \in \{u, v\}$, say $w = u$, then its neighbours in $B_{u,0}$ are purple with respect to $\mathcal{S}'$ if and only if they are purple with respect to $\mathcal{S}$. Also, v is not purple with respect to $\mathcal{S}'$, otherwise it would be so with respect to $\hat{\mathcal{S}}$. We conclude that u is also white with respect $\mathcal{S}'$, and the same argument holds for v by symmetry, establishing the base of induction.

The induction step follows the same lines. Since by the induction hypothesis the colours on $B_{u,i'-1}$ are the same at time $i'-1$, we can show that the relevant labels in the branches $B_{u,i'-1}$ have to be the same at time i'. From here, we repeat the previous argument to show that the colours have to be the same on $B_{u,i'}$, which concludes the proof of the claim. □

We now have

$$\mathbf{P}\left(C_v \mid C_u \wedge E\right) = \sum_{S_u, S_v} \mathbf{P}\left(S_u \wedge S_v \mid C_u \wedge E\right) \mathbf{P}\left(C_v \mid C_u \wedge E \wedge S_u \wedge S_v\right) \tag{3.1.3}$$

where S_u and S_v range over all the possible labellings of $B_{u,0}$ and $B_{v,0}$, respectively, such that $\mathbf{P}\left(S_u \wedge S_v \mid C_u \wedge E\right) > 0$, that is, such that $S_u \wedge S_v$ implies $C_u \wedge u \in W_i \wedge v \in W_i$. (This equivalence is justified by the fact that the vertices not labelled by S_u or S_v cannot influence the colours of $N_{u,d'}$ and $N_{v,d'}$ at time i as a consequence of Corollary 3.1.6.) Now, the probability $\mathbf{P}\left(C_v \mid C_u \wedge E \wedge S_u \wedge S_v\right)$ is equal to 1 if $S_u \wedge S_v$ also implies $C_u \wedge C_v \wedge u \in W_i \wedge v \in W_i$. By our previous claim, this does not depend on the particular S_u and C_u, so

$$\mathbf{P}\left(C_v \mid C_u \wedge E \wedge S_u \wedge S_v\right) = \mathbf{P}\left(C_v \mid E \wedge S_v\right).$$

The right-hand side of Equation (3.1.3) becomes

$$\begin{aligned} &\sum_{S_u, S_v} \mathbf{P}\left(S_u \wedge S_v \mid C_u \wedge E\right) \mathbf{P}\left(C_v \mid E \wedge S_v\right) \\ &= \sum_{S_v} \mathbf{P}\left(C_v \mid E \wedge S_v\right) \sum_{S_u} \mathbf{P}\left(S_u \wedge S_v \mid C_u \wedge E\right) \\ &= \sum_{S_v} \mathbf{P}\left(C_v \mid E \wedge S_v\right) \mathbf{P}\left(S_v \mid C_u \wedge E\right) \end{aligned} \tag{3.1.4}$$

Our claim also implies that $\mathbf{P}\left(S_v \mid C_u \wedge E\right) = \mathbf{P}\left(S_v \mid E\right)$, since it shows that the colouring of $N_{u,d'}$ is not affected by the labelling of $B_{v,0}$ conditional upon the event E. Using this fact, equation (3.1.4) can be rewritten as

$$\sum_{S_v} \mathbf{P}\left(C_v \wedge S_v \mid E\right) = \mathbf{P}\left(C_v \mid E\right).$$

Thus,

$$\mathbf{P}\left(C_u \wedge C_v \mid E\right) = \mathbf{P}\left(C_u \mid E\right) \mathbf{P}\left(C_v \mid E \wedge C_u\right) = \mathbf{P}\left(C_u \mid E\right) \mathbf{P}\left(C_v \mid E\right),$$

which concludes the proof of the theorem. ∎

Conditional independence of branches is an easy consequence Theorem 3.1.10. We prove that, conditional upon knowing the purple and yellow neighbours of a white vertex, the colourings of the branches rooted at the remaining neighbours are mutually independent (that is, the colourings of the branches rooted at the white neighbours of u are mutually independent). Before proving this, we note that it is indeed necessary to condition upon knowing the purple and yellow neighbours. We present an easy example to justify why conditional independence of branches does not hold when this is not the case. Let $d = 1$ and $\ell = 2$, and suppose that u is a white vertex, without conditioning on the colour of any of its neighbours. Clearly, the probability of a given neighbour being purple may be nonzero, since u can have up to one purple neighbour. However, once we know that one of the neighbours is purple, none of the other neighbours can be purple, since u would not be white otherwise, hence the colours of the neighbours of u are not independent. Conditional independence of branches also fails if the yellow neighbours are not known. For, suppose that $\ell = 1$ and that, for some every $i > 0$, $p_{i,0,k}$ is a small positive constant, if $k \leq 1$, and $p_{i,0,k} = 1$, if $k > 1$. Now, if one of the branches around the white u is known to contain a yellow neighbour of u, then the probability that another branch satisfies the same property may be much smaller, since a white vertex u can only have two yellow neighbours for one step, as it would be chosen with probability 1 on the next step.

Corollary 3.1.12 ***(Conditional independence of branches)*** *Let $u \in V$ and $A, B \subseteq N(u)$, and fix nonnegative integers i and d' such that $2((d+1)i + d + d') + d < g$. Consider colourings χ_t of $T_{u,t,d'}$, $u_t \in N(u) \setminus (A \cup B)$ and let C_t be the event that $T_{u,t,d'}$ has colouring χ_t at time i. Then, conditional upon the event $\{u \in W_i \wedge (N(u) \cap P_i = A) \wedge (N(u) \cap Y_i = B)\}$, the events C_t are mutually independent.*

Proof Let $\{u_1, \ldots, u_t\} \subseteq N(u) \setminus (A \cup B)$. We have

$$\begin{aligned}
&\mathbf{P}\left(\bigwedge_{t'=1}^{t} C_{t'} \mid u \in W_i \wedge (N(u) \cap P_i = A) \wedge (N(u) \cap Y_i = B)\right) \\
&= \mathbf{P}\left(C_t \mid u \in W_i \wedge (N(u) \cap P_i = A) \wedge (N(u) \cap Y_i = B)\right) \\
&\quad \times \mathbf{P}\left(\bigwedge_{t'=1}^{t-1} C_{t'} \mid u \in W_i \wedge (N(u) \cap P_i = A) \wedge (N(u) \cap Y_i = B) \wedge C_t\right)
\end{aligned} \tag{3.1.5}$$

Now, the event $u \in W_i \wedge (N(u) \cap P_i = A) \wedge (N(u) \cap Y_i = B) \wedge C_t$ can be rewritten as $u \in W_i \wedge u_t \in W_i \wedge D_u \wedge C_t$, where $D_u = (N(u) \cap P_i = A) \wedge (N(u) \cap Y_i = B) \wedge (u_{t'} \in W_i, t' < t)$. We apply Theorem 3.1.10 to the vertices u and u_t, and to the event $\bigwedge_{t'=1}^{t-1} C_{t'}$, which fixes a

colouring of some of the vertices in $N_{u,d'}$, to obtain

$$\mathbf{P}\left(\bigwedge_{t'=1}^{t-1} C_{t'} \mid u \in W_i \wedge (N(u) \cap P_i = A) \wedge (N(u) \cap Y_i = B) \wedge C_t\right)$$
$$= \mathbf{P}\left(\bigwedge_{t'=1}^{t-1} C_{t'} \mid u \in W_i \wedge (N(u) \cap P_i = A) \wedge (N(u) \cap Y_i = B)\right).$$

Thus, equation (3.1.5) becomes

$$\mathbf{P}(C_t \mid u \in W_i \wedge (N(u) \cap P_i = A) \wedge (N(u) \cap Y_i = B)$$
$$\times \mathbf{P}(\bigwedge_{t'=1}^{t-1} C_{t'} \mid u \in W_i \wedge (N(u) \cap P_i = A) \wedge (N(u) \cap Y_i = B))$$

and, repeating the same argument for the other branches, we obtain

$$\prod_{t'=1}^{t} \mathbf{P}(C_{t'} \mid u \in W_i \wedge (N(u) \cap P_i = A) \wedge (N(u) \cap Y_i = B)),$$

as required. ∎

In the previous result, we conditioned upon knowing the sets A and B of purple and yellow neighbours of a vertex u. However, Theorem 3.1.10 can also be used to show that the particular sets A and B do not affect the probability that one of the remaining branches has a given colouring.

Corollary 3.1.13 *Let $u \in V$, $A, B \subseteq N(u)$ and $u_t \in N(u) \setminus (A \cup B)$, and fix $i \in \mathbb{N}$ such that $2((d+1)i + d + d') + d + 1 < g$. Let χ_t be a colouring of the branch $T_{u,t,d'}$ and consider the event C_t that $T_{u,t,d'}$ has colouring χ_t at time i. Then*

$$\mathbf{P}(C_t \mid u \in W_i \wedge (N(u) \cap P_i = A) \wedge (N(u) \cap Y_i = B)) = \mathbf{P}(C_t \mid u \in W_i \wedge u_t \in W_i).$$

Proof Let C_u denote the colouring of $N(u) \setminus \{v\}$ such that the vertices in A are purple, the vertices in B are yellow and the remaining vertices are white. We have

$$\mathbf{P}(C_t \mid u \in W_i \wedge (N(u) \cap P_i = A) \wedge (N(u) \cap Y_i = B))$$
$$= \mathbf{P}(C_t \mid u \in W_i \wedge u_t \in W_i \wedge C_u).$$

Applying Theorem 3.1.10 to the white vertices u and u_t and the colourings C_t and C_u, we deduce that

$$\mathbf{P}(C_t \mid u \in W_i \wedge u_t \in W_i \wedge C_u) = \mathbf{P}(C_t \mid u \in W_i \wedge u_t \in W_i),$$

which is independent of C_u, hence of A and B. ∎

An important special case of this result is that $\mathbf{P}\left(v \in W_i^{j,k} \mid u \in W_i^{j',k'} \wedge v \in W_i\right)$ is independent of j', k' and can be denoted by $q_{i,j,k}$. A formula for $q_{i,j,k}$ can be obtained in terms of the probabilities $w_{i,j',k'}$, as seen in the following corollary.

Corollary 3.1.14

$$q_{i,j,k} = \frac{(r-j-k)w_{i,j,k}}{\sum_{j',k'}(r-j'-k')w_{i,j',k'}}.$$

Proof The expected number of pairs $(a,b) \in V^2$ such that $a \in W_i^{j,k}$ and $b \in W_i^{j',k'}$ is given by

$$e(j,j',k,k') = \sum_{a \in V} \sum_{b \in N(a)} \mathbf{P}\left(b \in W_i^{j',k'} \wedge a \in W_i\right) \mathbf{P}\left(a \in W_i^{j,k} \mid b \in W_i^{j',k'} \wedge a \in W_i\right).$$

By definition, we have $\mathbf{P}\left(a \in W_i^{j,k} \mid b \in W_i^{j',k'} \wedge a \in W_i\right) = q_{i,j,k}$.

Now, $b \in W_i^{j',k'}$ has exactly $r-j'-k'$ white neighbours, and, conditional on this fact, the probability that a given neighbour of u is white does not depend on the particular neighbour, by the independence of vertex labelling given in Corollary 3.1.7. This implies

$$\mathbf{P}\left(b \in W_i^{j',k'} \wedge a \in W_i\right) = w_{i,j',k'}\mathbf{P}(a \in W_i \mid b \in W_i^{j',k'}) = (r-j'-k')w_{i,j',k'}/r,$$

from which

$$e(j,j',k,k') = (r-j'-k')w_{i,j',k'}q_{i,j,k}n$$

But the expected number of pairs (a,b), where $a \in W_i^{j,k}$ and $b \in W_i$, is equal to $(r-j-k)w_{i,j,k}n$, since G is expected to contain $w_{i,j,k}n$ vertices in $W_i^{j,k}$, and each of these vertices is adjacent to exactly $r-j-k$ white vertices. We have

$$(r-j-k)w_{i,j,k}n = \sum_{j',k'}^{r-1} n(r-j'-k')w_{i,j',k'}q_{i,j,k}n,$$

that is,

$$q_{i,j,k} = \frac{(r-j-k)w_{i,j,k}}{\sum_{j',k'}(r-j'-k')w_{i,j',k'}}. \quad \blacksquare$$

3.2 Recurrence equations

Let $r \geq 3$, $\ell \leq r$ and d be fixed positive integers. Also fix $N \in \mathbb{N}$ and the probabilities p_0 and $(p_{i,j,k})_{0 \leq i \leq N,(j,k) \in \mathcal{I}}$. Let G be an r-regular graph with girth g greater than $2((d+1)N+d)+d+1$ and consider an application of Algorithm 3.1.1 to G with the quantities fixed above.

The work in the previous section ensures that the values of $w_{i,j,k} = \mathbf{P}(u \in W_i^{j,k})$, where u is a vertex of G, are independent of the particular vertex u and even of the particular graph G. Our objective in the current section is to calculate these probabilities. Although the events whose probability is being considered are not the same, this will resemble the derivation of the recurrence equations in Section 2.4.

In the case when $i = 0$, these numbers depend only on r, d, ℓ and p_0, and can be easily calculated. We now calculate them in the case of independent sets, since, by Proposition 3.1.2,

a locally greedy algorithm for independent sets is obtained if the parameters are $d = \ell = 1$. Because $\ell = 1$, a white vertex cannot have purple neighbours, so we only need to calculate the probability that a vertex v is white and has exactly k yellow neighbours after step 0. Clearly, this event occurs if and only if v does not have relevant label 0, none of its neighbours do, and exactly k of its neighbours are adjacent to a vertex with relevant label 0. Now, a vertex is assigned label 0 at random, independently of all other vertices, with probability p_0. In particular, the probability that neither v nor its neighbours have relevant label 0 is $(1-p_0)^{r+1}$. Conditional on this event, if we look at a fixed neighbour of v, the event that none of its neighbours have relevant label 0 has probability $(1-p_0)^{r-1}$ (since we already know that v does not have relevant label 0). Because exactly k neighbours of v must be adjacent to at least one vertex with relevant label zero, it follows that

$$w_{0,0,k} = (1-p_0)^{r+1}\binom{r}{k}(1-p_0)^{(r-1)(r-k)}\left(1-(1-p_0)^{r-1}\right)^k.$$

With similar arguments, the following formula can be obtained for $d = 1$ and $\ell \in \{1, \ldots, r-1\}$ may be obtained. Observe that this case includes the locally greedy algorithm for induced forests suggested in Proposition 3.1.2, since the parameters are $d = 1$ and $\ell = 2$.

$$\begin{aligned} w_{0,j,k} = \binom{r}{j}\binom{r-j}{k}(1-p_0)^{r-j+1}\left(1-\sum_{s=0}^{\ell-1}\binom{r-1}{s}p_0^s(1-p_0)^{r-s)}\right)^k \\ \times\left(\sum_{s=0}^{\ell-1}\binom{r-1}{s}p_0^s(1-p_0)^{r-s}\right)^{r-j-k} \end{aligned} \tag{3.2.1}$$

Henceforth, for fixed r, d and ℓ, we shall use $w_{0,j,k}(p_0)$ to denote the probability of the event $u \in W_0^{j,k}$ as a function of the initial probability.

We now consider $w_{i,j,k}$ in the case $i > 0$. We shall express $\mathbf{w}_i$ as a function of $\mathbf{p}_i$ and $\mathbf{w}_{i-1}$, giving a recurrence relation that allows us to inductively calculate all the probabilities. The recurrence equations we aim to find are of the form

$$\mathbf{w}_i = \mathbf{w}_{i-1} + \mathbf{F}(\mathbf{p}_i, \mathbf{w}_{i-1}) + \mathbf{E}(\mathbf{p}_i, \mathbf{w}_{i-1}),\ i = 1, \ldots, N, \tag{3.2.2}$$

where $\mathbf{F} = (F_{j,k})_{(j,k)\in\mathcal{I}}$ and $\mathbf{E} = (E_{j,k})_{(j,k)\in\mathcal{I}}$ are vector functions with domain $\mathbb{R}^{2|\mathcal{I}|}$ and range $\mathbb{R}^{|\mathcal{I}|}$ satisfying the following properties:

- $F_{j,k}$ and $E_{j,k}$ are polynomials in the first $|\mathcal{I}|$ variables (that is, on the variables corresponding to the vector $\mathbf{p}_i$) whose coefficients are rational functions on the remaining $|\mathcal{I}|$ variables (which correspond to $\mathbf{w}_{i-1}$).
- $F_{j,k}$, as a polynomial in the first $|\mathcal{I}|$ variables, is homogeneous of degree one (that is, all its monomials have degree one in $\mathbf{p}_i$).

- $E_{j,k}$, as a polynomial in the first $|\mathcal{I}|$ variables, contains only monomials of degree larger than one.

Clearly, if such $\mathbf{F}$ and $\mathbf{E}$ exist, they are unique, and $\mathbf{F}$ will be called *the function associated with Algorithm 3.1.1* in this case, whereas $\mathbf{E}$ is *the error function associated with Algorithm 3.1.1.* The reason for these names is that, when the probabilities are small, the value of $E_{j,k}$ is negligible in comparison with $F_{j,k}$, so the behaviour of the Algorithm is essentially described by F.

In the remainder of this section, we show that recurrence equations of this form may be found for the case $d = 1$ and $\ell = 1$ and the case $d = 1$ and $\ell = 2$, which will play a fundamental role in the analysis of locally greedy algorithms for independent sets and induced forests. Our first lemma, however, holds for any $\ell \in \{1, \dots, r\}$.

Before stating this result, we prove an auxiliary result, which is once more a form of conditional independence of branches, but is proved only in the case $d = 1$. Here, conditional upon a vertex u being white at time i and not having relevant label $i + 1$, we show that, at time $i + 1$, the colourings of the branches around u rooted at white vertices at time i are mutually independent.

Corollary 3.2.1 *Let $u \in V$ and $A, B \subseteq N(u)$, and fix nonnegative integers i and d' such that $2((d+1)(i+1)+d+d')+d < g$. Consider colourings χ_t of $T_{u,t,d'}$, $u_t \in N(u) \setminus (A \cup B)$ and let C_t be the event that $T_{u,t,d'}$ has colouring χ_t at time $i+1$. Then, conditional upon the event $E(A,B) = u \in W_i \wedge (N(u) \cap P_i = A) \wedge (N(u) \cap Y_i = B) \wedge u \notin P_{i+1}$, the events C_t are mutually independent. Moreover, the probabilities $\mathbf{P}(C_t \mid E(A,B))$ are independent of A and B, for every t.*

Proof We look at $\mathbf{P}\left(\bigwedge_t C_t \mid E(A,B)\right)$. In other words, we aim to compute the probability that each branch rooted at a white neighbour u_t of u at time i has colouring C_t at time $i + 1$. We are given the colours of u and its neighbours at time i, and the additional piece of information that u does not have relevant label $i + 1$.

Using this, the colouring of a given branch $T_{u,t,d'}$ at time $i+1$ only depends on the vertices with relevant label $i + 1$ in the branch $T_{u,t,d'+1}$, by the definition of colouring at time $i + 1$ (see Definition 3.1.3). But this dependency is actually on the colouring of this branch at time i, since the labels are given in advance, independently of each other. Now, conditional on the event $u \in W_i \wedge (N(u) \cap P_i = A) \wedge (N(u) \cap Y_i = B)$, the colourings of the branches rooted at vertices in $N(u) \setminus (A \cup B)$ are mutually independent by Corollary 3.1.12. Furthermore, they do not depend on the sets A and B, by Corollary 3.1.13. This establishes our result. ∎

We are now ready to establish the following recurrence formula for $w_{i,j,k} = \mathbf{P}(u \in W_i^{j,k})$.

Lemma 3.2.2 *Fix $d = 1$ and let r, ℓ and N be positive integers $\ell \leq r$. Fix probabilities p_0 and $(p_{i,j,k})_{1 \leq i \leq N, (j,k) \in \mathcal{I}}$. Let G be an r-regular graph with girth greater than $4N + 4$ and*

consider adjacent vertices u and v in G. Then, if $(j,k) \in \mathcal{I}$ and $1 \le i \le N$,

$$w_{i,j,k} = \sum_{j'=0}^{j} \sum_{k'=0}^{k} \binom{r-j'-k'}{j-j'} \binom{r-j-k'}{k-k'} (1-p_{i,j',k'}) w_{i-1,j',k'} \\ \times \mathbf{P}(v \in W_i \mid E_{i-1})^{r-j-k} \mathbf{P}(v \in P_i \mid E_{i-1})^{j-j'} \mathbf{P}(v \in Y_i \mid E_{i-1})^{k-k'},$$

where E_{i-1} denotes the event $\{u \in W_{i-1} \wedge v \in W_{i-1} \wedge u \notin P_i\}$.

Proof By definition of $w_{i,j,k}$ and using conditional probability,

$$\begin{aligned} w_{i,j,k} &= \mathbf{P}(u \in W_i^{j,k}) \\ &= \sum_{j'=0}^{j} \sum_{k'=0}^{k} \mathbf{P}(u \in W_{i-1}^{j',k'}) \mathbf{P}(u \in W_i^{j,k} \mid u \in W_{i-1}^{j',k'}) \\ &= \sum_{j'=0}^{j} \sum_{k'=0}^{k} w_{i-1,j',k'} \mathbf{P}(u \notin P_i \mid u \in W_{i-1}^{j',k'}) \mathbf{P}(u \in W_i^{j,k} \mid u \in W_{i-1}^{j',k'} \wedge u \notin P_i) \\ &= \sum_{j'=0}^{j} \sum_{k'=0}^{k} w_{i-1,j',k'} (1-p_{i,j',k'}) \mathbf{P}(u \in W_i^{j,k} \mid u \in W_{i-1}^{j',k'} \wedge u \notin P_i). \end{aligned} \tag{3.2.3}$$

Note that, when we say that $\mathbf{P}(u \notin P_i \mid u \in W_{i-1}^{j',k'}) = 1 - p_{i,j',k'}$, we are assuming that u becomes purple at time i if and only if it has relevant label 1, which is the case when $d = 1$.

To analyse the term $\mathbf{P}(u \in W_i^{j,k} \mid u \in W_{i-1}^{j',k'} \wedge u \notin P_i)$, we sum over all possible distributions of colours around u at time $i-1$, that is, over the set $\mathcal{A} = \{(A,B) : A, B \subseteq N(u),\ |A| = j',\ |B| = k',\ A \cap B = \emptyset\}$. We have

$$\begin{aligned} &\mathbf{P}(u \in W_i^{j,k} \mid u \in W_{i-1}^{j',k'} \wedge u \notin P_i) \\ &\quad = \sum_{(A,B) \in \mathcal{A}} \mathbf{P}(E(A,B) \mid u \in W_{i-1}^{j',k'} \wedge u \notin P_i) \\ &\quad \times \mathbf{P}(u \in W_i^{j,k} \mid E(A,B) \wedge u \in W_{i-1} \wedge u \notin P_i), \end{aligned} \tag{3.2.4}$$

where $E(A,B) = \{(N(u) \cap P_{i-1} = A) \wedge (N(u) \cap Y_{i-1} = B)\}$.

But, by assuming that u does not have relevant label i, u moves from $W_{i-1}^{j',k'}$ to $W_i^{j,k}$ if and only if exactly $j - j'$ of its white neighbours at time $i-1$ turn purple and $k - k'$ of the white neighbours at time $i-1$ turn yellow at time i. Here, we are again using the restriction $d = 1$. In more general cases, a yellow vertex might become purple by being at distance at most $d-1$ of a vertex with relevant label i. So, u is in $W_i^{j,k}$ conditional upon the event $\{E(A,B) \wedge u \in W_{i-1} \wedge u \notin P_i\}$ if and only if the event $F(A',B') = (N(u) \cap P_i \cap W_{i-1} = A') \wedge (N(u) \cap Y_i \cap W_{i-1} = B')$ holds for some sets $A', B' \subseteq N(u) \setminus (A \cup B)$, $|A'| = j - j'$, $|B'| = k - k'$.

Let $\mathcal{A}' = \{(A', B') : A', B' \subseteq N(u) \setminus (A \cup B),\ |A'| = j - j',\ |B| = k - k',\ A' \cap B' = \emptyset\}$. Then

$$\begin{aligned}
&\mathbf{P}(u \in W_i^{j,k} \mid E(A,B) \wedge u \in W_{i-1} \wedge u \notin P_i) \\
&= \sum_{(A',B') \in \mathcal{A}'} \mathbf{P}(F(A',B') \mid E(A,B) \wedge u \in W_{i-1} \wedge u \notin P_i) \\
&\quad \times \mathbf{P}(u \in W_i^{j,k} \mid F(A',B') \wedge E(A,B) \wedge u \in W_{i-1} \wedge u \notin P_i)) \\
&= \sum_{(A',B') \in \mathcal{A}'} \mathbf{P}(F(A',B') \mid E(A,B) \wedge u \in W_{i-1} \wedge u \notin P_i),
\end{aligned} \tag{3.2.5}$$

since $\mathbf{P}(u \in W_i^{j,k} \mid F(A',B') \wedge E(A,B) \wedge u \in W_{i-1} \wedge u \notin P_i)) = 1$.

We look at $\mathbf{P}(F(A',B') \mid E(A,B) \wedge u \in W_{i-1} \wedge u \notin P_i)$. It involves the colourings at time i of branches known to be rooted at white vertices at time $i-1$. By Corollary 3.2.1, the colourings of these branches at time i are mutually independent, and the probabilities that given colourings are achieved are independent of A and B. As a consequence, if E_{i-1} denotes the event $\{u \in W_{i-1} \wedge v \in W_{i-1} \wedge u \notin P_i\}$ and $N(A, A', B, B') = N(u) \setminus (A \cup A' \cup B \cup B')$, we have

$$\begin{aligned}
&\mathbf{P}(F(A',B') \mid E(A,B) \wedge u \in W_{i-1} \wedge u \notin P_i)) \\
&= \prod_{v \in N(A,A',B,B')} \mathbf{P}(v \in W_i \mid E_{i-1}) \prod_{v \in A'} \mathbf{P}(v \in P_i \mid E_{i-1}) \prod_{v \in B'} \mathbf{P}(v \in Y_i \mid E_{i-1}) \\
&= \mathbf{P}(v \in W_i \mid E_{i-1})^{r-j-k} \mathbf{P}(v \in P_i \mid E_{i-1})^{j-j'} \mathbf{P}(v \in Y_i \mid E_{i-1})^{k-k'}.
\end{aligned} \tag{3.2.6}$$

The last equality comes from the fact that all these probabilities are independent of the particular vertex v by Corollary 3.1.7.

Equation (3.2.5) sums this same expression $|\mathcal{A}'|$ times, which leads to the multiplicative factor $\binom{r-j'-k'}{j-j'}\binom{r-j-k'}{k-k'}$. Now, the terms $\mathbf{P}(u \in W_i^{j,k} \mid E(A,B) \wedge u \in W_{i-1} \wedge u \notin P_i)$ in equation (3.2.4) do not depend on (A, B), so that equation (3.2.3) becomes

$$\begin{aligned}
w_{i,j,k} = \sum_{j'=0}^{j} \sum_{k'=0}^{k} &\binom{r-j'-k'}{j-j'} \binom{r-j-k'}{k-k'} (1 - p_{i,j',k'}) w_{i-1,j',k'} \\
&\times \mathbf{P}(v \in W_i \mid E_{i-1})^{r-j-k} \mathbf{P}(v \in P_i \mid E_{i-1})^{j-j'} \mathbf{P}(v \in Y_i \mid E_{i-1})^{k-k'},
\end{aligned}$$

establishing our result. ∎

This lemma tells us that a recurrence relation for $w_{i,j,k}$ will be obtained if we express the probabilities $\mathbf{P}(v \in W_i \mid E_{i-1})$, $\mathbf{P}(v \in P_i \mid E_{i-1})$ and $\mathbf{P}(v \in Y_i \mid E_{i-1})$ in terms of $\mathbf{w}_{i-1}$ and $\mathbf{p}_i$. These probabilities depend on the value of ℓ. Indeed, if $\ell = 1$, for v to be white at time i conditional on the event $\{u \in W_{i-1} \wedge v \in W_{i-1} \wedge u \notin P_i\}$, it must be that neither v nor its white neighbours at time $i-1$ have relevant label i. However, for larger values of ℓ, some of the white neighbours of v are allowed to have relevant label i. We now look more closely at the cases when $\ell = 1$ and $\ell = 2$.

3.3 Recurrence equations for independent sets

We consider the case $d = 1$ and $\ell = 1$, for which white vertices never have purple neighbours. Since j is always equal to zero, we omit the value of j from the notation, that is, $W_i^{j,k} = W_i^k$, $w_{i,j,k} = w_{i,k}$ and $q_{i,j,k} = q_{i,k}$. To obtain a recurrence relation for $\mathbf{w}_i$ from Lemma 3.2.2, we need to calculate $\mathbf{P}(u \in W_i \mid E_{i-1})$ and $\mathbf{P}(u \in Y_i \mid E_{i-1})$ in terms of $\mathbf{p}_i$ and $\mathbf{w}_{i-1}$, where E_{i-1} is the event $\{u \in W_{i-1} \wedge v \in W_{i-1} \wedge u \notin P_i\}$.

We calculate $\mathbf{P}(u \in W_i \mid E_{i-1})$. In the first step, we use conditional probabilities to fix the number of yellow neighbours of v at time $i-1$.

$$\mathbf{P}\left(v \in W_i \mid E_{i-1}\right) = \sum_{k''=0}^{r-1} \mathbf{P}\left(v \in W_{i-1}^{k''} \mid E_{i-1}\right) \mathbf{P}\left(v \in W_i \mid v \in W_{i-1}^{k''} \wedge E_{i-1}\right).$$

Note that, once we condition on $\{u \in W_{i-1}\}$, the event $\{u \notin P_i\}$ influences the colouring of the graph at time $i-1$ through the number of purple and yellow neighbours of u. Now, as a consequence of Corollary 3.1.13, the probability $q_{i-1,k''}$ that v lies in $W_{i-1}^{k''}$ given that u and v are white at time $i-1$ is independent of the set of purple and yellow neighbours of u, so $\mathbf{P}\left(v \in W_{i-1}^{k''} \mid E_{i-1}\right) = q_{i-1,k''}$.

Now, the term $\mathbf{P}\left(v \in W_i \mid v \in W_{i-1}^{k''} \wedge E_{i-1}\right)$ in the above equation may be rewritten as

$$\begin{aligned} &\mathbf{P}\left(v \in W_i \mid v \in W_{i-1}^{k''} \wedge E_{i-1}\right) \\ &\quad = \mathbf{P}\left(v \notin P_i \mid v \in W_{i-1}^{k''} \wedge E_{i-1}\right) \mathbf{P}\left(v \in W_i \mid v \notin P_i \wedge v \in W_{i-1}^{k''} \wedge E_{i-1}\right). \end{aligned} \tag{3.3.1}$$

We have $\mathbf{P}\left(v \notin P_i \mid v \in W_{i-1}^{k''} \wedge E_{i-1}\right) = 1 - p_{i,k''}$, since v is not in P_i if and only if it does not have relevant label i. We now look at the term $\mathbf{P}\left(v \in W_i \mid v \notin P_i \wedge v \in W_{i-1}^{k''} \wedge E_{i-1}\right)$. Since we are conditioning upon u and v not having relevant label i, the event $v \in W_i$ holds if and only if none of the remaining $r - k'' - 1$ white neighbours of v at time $i-1$ have relevant label i. Let S be the set of white neighbours of v at time $i-1$, with the exception of u. We have

$$\begin{aligned} &\mathbf{P}\left(v \in W_i \mid v \notin P_i \wedge v \in W_{i-1}^{k''} \wedge E_{i-1}\right) \\ &\quad = \mathbf{P}\left(\bigwedge_{w \in S} w \notin P_i \mid v \notin P_i \wedge v \in W_{i-1}^{k''} \wedge E_{i-1}\right). \end{aligned}$$

This last expression concerns the colourings, at time i, of branches around a white vertex v rooted at white vertices at time $i-1$. Moreover, v does not contain relevant label i. We may apply Corollary 3.2.1 to conclude that these colourings are independent, so that

$$\mathbf{P}\left(v \in W_i \mid v \notin P_i \wedge v \in W_{i-1}^{k''} \wedge E_{i-1}\right) = \prod_{w \in S} \mathbf{P}\left(w \notin P_i \mid v \notin P_i \wedge v \in W_{i-1}\right).$$

In this equation, the term $\mathbf{P}\left(w \notin P_i \mid v \notin P_i \wedge v \in W_{i-1}^{k''} \wedge E_{i-1}\right)$ has been turned into the term $\mathbf{P}\left(w \notin P_i \mid v \notin P_i \wedge v \in W_{i-1}\right)$. This can be done because, by Corollary 3.2.1, events such as the number of yellow neighbours of v or the colouring of u and its neighbours at time $i-1$ do not affect the colour of w at time i conditional upon $v \notin P_i \wedge v \in W_{i-1}$.

Finally, for any $w \in S$,

$$\begin{aligned}
&\mathbf{P}\left(w \notin P_i \mid v \notin P_i \wedge v \in W_{i-1}\right) \\
&\quad = \sum_{k^\star=0}^{r-1} \mathbf{P}\left(w \in W_{i-1}^{k^\star} \mid v \notin P_i \wedge v \in W_{i-1}\right) \mathbf{P}\left(w \notin P_i \mid w \in W_{i-1}^{k^\star} \wedge v \notin P_i \wedge v \in W_{i-1}\right) \\
&\quad = \sum_{k^\star=0}^{r-1} q_{i-1,k^\star}\left(1 - p_{i,k^\star}\right).
\end{aligned}$$

Using this, equation (3.3.1) may be rewritten as

$$\mathbf{P}\left(v \in W_i \mid v \notin P_i \wedge v \in W_{i-1}^{k''} \wedge E_{i-1}\right) = \left(1 - p_{i,k''}\right)\left(\sum_{k^\star=0}^{r-1} q_{i-1,k^\star}\left(1 - p_{i,k^\star}\right)\right)^{r-k''-1},$$

from which we conclude that

$$\mathbf{P}\left(v \in W_i \mid E_{i-1}\right) = \sum_{k''=0}^{r-1} q_{i-1,k''}(1 - p_{i,k''})\left(\sum_{k^\star=0}^{r-1} q_{i-1,k^\star}(1 - p_{i,k^\star})\right)^{r-k''-1}. \tag{3.3.2}$$

We also have

$$\mathbf{P}\left(v \in Y_i \mid E_{i-1}\right) = \sum_{k''=0}^{r-1} q_{i-1,k''}(1 - p_{i,k''})\left(1 - \left(\sum_{k^\star=0}^{r-1} q_{i-1,k^\star}(1 - p_{i,k^\star})\right)^{r-k''-1}\right). \tag{3.3.3}$$

The argument for this case follows exactly the same steps of the case $v \in W_i$, until we reach an equivalent version of equation (3.3.1), namely

$$\begin{aligned}
&\mathbf{P}\left(v \in Y_i \mid v \in W_{i-1}^{k''} \wedge E_{i-1}\right) \\
&\quad = \mathbf{P}\left(v \notin P_i \mid v \in W_{i-1}^{k''} \wedge E_{i-1}\right) \mathbf{P}\left(v \in Y_i \mid v \notin P_i \wedge v \in W_{i-1}^{k''} \wedge E_{i-1}\right).
\end{aligned}$$

Here, $\mathbf{P}\left(v \in Y_i \mid v \notin P_i \wedge v \in W_{i-1}^{k''} \wedge E_{i-1}\right) = 1 - \mathbf{P}\left(v \in W_i \mid v \notin P_i \wedge v \in W_{i-1}^{k''} \wedge E_{i-1}\right)$, since any non-purple vertex is either white or yellow, which leads to the formula in equation (3.3.3).

Now that we have expressions for $\mathbf{P}\left(v \in W_i \mid E_{i-1}\right)$ and $\mathbf{P}\left(v \in Y_i \mid E_{i-1}\right)$ in terms of $\mathbf{w}_{i-1}$

and $\mathbf{p}_i$, we may apply Lemma 3.2.2 to obtain

$$\begin{aligned} w_{i,k} &= \mathbf{P}\left(u \in W_i^k\right) \\ &= \sum_{k'=0}^{k} w_{i-1,k'}\left(1-p_{i,k'}\right)\binom{r-k'}{k-k'} \\ &\quad \times \left(\sum_{k''=0}^{r-1} q_{i-1,k''}(1-p_{i,k''})\left(\sum_{k^\star=0}^{r-1} q_{i-1,k^\star}(1-p_{i,k^\star})\right)^{r-k''-1}\right)^{r-k} \\ &\quad \times \left(\sum_{k''=0}^{r-1} q_{i-1,k''}(1-p_{i,k''})\left(1-\left(\sum_{k^\star=0}^{r-1} q_{i-1,k^\star}(1-p_{i,k^\star})\right)^{r-k''-1}\right)\right)^{k-k'}. \end{aligned}$$

Expanding this and using the formula for $q_{i-1,k}$ given in Corollary 3.1.14, we obtain

$$\begin{aligned} w_{i,k} &= w_{i-1,k} - p_{i,k}w_{i-1,k} - (r-k)w_{i-1,k}\sum_{k''=0}^{r-1}\frac{p_{i,k''}(r-k'')w_{i-1,k''}}{s_{i-1}} \\ &+ \left((r-k+1)w_{i-1,k-1}\delta_{k\geq 1} - (r-k)w_{i-1,k}\right)\left(\sum_{k''=0}^{r-2}\frac{(r-k'')(r-k''-1)w_{i-1,k''}}{s_{i-1}^2}\right) \\ &\times \left(\sum_{k^\star=0}^{r-1} p_{i,k^\star}(r-k^\star)w_{i-1,k^\star}\right) + E_k(\mathbf{p},\mathbf{w}), \end{aligned} \tag{3.3.4}$$

where $s_i = \sum_{t=0}^{r-1}(r-t)w_{i,t}$ and $E_k(\mathbf{p},\mathbf{w})$ is a polynomial in the indeterminates $p_{i,k}$ for which every monomial has degree at least two. The coefficients are rational functions in the indeterminates $\mathbf{w}_{i-1}$ whose denominators are powers of $s_{i-1} = s_{i-1}(\mathbf{w})$. In particular, the function $\mathbf{E} = (E_k)_{k=0}^r$ is the error function defined in equation (3.2.2). The function associated with the algorithm $\mathbf{F} = (F_k)_{k=0}^r$ is given by

$$\begin{aligned} F_k(\mathbf{p},\mathbf{w}) &= -p_{i,k}w_{i-1,k} - (r-k)w_{i-1,k}\sum_{k''=0}^{r-1}\frac{p_{i,k''}(r-k'')w_{i-1,k''}}{s_{i-1}} \\ &+ \left((r-k+1)w_{i-1,k-1}\delta_{k\geq 1} - (r-k)w_{i-1,k}\right) \\ &\times \left(\sum_{k''=0}^{r-2}\frac{(r-k'')(r-k''-1)w_{i-1,k''}}{s_{i-1}^2}\right)\left(\sum_{k^\star=0}^{r-1} p_{i,k^\star}(r-k^\star)w_{i-1,k^\star}\right). \end{aligned} \tag{3.3.5}$$

3.4 Recurrence equations for induced forests

We now consider the case when $d = 1$ and $\ell = 2$, for which $\mathcal{I} = \{(j,k) : j,k \geq 0, j \in \{0,1\}, j+k \leq r\}$. Recall that we want to calculate $\mathbf{P}\left(v \in W_i \mid E_{i-1}\right)$, $\mathbf{P}\left(v \in P_i \mid E_{i-1}\right)$ and $\mathbf{P}\left(v \in Y_i \mid E_{i-1}\right)$ to obtain a recurrence equation for $w_{i,j,k}$ through Lemma 3.2.2.

The formula for $\mathbf{P}\left(v \in P_i \mid E_{i-1}\right)$ can be obtained as follows. First, we fix the colouring of the neighbourhood of v at time $i-1$. Now, there is the possibility that a white vertex has

a purple neighbour, so

$$\mathbf{P}\left(v \in P_i \mid E_{i-1}\right) = \sum_{k''=0}^{r-2} \mathbf{P}\left(v \in W_{i-1}^{1,k''} \mid E_{i-1}\right) \mathbf{P}\left(v \in P_i \mid v \in W_{i-1}^{1,k''} \wedge E_{i-1}\right) + \sum_{k''=0}^{r-1} \mathbf{P}\left(v \in W_{i-1}^{0,k''} \mid E_{i-1}\right) \mathbf{P}\left(v \in P_i \mid v \in W_{i-1}^{0,k''} \wedge E_{i-1}\right),$$

Here, we have $\mathbf{P}\left(v \in W_{i-1}^{1,k''} \mid E_{i-1}\right) = q_{i-1,1,k''}$ and $\mathbf{P}\left(v \in W_{i-1}^{0,k''} \mid E_{i-1}\right) = q_{i-1,0,k''}$, since, as with the case $d = 1$ and $\ell = 1$, these probabilities do not depend on the set of purple and yellow vertices of u at time $i-1$ by Corollary 3.1.13. On the other hand,

$$\mathbf{P}\left(v \in P_i \mid v \in W_{i-1}^{j'',k''} \wedge E_{i-1}\right) = p_{i,j'',k''},$$

since $v \in P_i$ if and only if it has relevant label i. As a consequence,

$$\mathbf{P}\left(v \in P_i \mid E_{i-1}\right) = \sum_{j''=0}^{1} \sum_{k''=0}^{r-j''-1} p_{i,j'',k''} q_{i-1,j'',k''}.$$

Now, we calculate $\mathbf{P}\left(v \in W_i \mid E_{i-1}\right)$. As before, this can be rewritten as

$$\begin{aligned}\mathbf{P}\left(v \in W_i \mid E_{i-1}\right) = &\sum_{k''=0}^{r-2} \mathbf{P}\left(v \in W_{i-1}^{1,k''} \mid E_{i-1}\right) \mathbf{P}\left(v \in W_i \mid v \in W_{i-1}^{1,k''} \wedge E_{i-1}\right) \\ &+ \sum_{k''=0}^{r-1} \mathbf{P}\left(v \in W_{i-1}^{0,k''} \mid E_{i-1}\right) \mathbf{P}\left(v \in W_i \mid v \in W_{i-1}^{0,k''} \wedge E_{i-1}\right)\end{aligned} \tag{3.4.1}$$

The term $\mathbf{P}\left(v \in W_i \mid v \in W_{i-1}^{1,k''} \wedge E_{i-1}\right)$ is equal to

$$\mathbf{P}\left(v \notin P_i \mid v \in W_{i-1}^{1,k''} \wedge E_{i-1}\right) \mathbf{P}\left(v \in W_i \mid v \notin P_i \wedge v \in W_{i-1}^{1,k''} \wedge E_{i-1}\right). \tag{3.4.2}$$

Clearly, $\mathbf{P}\left(v \notin P_i \mid v \in W_{i-1}^{1,k''} \wedge E_{i-1}\right) = 1 - p_{i,1,k''}$. For the other term, note that, because we are conditioning on u and v not having relevant label i, vertex v is in W_i if and only if none of its remaining $r - k'' - 2$ white neighbours have relevant label i. As in the case $d = 1$ and $\ell = 1$, if S denotes the set of white neighbours of v other than u at time $i-1$, we may apply Corollary 3.2.1 to obtain

$$\mathbf{P}\left(v \in W_i \mid v \notin P_i \wedge v \in W_{i-1}^{1,k''} \wedge E_{i-1}\right) = \prod_{w \in S} \mathbf{P}\left(w \notin P_i \mid v \in W_{i-1} \wedge v \notin P_i\right),$$

with

$$\begin{aligned}\mathbf{P}\left(v \in W_i \mid v \in W_{i-1} \wedge v \notin P_i\right) &= \sum_{j^\star=0}^{1} \sum_{k^\star=0}^{r-j^\star-1} \mathbf{P}\left(w \in W_{i-1}^{j^\star,k^\star} \mid v \notin P_i \wedge v \in W_{i-1}\right) \\ &\times \mathbf{P}\left(w \notin P_i \mid w \in W_{i-1}^{j^\star,k^\star} \wedge v \notin P_i \wedge v \in W_{i-1}\right) \\ &= \sum_{j^\star=0}^{1} \sum_{k^\star=0}^{r-j^\star-1} q_{i-1,j^\star,k^\star} \left(1 - p_{i,j^\star,k^\star}\right).\end{aligned} \tag{3.4.3}$$

This implies that

$$\mathbf{P}\left(v \in W_i \mid v \in W_{i-1}^{1,k''} \wedge E_{i-1}\right) = (1 - p_{i,1,k',\prime}) \left(\sum_{j^\star=0}^{1} \sum_{k^\star=0}^{r-j^\star-1} q_{i-1,j^\star,k^\star} \left(1 - p_{i,j^\star,k^\star}\right)\right)^{r-k''-2}.$$

The term $\mathbf{P}\left(v \in W_i \mid v \in W_{i-1}^{0,k''} \wedge E_{i-1}\right)$ in equation (3.4.1) can be treated similarly, and we obtain

$$\mathbf{P}\left(v \in W_i \mid v \in W_{i-1}^{0,k''} \wedge E_{i-1}\right) = \left(\sum_{j^\star=0}^{1} \sum_{k^\star=0}^{r-j^\star-1} q_{i-1,j^\star,k^\star} \left(1 - p_{i,j^\star,k^\star}\right)\right)^{r-k''-1}$$
$$+ \binom{r-k''-1}{1} \left(\sum_{j^\star=0}^{1} \sum_{k^\star=0}^{r-j^\star-1} q_{i-1,j^\star,k^\star} p_{i,j^\star,k^\star}\right) \left(\sum_{j^\star=0}^{1} \sum_{k^\star=0}^{r-j^\star-1} q_{i-1,j^\star,k^\star} \left(1 - p_{i,j^\star,k^\star}\right)\right)^{r-k''-2}.$$

Observe that the first term in this sum refers to the case in which none of the white neighbours of v is in P_i, while the second term refers to the case when exactly one of them is in P_i.

Substituting these values into equation (3.2.5), we obtain the formula

$$\mathbf{P}\left(v \in W_i \mid E_{i-1}\right) = \sum_{k''=0}^{r-2} \left(1 - p_{i,1,k''}\right) q_{i-1,1,k''} \left(\sum_{j^\star=0}^{1} \sum_{k^\star=0}^{r-j^\star-1} q_{i-1,j^\star,k^\star} \left(1 - p_{i,j^\star,k^\star}\right)\right)^{r-k''-2}$$
$$+ \sum_{k''=0}^{r-1} \left(1 - p_{i,0,k''}\right) q_{i-1,0,k''} \left(\sum_{j^\star=0}^{1} \sum_{k^\star=0}^{r-j^\star-1} q_{i-1,j^\star,k^\star} \left(1 - p_{i,j^\star,k^\star}\right)\right)^{r-k''-1}$$
$$+ \binom{r-k''-1}{1} \left(1 - p_{i,0,k''}\right) q_{i-1,0,k''} \left(\sum_{j^\star=0}^{1} \sum_{k^\star=0}^{r-j^\star-1} q_{i-1,j^\star,k^\star} p_{i,j^\star,k^\star}\right)$$
$$\times \left(\sum_{j^\star=0}^{1} \sum_{k^\star=0}^{r-j^\star-1} q_{i-1,j^\star,k^\star} \left(1 - p_{i,j^\star,k^\star}\right)\right)^{r-k''-2}.$$

A formula for the term $\mathbf{P}\left(v \in Y_i \mid E_{i-1}\right)$ may be obtained by proceeding as in the calculation of $\mathbf{P}\left(v \in W_i \mid E_{i-1}\right)$, but with $\{v \in Y_i\}$ replacing $\{v \in W_i\}$, up until equation (3.4.3). In this equation, we see that $\mathbf{P}\left(v \in Y_i \mid v \notin P_i \wedge v \in W_{i-1}\right) = 1 - \mathbf{P}\left(v \in W_i \mid v \notin P_i \wedge v \in W_{i-1}\right)$, which has already been calculated. This gives

$$\mathbf{P}\left(v_t \in Y_i \mid E\right) = \sum_{k''=0}^{r-2} \left(1 - p_{i,1,k''}\right) q_{i-1,1,k''} \left(1 - \left(\sum_{j^\star,k^\star=0}^{1,r-1} q_{i-1,j^\star,k^\star} \left(1 - p_{i,j^\star,k^\star}\right)\right)^{r-k''-2}\right)$$

By Lemma 3.2.2,

$$\begin{aligned} w_{i,j,k} &= \mathbf{P}(u \in W_i^{j,k}) \\ &= \sum_{j'=0}^{j} \sum_{k'=0}^{k} w_{i-1,j',k'}(1-p_{i,j',k'}) \frac{(r-j'-k')!}{(r-j-k)!(j-j')!(k-k')!} \mathbf{P}\left(v_t \in W_i \mid E\right)^{r-j-k} \\ &\qquad \times \ \mathbf{P}\left(v_t \in P_i \mid E\right)^{j-j'} \mathbf{P}\left(v_t \in Y_i \mid E\right)^{k-k'} \end{aligned}$$

By substituting the expressions obtained and by using the formula for $q_{i-1,j,k}$ given in Corollary 3.1.14, we have

$$\begin{aligned} w_{i,j,k} =& w_{i-1,j,k} - p_{i,j,k} w_{i-1,j,k} \\ & - \frac{(r-j-k) w_{i-1,j,k}}{s_{i-1}} \sum_{j''=0}^{1} \sum_{k''=0}^{r-1-j''} p_{i,j'',k''}(r-j''-k'') w_{i-1,j'',k''} \\ & + \delta_{j=1} \frac{(r-j-k+1) w_{i-1,0,k}}{s_{i-1}} \left(\sum_{j''=0}^{1} \sum_{k''=0}^{r-1-j''} (r-j''-k'') w_{i-1,j'',k''} p_{i,j'',k''} \right) \\ & + \frac{\delta_{k\geq 1}(r-j-k+1) w_{i-1,j,k-1} - (r-j-k) w_{i-1,j,k}}{s_{i-1}^2} \\ & \times \lambda_i \left(\sum_{j^\star=0}^{1} \sum_{k^\star=0}^{r-j^\star-1} (r-j^\star-k^\star) w_{i-1,j^\star,k^\star} p_{i,j^\star,k^\star} \right) \\ & + E_{j,k}(\mathbf{p}_i, \mathbf{w}_{i-1}) \end{aligned} \tag{3.4.4}$$

where

$$s_i = \sum_{j^\star=0}^{1} \sum_{k^\star=0}^{r-1-j} (r-j^\star-k^\star) w_{i,j^\star,k^\star}, \ \lambda_i = \left(\sum_{k''=0}^{r-3} (r-k''-1)(r-k''-2) w_{i-1,1,k''} \right),$$

and $E_{j,k}$ is the error function, as defined in equation (3.2.2). Again, the denominators of the rational coefficients of $E_{j,k}$ with respect to $\mathbf{w}_{i-1}$ are powers of s_{i-1}. All the other terms in the above expansion are monomials of degree 1 with respect to $\mathbf{p}_i$, which gives us the function $\mathbf{F} = (F_{j,k})_{j,k\in\mathcal{I}}$ associated with the algorithm.

Chapter 4

The analysis of locally greedy algorithms

In Chapter 3, we have seen that it is possible to obtain recurrence equations for the probabilities $w_{i,j,k} = \mathbf{P}(u \in W_i^{j,k})$ associated with an application of a locally greedy algorithm with fixed parameters to a fixed input graph. Recall that similar results were obtained in Chapter 2 with regard to the greedy algorithm, even though the variables in the recurrence equations in that chapter were the probabilities of a different set of events. In this chapter, we aim to extend the remainder of the analysis in Chapter 2 to estimate the expected performance of a locally greedy algorithm. That is, we would like to approximate the probabilities $w_{i,j,k}$, the solutions to the system of recurrence equations given by equation (3.2.2) and with initial conditions defined in (3.2.1), by the solutions of an associated system of differential equations. The new bounds on the independence number and on the size of a largest induced forest for regular graphs with large girth would then be obtained by solving these associated systems of differential equations for the appropriate choice of parameters, where $d = \ell = 1$ in the case of independent sets, and $d = 1$ and $\ell = 2$ for induced forests.

However, we are faced with the additional difficulty of defining the probabilities p_0 and $p_{i,j,k}$ in a way that optimises the performance of the algorithm, since a considerable degree of latitude seems to be allowed in the assignment of these probabilities. As a consequence, before introducing a framework in which the probabilities are fixed, we shall first describe some properties that we would like our locally greedy algorithm to satisfy. This will be motivated by the analysis of the performance of prioritised algorithms in the probabilistic space of random regular graphs, which has originally been done in [56] (see also [57] for a more detailed account). In the context of random regular graphs, the objective is to determine the expected performance of the algorithm on a large typical input graph, that is, on an r-regular graph on n vertices that is chosen uniformly among all r-regular graphs on n vertices. This simplifies the analysis, since we may avoid complicated configurations by proving that they

occur with negligible probability. On the other hand, the results obtained are asymptotic, that is, it is shown that, with an input graph chosen uniformly among all r-regular graphs on n vertices, the probability that the algorithm performs as argued in the analysis tends to 1 as n tends to infinity. Our objective here is different, since we wish to analyse a locally greedy algorithm for which the input graph is fixed in advance, and the results aimed at refer to all graphs with sufficiently large girth. Nevertheless, we shall see that, even though the analysis done in the case of random regular graphs cannot readily adapted to our case, it is a good source of motivation in the development of a framework that works in large girth graphs.

In Section 4.1, we look at random algorithms for a particular problem, the problem of obtaining a large independent set in an r-regular graph G. We discuss how to introduce prioritisation in such an algorithm, as well as its performance on a typical graph. This discussion is not rigorous, but it sheds light on later definitions.

In Section 4.2, we introduce a framework for defining the probabilities in the input of a locally greedy algorithm which, in some sense, lets us define an algorithm with the properties suggested in Section 4.1, but, at same time, can still be analysed through the method of Chapter 2. As an illustration, we discuss an application of this framework to the analysis of a simple locally greedy algorithm in Section 4.3.

The remaining sections of this chapter are devoted to defining the probabilities so as to improve the performance of a locally greedy algorithm. This will be used in the proof of the new bounds on the independence number and on the size of a largest induced forest for regular graphs with large girth, which are obtained in Chapter 5 and Chapter 6, respectively.

4.1 An example: independent sets in r-regular graphs

As an illustration of the ideas involved in our approach, we discuss the particular case of finding large independent sets in r-regular graphs. We solve this problem by considering a deletion algorithm with parameters $d = \ell = 1$, which produces an independent set of the input graph, as seen in Proposition 3.1.2.

Here, white vertices are never adjacent to purple vertices. So, we may define an equivalent version of this algorithm, which, instead of colouring the chosen vertices purple and their neighbours yellow, deletes both purple and yellow vertices. This new version is called a deletion algorithm for independent sets. Clearly, the vertices with degree $r - k$ in the graph after i steps of the local deletion algorithm correspond to the white vertices with k yellow neighbours in the colouring at time i in an application of Algorithm 3.1.1 with these parameters. This new version of the algorithm may be written as follows.

Algorithm 4.1.1 ***Deletion algorithm for independent sets***
Input: An r-regular graph G, a positive integer N, an initial probability p_0 and vectors of probabilities $\mathbf{p}_i = (p_{i,k} : 0 \leq k \leq r)$, $i = 1, \ldots, N$.

1. *In the first step, add each vertex in G to a set S_0 with probability p_0, at random, independently of all others. Let $G_0 = G \setminus (S_0 \cup N_0)$, where N_0 is the set of vertices in G adjacent with vertices in S_0.*

2. *At each step i, vertices of G_{i-1} are randomly and independently added to a set S_i, where a vertex with degree $r-k$ is added with probability $p_{i,k}$. Let $G_i = G_{i-1} \setminus (S_i \cup N_i)$, where N_i is the set of vertices in G_{i-1} adjacent with vertices in S_i. Repeat this iteratively for N steps.*

3. *Create a set $\bar{S} \subseteq S = \bigcup_{i=0}^{N} S_i$ by deleting every pair of adjacent vertices in S.*

Output: The set $\bar{S}$.

The performance of an instance of this algorithm obviously depends on the definition of the probabilities p_0 and $p_{i,k}$. A very simple way of assigning these probabilities is to set $p_0 = p_{i,k} = p$ for every $i \in \{1, \ldots, N\}$ and every $k \in \{0, \ldots, r\}$, where p is some positive constant. This yields a greedy algorithm for independent sets in r-regular graphs, in the sense that, at every step, vertices that are not in the independent set nor adjacent with a vertex in the independent set are chosen with the same probability. An algorithm with an equivalent definition, but which ignores degrees of vertices in G_{i-1}, given that they are irrelevant for the algorithm, has been previously analysed in [39].

Such a choice of probabilities may have the advantage of simplifying the analysis of our algorithm, but it clearly does not use its full power, which resides precisely in the possibility of defining the probabilities according to the degree of the vertex in G_{i-1}. In principle, one would expect that a vertex of smaller degree should be chosen with larger probability, since the addition of such a vertex to S_i causes the deletion of fewer vertices when G_i is obtained from G_{i-1}. In order to take maximum advantage of this fact, one would ideally choose, at a given step i, only vertices in G_{i-1} with minimum degree. Such an algorithm has been called *degree-greedy*. For simplicity, we discuss the behaviour of a degree-greedy algorithm for which a single vertex is added to the independent set at each step, with the vertex added at step i being randomly chosen amongst all vertices with minimum degree in G_{i-1}. We suppose that the graph to which this algorithm is applied is a large typical graph, or a random graph. We also assume that the number N of steps performed by the algorithm is such that the algorithm only ends when G_N is the empty graph.

Initially, all the vertices in G have degree r, so any vertex of the graph may be chosen, and, once a vertex v is chosen, the graph G_0 is obtained by deleting v and its neighbours from G. Clearly, the vertices at distance two from v in the original graph have degree smaller than r in G_0. In fact, all of them have degree exactly $r-1$ unless v is part of a 4-cycle. ¿From now on, vertices of degree smaller than r will be chosen. Early in the process, it is unlikely that any vertices of degree smaller than $r-1$ will be created, and any such vertices

created are chosen in the next few steps until none remain. We will say that the algorithm is in Phase 1, for which adding a vertex of degree $r-1$ to the independent set is the *basic* operation. As our algorithm evolves, however, the occasional vertices of degree $r-2$ created will become increasingly common, until we reach a point for which, as a vertex of degree $r-2$ is chosen, more vertices of degree $r-2$ tend to be created, that is, the vertices of degree $r-2$ begin to regenerate themselves faster than they are consumed. Our algorithm now enters in Phase 2, whose basic operation is to choose a vertex of degree $r-2$. In general, when the algorithm is in Phase k, the basic operation consists of adding a vertex of degree $r-k$ to the independent set. Again, vertices of smaller degree may be created during this phase, and they are all added to the independent set before another basic operation is performed. There will be a point in which vertices of degree $r-k-1$ regenerate themselves faster than they are consumed, which marks the transition of the algorithm to Phase $k+1$, in which the basic operation is choosing a vertex of degree $r-k-1$. Vertices of degree 0 cannot accumulate, since no vertices of degree 0 are ever created due to the selection a vertex of degree 0. As a consequence, Phase $k-1$, whose basic operation is choosing a vertex of degree 1, is the last possible phase of this algorithm.

Note that, in this discussion, we have assumed that the phases proceed in an orderly fashion, which is not at all obvious. On the one hand, we implicitly supposed that, in Phase k, as vertices of degree k are the basic vertices processed, the vertices of degree $k-1$ begin regenerating themselves fast enough for the next phase to start before the vertices of degree k are exhausted. If this were not the case, we would reach a point in which the minimum degree of the deletion graph created by the algorithm would be larger than $r-k$, and the algorithm would have to reverse to an earlier phase. On the other hand, we assumed that a phase whose basic vertices have degree $r-k$ is followed by a phase in which the basic vertices have degree $r-k-1$, that is, we assumed that the algorithm would not skip the vertices of degree $r-k-1$ and jump to a phase whose basic vertices have smaller degree, say degree $r-k-2$.

By performing numerical calculations, Wormald shows in [57] that, if the degree-greedy algorithm is applied to a random r-regular graph with $r \leq 100$, these phases can be defined precisely, and that this algorithm a.a.s. proceeds in the way described above. In Section 4.4, we shall discuss in more depth how this degree-greedy algorithm can be analysed in the context of random regular graphs, since this will be useful in defining probabilities $p_{i,k}$ that lead to a locally greedy algorithm whose performance approximates the degree-greedy algorithm.

4.2 A framework for analysis

Our objective in this section is to present conditions under which the performance of the locally greedy algorithm can be analysed. We bear in mind that, to prove our main results, we shall fix the parameters of Algorithm 3.1.1 in a way that approximates the degree-greedy behaviour discussed in the previous section.

So, we first discuss, in the particular case of independent sets, the influence of each of the parameters of Algorithm 4.1.1. The initial probability p_0 will act as an initial step, as was the case when a vertex of degree r was initially selected in the degree-greedy algorithm of Section 4.1. To apply Algorithm 4.1.1, we also need to define the number of steps N and the probabilities $p_{i,k}$. Note that, once these values are fixed, the vector $\mathbf{w}_i = (w_{i,k})_{k=0}^{r}$, where $w_{i,k}$ is the probability that a vertex $u \in V(G)$ has degree $r-k$ in G_i, satisfies the system of recurrence equations

$$\begin{aligned} &\mathbf{w}_i = \mathbf{w}_{i-1} + \mathbf{F}(\mathbf{p}_i, \mathbf{w}_{i-1}) + \mathbf{E}(\mathbf{p}_i, \mathbf{w}_{i-1}),\ i = 1, \ldots, N, \\ &\mathbf{w}_0 = \mathbf{w}_0(p_0), \end{aligned}$$

where $\mathbf{F}$ and $\mathbf{E}$ are the functions associated with Algorithm 3.1.1 in the case $d = 1$ and $\ell = 1$, which are given in equation (3.3.4), and $\mathbf{w}_0(p_0)$ denotes the initial conditions defined in (3.2.1). Recall that the error term $\mathbf{E}(\mathbf{p}_i, \mathbf{w}_{i-1})$ is a polynomial in the variables $\mathbf{p}_i$ for which every monomial has total degree at least 2. In particular, for the error to be negligible, we need small probabilities $p_{i,k}$. Moreover, in order to give a degree-greedy character to the algorithm, we should allow for the definition of the vectors of probabilities $\mathbf{p}_i$ and $\mathbf{p}_{i+1}$ to differ considerably in a finite set of steps $i \in \{1, \ldots, N\}$, these being the steps in which the algorithm goes through a phase transition.

We now present an overview of our strategy. For fixed parameters r, d and ℓ, and a given initial probability p_0, the vector of initial conditions $\mathbf{w}_0(p_0)$ is defined as in (3.2.1). Again, we shall restrict to the case $d = 1$, since this simplifies the discussion slightly, in addition to being the relevant case for the theorems to be proved.

Suppose that this choice of d and ℓ determines functions $\mathbf{F} = (F_{j,k})_{(j,k)\in\mathcal{I}}$ and $\mathbf{E} = (E_{j,k})_{(j,k)\in\mathcal{I}}$ associated with the algorithm, as defined in (3.2.2). To apply Algorithm 3.1.1 to an r-regular graph, we still need to define the number of steps N and the probabilities $p_{i,j,k}$, $1 \le i \le N$, $(j,k) \in \mathcal{I}$. For the latter, we will specify bounded nonnegative functions

$$\hat{p}_{j,k} : \mathbb{R} \to \mathbb{R},\ \text{for every } (j,k) \in \mathcal{I},$$

which are fixed in the following discussion. Now, given a sufficiently small constant $\epsilon > 0$, we may define the probabilities as

$$p_{i,j,k} = p_{i,j,k}(\epsilon) = \epsilon \hat{p}_{j,k}(i\epsilon).$$

Having the probabilities defined in this way is convenient, since it allows us to consider arbitrarily small probabilities by letting ϵ go to zero. Moreover, we let these functions $\hat{p}_{j,k}$ have discontinuities in some finite set $\{x_1, \ldots, x_{m-1}\}$, which are related with the steps of the algorithm in which a phase transition occurs.

Let $\epsilon > 0$. We consider the application of Algorithm 3.1.1 with $d = 1$ and parameters r and ℓ. The probabilities are p_0 and $p_{i,j,k} = \epsilon \hat{p}_{j,k}(i\epsilon)$, for $i \in \{1, \ldots, N\}$ and $(j,k) \in \mathcal{I}$. Here, the number of steps N is given by $N = \lfloor x_m/\epsilon \rfloor$, where x_m is a positive real number larger than x_{m-1} that will be defined later and naturally arises as a termination condition for the algorithm. We also let an input graph G have girth at least $4N+4$ to ensure that the independence results in Section 3.1 hold.

Let $w^{\epsilon}_{i,j,k}$ denote the probability $\mathbf{P}(u \in W_i^{j,k})$ when this algorithm is applied with some fixed $\epsilon > 0$. Then, by our previous assumption, the quantities $w^{\epsilon}_{i,j,k}$ satisfy a system of recurrence equations of the form

$$\begin{aligned} &w^{\epsilon}_{i,j,k} = w^{\epsilon}_{i-1,j,k} + F_{j,k}(\epsilon\hat{\mathbf{p}}(i\epsilon), \mathbf{w}^{\epsilon}_{i-1}) + E_{j,k}(\epsilon\hat{\mathbf{p}}(i\epsilon), \mathbf{w}^{\epsilon}_{i-1}),\ i = 1, \ldots, N,\ (j,k) \in \mathcal{I}, \\ &w^{\epsilon}_{0,j,k} = w_{0,j,k}(p_0),\ (j,k) \in \mathcal{I}, \end{aligned} \tag{4.2.1}$$

with the functions $\hat{\mathbf{p}} = (\hat{p}_{j,k})$ being fixed. We shall show that, as ϵ tends to zero, the solutions to this system of recurrence equations are approximated by the solution to the systems of differential equations given as follows, provided that some technical conditions to be specified later hold. For $t \in \{1, \ldots, m\}$,

$$\begin{aligned} &\frac{dy_{j,k}}{dx} = F_{j,k}(\hat{\mathbf{p}}(x), \mathbf{y}),\ (j,k) \in \mathcal{I},\ \text{for } x \text{ in the interval } [x_{t-1}, x_t) \\ &y_{j,k}(x_{t-1}) = \beta_{t-1,j,k},\ (j,k) \in \mathcal{I}, \end{aligned} \tag{4.2.2}$$

where $\beta_{0,j,k} = w_{0,j,k}(p_0)$ and, inductively,

$$\beta_{t,j,k} = \lim_{x \to x_t^-} y_{j,k}(x),\ \text{if } t > 1.$$

These systems of differential equations will be called the *systems of differential equations associated with the algorithm.*

By inductively solving this system of differential equations for $t = 1, \ldots, m$, we derive information about the expected performance of Algorithm 3.1.1, from which a bound on the desired graph property is deduced. Observe that, as ϵ approaches zero, the number of steps N undertaken by the algorithm increases, and so does the girth of the graphs to which the algorithm can be applied without affecting the independence results of Section 3.1.

This framework is now developed in more detail. Again, let $d = 1$, and let $0 < l \leq r$ and $p_0 \in [0,1]$ be fixed. Let $(w_{0,j,k}(p_0))_{(j,k)\in\mathcal{I}}$ be the initial conditions associated with Algorithm 3.1.1 with these parameters, given in equation (3.2.1).

For $\gamma, M > 0$, we define a region $\Omega_{\gamma,M}$ given by

$$\{(x, \mathbf{y}) \in \mathbb{R}^{|\mathcal{I}|+1} : 0 \leq x \leq M,\ y_{0,0} \geq \gamma \text{ and } 0 \leq y_{j,k} \leq M,\ \text{for every } (j,k) \in \mathcal{I}\}, \tag{4.2.3}$$

and we fix γ, M in such a way that the vector of initial conditions $(0, (w_{0,j,k}(p_0))_{(j,k)\in\mathcal{I}})$ lies in $\Omega_{\gamma,M}$.

For a given positive integer m, consider real numbers $x_0 = 0 < x_1 < x_2 < \cdots < x_m \leq M$ and, for $(j,k) \in \mathcal{I}$, let $\hat{p}_{j,k} : [0, x_m) \to \mathbb{R}^+$ be functions bounded by a positive constant C which are piecewise continuous, with discontinuities restricted to the set $\{x_1, \ldots, x_{m-1}\}$. The real numbers in this set are called *the points of phase transition.*

Assume that, for this choice of d and ℓ, there exist functions $\mathbf{F} = (F_{j,k})_{(j,k)\in\mathcal{I}}$ and $\mathbf{E} = (E_{j,k})_{(j,k)\in\mathcal{I}}$ associated with the algorithm, as defined in (3.2.2). Recall that $\mathbf{F}$ and $\mathbf{E}$ are vector functions such that, for every $(j,k) \in \mathcal{I}$, $F_{j,k} = F_{j,k}(\mathbf{p}, \mathbf{w})$ and $E_{j,k} = E_{j,k}(\mathbf{p}, \mathbf{w})$ are polynomials whose variables are coordinates of $\mathbf{p}$ and whose coefficients are rational functions in the coordinates of $\mathbf{w}$. Moreover, $F_{j,k}$ is homogeneous with degree 1 in the variables $\mathbf{p}$, while all the monomials of $E_{j,k}$ have total degree at least 2.

We shall consider the functions $f_{j,k} : \Omega_{\gamma,M} \to \mathbb{R}$ given by

$$f_{j,k}(x, \mathbf{y}) = F_{j,k}(\hat{\mathbf{p}}(x), \mathbf{y}).$$

These functions describe the expected rate of change of the variable indexed by $(j,k) \in \mathcal{I}$ in an application of the algorithm. To emphasise the fact that each phase will be analysed separately, we use the notation

$$f_{j,k}^{(t)}(x, \mathbf{y}) = f_{j,k}(x, \mathbf{y})$$

for every $(j,k) \in \mathcal{I}$, $t \in \{1, \ldots, m\}$ and $x \in [x_{t-1}, x_t)$.

Assumption 4.2.1 *The functions $\mathbf{F} = (F_{j,k})$ and $\mathbf{E} = (E_{j,k})$ satisfy the following properties.*

(P_1) The coefficients of the polynomials $E_{j,k}$, which are rational functions in the variables $\mathbf{w}$, do not have poles in the region $\Omega_{\gamma,M}$.

(P_2) The functions $f_{j,k}^{(t)}$ are Lipschitz in the region

$$\Omega_{\gamma,M} \cap \left([x_{t-1}, x_t) \times \mathbb{R}^{|\mathcal{I}|} \right),$$

for each $t \in \{1, \ldots, m\}$ and $(j,k) \in \mathcal{I}$.

(P_3) There exist functions $\hat{\mathbf{w}}(x) = (\hat{w}_{j,k}(x))_{(j,k)\in\mathcal{I}}$ defined for x in the interval $[0, x_m)$ such that $(x, \hat{\mathbf{w}}(x)) \in \Omega_{\gamma,M}$ for every x and that

$$\begin{aligned} &\frac{d\hat{w}_{j,k}}{dx} = f_{j,k}^{(t)}(x, \hat{\mathbf{w}}) \text{ in the interval } [x_{t-1}, x_t),\ t = 1, \ldots, m \\ &\hat{w}_{j,k}(x_{t-1}) = \beta_{t-1,j,k}, \end{aligned} \tag{4.2.4}$$

where $\beta_{t-1,j,k}$ is equal to the initial condition $w_{0,j,k}(p_0)$, if $t = 1$, and is inductively defined as

$$\lim_{x \to x_{t-1}^-} \hat{w}_{j,k}(x) \text{ if } t > 1.$$

Intuitively, the first condition ensures that the influence of the error term $E_{j,k}(\epsilon\hat{p}(i\epsilon), \mathbf{w}_{i-1}^{\epsilon})$ in equation (4.2.1) is negligible in comparison with the influence of $F_{j,k}(\epsilon\hat{p}(i\epsilon), \mathbf{w}_{i-1}^{\epsilon})$. The second condition guarantees that the functions describing the behaviour of the algorithm are well-behaved in each phase.

We now relate this framework with the expected behaviour of Algorithm 3.1.1, and, in particular, we shall require the properties of the above assumption.

Our results will refer to an application of Algorithm 3.1.1, where $d = 1$ and the parameters r and ℓ are fixed. Suppose that the initial probability p_0 is given, and, for some $\epsilon > 0$, the number of steps is $N = \lfloor x_m/\epsilon \rfloor$ and the probabilities $p_{i,j,k}$ are given by $\epsilon\hat{p}_{j,k}(i\epsilon)$. Also assume that the input graph G is r-regular with girth larger than $4N + 4$ so that the independence results of Section 3.1 hold. Since the functions $\hat{p}_{j,k}$ are bounded by a constant C, the additional condition $\epsilon < 1/C$ ensures that the probabilities $p_{i,j,k}$ are always between 0 and 1.

Recall that, by definition of the function $\mathbf{F}$ associated with the algorithm, we have that

$$w_{i,j,k}^{\epsilon} = \mathbf{P}\left(u \in W_i^{j,k}\right)$$

satisfies the recurrence equation

$$w_{i,j,k}^{\epsilon} = w_{i-1,j,k}^{\epsilon} + F_{j,k}(\epsilon\hat{p}(i\epsilon), \mathbf{w}_{i-1}^{\epsilon}) + E_{j,k}(\epsilon\hat{p}(i\epsilon), \mathbf{w}_{i-1}^{\epsilon}). \tag{4.2.5}$$

Because $F_{j,k}$ is homogeneous with degree 1 in the variables $\mathbf{p}$, we have

$$F_{j,k}(\epsilon\hat{p}(i\epsilon), \mathbf{w}_{i-1}^{\epsilon}) = \epsilon F_{j,k}(\hat{p}(i\epsilon), \mathbf{w}_{i-1}^{\epsilon}) = \epsilon f_{j,k}(i\epsilon, \mathbf{w}_{i-1}^{\epsilon}).$$

Moreover, recall that $E_{j,k}(\mathbf{p}, \mathbf{w})$ is a polynomial in the variables $\mathbf{p}$ whose monomials have degree at least two and whose coefficients are rational functions in the variables $\mathbf{w}$ with no poles if $(x, \mathbf{w}) \in \Omega_{\gamma,M}$. So, there exists a constant K_1 for which

$$|E_{j,k}(\epsilon\hat{p}(i\epsilon), \mathbf{w}_{i-1}^{\epsilon})| < K_1\epsilon^2$$

whenever $(\epsilon\hat{p}(i\epsilon), \mathbf{w}_{i-1}^{\epsilon}) \in \Omega_{\gamma,M}$.

The next result shows that, as $\epsilon > 0$ goes to zero, the solutions of the system of differential equations (4.2.4) approximate the solution to the system of recurrence equations (4.2.5). To ensure that the analysis of these systems occurs within the region $\Omega_{\gamma,M}$, we define the time N_f as the last step for which $(i\epsilon, \mathbf{w}_i^{\epsilon})$ is inside $\Omega_{\gamma,M}$, namely

$$N_f = N_f(\epsilon) = \min\{i - 1 \,:\, i \leq N \text{ and } (i\epsilon, \mathbf{w}_i^{\epsilon}) \notin \Omega_{\gamma,M}\}. \tag{4.2.6}$$

If the latter set is empty, we use the convention that $N_f = N$. We observe that, because the solutions of (4.2.5) are events in a well defined probability space, the only reason for $(i\epsilon, \mathbf{w}_i^{\epsilon})$ to leave $\Omega_{\gamma,M}$ is that $w_{i,0,0}^{\epsilon}$ becomes smaller than γ or that some of $w_{i,j,k}^{\epsilon}$ becomes larger than M. Note that the latter never occurs if M is larger than 1, which will be typically the case in our applications.

Theorem 4.2.2 *For any $\xi > 0$, there exists $\epsilon' > 0$ such that, if $0 < \epsilon < \epsilon'$,*

$$\left|w^{\epsilon}_{i,j,k} - \hat{w}_{j,k}(i\epsilon)\right| < \xi,\ i = 0, 1, \ldots, N_f = N_f(\epsilon).$$

Proof The proof of this fact is basically the proof of convergence of Euler's method for the solution of differential equations. Some additional work will be required when analysing the transition between systems of differential equations at the points $x_1, \ldots, x_{m-1}$.

The following will be a useful tool.

Claim 4.2.3 *Let $a_0, a_1, \ldots$ be a sequence of non-negative real numbers that satisfy $a_{n+1} \leq (1 + A)a_n + B$, $n = 0, 1, \ldots$, where A and B are positive constants. Then*

$$a_n \leq (a_0 + B/A)e^{An}.$$

Proof of the Claim We prove, by induction on n, that

$$a_n \leq (a_0 + B/A)(1 + A)^n - B/A.$$

The base case $n = 0$ is trivial. Now, supposing that this inequality is true for some value of $n \geq 0$, we have that

$$a_{n+1} \leq (1+A)a_n+B \leq (1+A)[(a_0+B/A)(1+A)^n-B/A]+B = (a_0+B/A)(1+A)^{n+1}-B/A,$$

concluding the induction. From this, because $1 + x \leq e^x$ for all values of x and because a_0 is non-negative and A, B are positive, we deduce that

$$a_n \leq (a_0 + B/A)(1 + A)^n - B/A \leq (a_0 + B/A)e^{An}. \quad \square$$

Let $\xi > 0$. We have to show the existence of $\epsilon' > 0$ with the property that, for $0 < \epsilon < \epsilon'$,

$$|w^{\epsilon}_{i,j,k} - \hat{w}_{j,k}(i\epsilon)| < \xi,\ i = 0, 1, \ldots, N_f.$$

Before proceeding with the proof, we fix two constants. Since each function $f^{(t)}_{j,k}$, $(j, k) \in \mathcal{I}$ and $1 \leq t \leq m$, is Lipschitz in the respective region $\Omega_{\gamma,M} \cap ([x_{t-1}, x_t) \times \mathbb{R}^{|\mathcal{I}|})$, we may choose the constant K as an upper bound on the Lipschitz constants of all the functions $f^{(t)}_{j,k}$, $(j, k) \in \mathcal{I}$ and $1 \leq t \leq m$, where the Lipschitz constants are taken with respect to the L_1-norm. Moreover, given that each function $f^{(t)}_{j,k}$ is Lipschitz in the bounded region under consideration, it must be bounded in this region, so we may fix L as an upper bound on all $|f^{(t)}_{j,k}(x, \mathbf{y})|$, where $(x, \mathbf{y}) \in \Omega_{\gamma,M} \cap ([x_{t-1}, x_t) \times \mathbb{R}^{|\mathcal{I}|})$.

Let $\epsilon > 0$ be fixed and define the error term

$$e_i = \sum_{(j,k)\in\mathcal{I}} |\hat{w}_{j,k}(i\epsilon) - w^{\epsilon}_{i,j,k}|,$$

for every $i \in \{0, \ldots, N_f\}$. Since the initial conditions are the same in both cases, we have $e_0 = 0$.

We now find an upper bound for e_{i+1} in terms of e_i. Two cases have to be considered. The first case occurs when, for some $t \in \{1, 2, \ldots, m\}$, we have $x_{t-1} \leq i\epsilon \leq (i+1)\epsilon \leq x_t$, that is, the step from i to $i+1$ is fully contained in the interval $[x_{t-1}, x_t]$. As a consequence, the system of differential equations is determined by $f^{(t)}$ only. We have

$$\begin{aligned} |\hat{w}_{j,k}((i+1)\epsilon) - w^{\epsilon}_{i+1,j,k}| \leq |\hat{w}_{j,k}(i\epsilon) - w^{\epsilon}_{i,j,k}| + |-w^{\epsilon}_{i+1,j,k} + w^{\epsilon}_{i,j,k} + \epsilon f^{(t)}_{j,k}(i\epsilon, \mathbf{w}^{\epsilon}_i)| \\ + |\hat{w}_{j,k}((i+1)\epsilon) - \hat{w}_{j,k}(i\epsilon) - \epsilon f^{(t)}_{j,k}(i\epsilon, \hat{\mathbf{w}}(i\epsilon))| + \epsilon |f^{(t)}_{j,k}(i\epsilon, \hat{\mathbf{w}}(i\epsilon)) - f^{(t)}_{j,k}(i\epsilon, \mathbf{w}^{\epsilon}_i)|. \end{aligned} \tag{4.2.7}$$

We shall bound the terms on the right-hand side. We first look at

$$|-w^{\epsilon}_{i+1,j,k} + w^{\epsilon}_{i,j,k} + \epsilon f^{(t)}_{j,k}(i\epsilon, \mathbf{w}^{\epsilon}_i)|.$$

By definition,

$$w^{\epsilon}_{i+1,j,k} - w^{\epsilon}_{i,j,k} - \epsilon f^{(t)}_{j,k}(i\epsilon, \mathbf{w}^{\epsilon}_i) = E_{j,k}(\epsilon \hat{p}(i\epsilon), \mathbf{w}^{\epsilon}_i).$$

Now, our choice of N_f ensures that $(i\epsilon, \mathbf{w}^{\epsilon}_i) \in \Omega_{\gamma,M}$, so

$$|E_{j,k}(\epsilon \hat{p}(i\epsilon), \mathbf{w}^{\epsilon}_i)| < K_1 \epsilon^2$$

for some given constant $K_1 > 0$. Hence,

$$\left| -w^{\epsilon}_{i+1,j,k} + w^{\epsilon}_{i,j,k} + \epsilon f^{(t)}_{j,k}(i\epsilon, \mathbf{w}^{\epsilon}_i) \right| \leq K_1 \epsilon^2.$$

The second term to be bounded is

$$\left| \hat{w}_{j,k}((i+1)\epsilon) - \hat{w}_{j,k}(i\epsilon) - \epsilon f^{(t)}_{j,k}(i\epsilon, \hat{\mathbf{w}}(i\epsilon)) \right|.$$

Since $\hat{w}_{j,k}(x)$ is differentiable for x in the interval $[x_{t-1}, x_t)$ with derivative $f^{(t)}_{j,k}(x, \hat{\mathbf{w}}(x))$, we have

$$\epsilon \min_{x \in [i\epsilon, (i+1)\epsilon)} f^{(t)}_{j,k}(x, \hat{\mathbf{w}}(x)) \leq \hat{w}_{j,k}((i+1)\epsilon) - \hat{w}_{j,k}(i\epsilon) \leq \epsilon \max_{x \in [i\epsilon, (i+1)\epsilon)} f^{(t)}_{j,k}(x, \hat{\mathbf{w}}(x)). \tag{4.2.8}$$

Observe that, because $\hat{w}(x_t) = \lim_{x \to x_t^-} \hat{w}(x)$, this conclusion also holds in the case when $(i+1)\epsilon = x_t$. But $f^{(t)}_{j,k}$ is Lipschitz with Lipschitz constant at most K over the region

$$\Omega_{\gamma,M} \cap \left([x_{t-1}, x_t) \times \mathbb{R}^{|I|} \right),$$

with the Lipschitz constant being given with respect to the L_1-norm. As a consequence, we have

$$|f^{(t)}_{j,k}(x, \hat{\mathbf{w}}(x)) - f^{(t)}_{j,k}(i\epsilon, \hat{\mathbf{w}}(i\epsilon))| < K(\epsilon + e_i),$$

and, in particular,

$$\begin{aligned}|\hat{w}_{j,k}((i+1)\epsilon) - \hat{w}_{j,k}(i\epsilon) - \epsilon f_{j,k}^{(t)}(i\epsilon, \hat{\mathbf{w}}(i\epsilon))| \\ \leq \epsilon \max_{x\in[i\epsilon,(i+1)\epsilon)} |f_{j,k}^{(t)}(x, \hat{\mathbf{w}}(x)) - f_{j,k}^{(t)}(i\epsilon, \hat{\mathbf{w}}(i\epsilon))| \leq \epsilon K(\epsilon + e_i).\end{aligned} \tag{4.2.9}$$

Finally, we see that the term

$$\epsilon \left| f_{j,k}^{(t)}(i\epsilon, \hat{\mathbf{w}}(i\epsilon)) - f_{j,k}^{(t)}(i\epsilon, \mathbf{w}_i^\epsilon) \right|$$

can be easily bounded because, by our choice of N_f, the Lipschitz condition may also be applied. We obtain

$$\epsilon \left| f_{j,k}^{(t)}(i\epsilon, \hat{\mathbf{w}}(i\epsilon)) - f_{j,k}^{(t)}(i\epsilon, \mathbf{w}_i^\epsilon) \right| < \epsilon K e_i.$$

Equation (4.2.7) now becomes

$$\begin{aligned}\left|\hat{w}_{j,k}((i+1)\epsilon) - w_{i+1,j,k}^\epsilon\right| &\leq \left|\hat{w}_{j,k}(i\epsilon) - w_{i,j,k}^\epsilon\right| + \epsilon^2 K_1 + \epsilon K(\epsilon + e_i) + \epsilon K e_i \\ &\leq \left|\hat{w}_{j,k}(i\epsilon) - w_{i,j,k}^\epsilon\right| + 2K\epsilon e_i + (K + K_1)\epsilon^2.\end{aligned}$$

Summing this over all pairs $(j,k) \in \mathcal{I}$, we obtain

$$e_{i+1} = \sum_{(j,k)\in\mathcal{I}} \left|\hat{w}_{j,k}(i\epsilon) - w_{i,j,k}^\epsilon\right| \leq e_i + 2K|\mathcal{I}|\epsilon e_i + (K + K_1)|\mathcal{I}|\epsilon^2. \tag{4.2.10}$$

Note that, by definition, $|\mathcal{I}| = \ell(2r - \ell + 3)/2$.

By applying Claim 4.2.3 to e_i with

$$A = 2K|\mathcal{I}|\epsilon,\ B = (K + K_1)|\mathcal{I}|\epsilon^2 \text{ and } n = \lfloor x_1/\epsilon \rfloor,$$

we conclude that the error accumulated up to time $\lfloor x_1/\epsilon \rfloor$ is at most

$$\frac{\epsilon(K + K_1)}{2K} e^{2K|I|x_1}. \tag{4.2.11}$$

We now look at the bound for e_{i+1} in terms of e_i in the transition between two systems of differential equations. In this case, $i\epsilon \leq x_t$ but $(i+1)\epsilon > x_t$. We write

$$\begin{aligned}|\hat{w}_{j,k}((i+1)\epsilon) - w_{i+1,j,k}^\epsilon| \leq |\hat{w}_{j,k}(i\epsilon) - w_{i,j,k}^\epsilon| + |-w_{i+1,j,k}^\epsilon + w_{i,j,k}^\epsilon + \epsilon f_{j,k}^{(t)}(i\epsilon, \mathbf{w}_i^\epsilon)| \\ + |\hat{w}_{j,k}(x_t) - \hat{w}_{j,k}(i\epsilon) - (x_t - i\epsilon) f_{j,k}^{(t)}(i\epsilon, \hat{\mathbf{w}}(i\epsilon))| \\ + |\hat{w}_{j,k}((i+1)\epsilon) - \hat{w}_{j,k}(x_t)| + |[(i+1)\epsilon - x_t] f_{j,k}^{(t)}(i\epsilon, \hat{\mathbf{w}}(i\epsilon))| \\ + \epsilon |f_{j,k}^{(t)}(i\epsilon, \mathbf{w}_i^\epsilon) - f_{j,k}^{(t)}(i\epsilon, \hat{\mathbf{w}}(i\epsilon))|.\end{aligned} \tag{4.2.12}$$

We use the arguments of the first case with respect to the region

$$\Omega_{\gamma,M} \cap ([i\epsilon, x_t) \times \mathbb{R}^{|I|}),$$

which, as argued before, can be extended to the point $x = x_t$ because $\hat{w}_{j,k}(x_t) = \lim_{x \to x_t^-} \hat{w}_{j,k}(x)$. This establishes that

$$|w^{\epsilon}_{i+1,j,k} - w^{\epsilon}_{i,j,k} - \epsilon f^{(t)}_{j,k}(i\epsilon, \mathbf{w}_i)| \leq K_1 \epsilon^2$$

and that

$$|\hat{w}_{j,k}(x_t) - \hat{w}_{j,k}(i\epsilon) - (x_t - i\epsilon) f^{(t)}_{j,k}(i\epsilon, \hat{\mathbf{w}}(i\epsilon))| < \epsilon K(\epsilon + e_i).$$

In addition, we have

$$\epsilon |f^{(t)}_{j,k}(i\epsilon, \mathbf{w}^{\epsilon}_i) - f^{(t)}_{j,k}(i\epsilon, \hat{\mathbf{w}}(i\epsilon))| < \epsilon K(\epsilon + e_i).$$

To bound the other terms in the right-hand side of equation (4.2.12), we use the upper bound L on the value of $|f^{(t)}_{j,k}(x, \mathbf{w})|$ in the region $\Omega_{\gamma,M} \cap ([x_{t-1}, x_t) \times \mathbb{R}^{|I|})$. Recall that this bound is a consequence of $f^{(t)}_{j,k}$ being Lipschitz in this bounded region.

This leads to,

$$\left| [(i+1)\epsilon - x_t] f^{(t)}_{j,k}(i\epsilon, \hat{\mathbf{w}}(i\epsilon)) \right| \leq \epsilon L.$$

Finally, if we use equation (4.2.8) for phase $t+1$ instead of phase t, we obtain

$$\begin{aligned} |\hat{w}_{j,k}((i+1)\epsilon) - \hat{w}_{j,k}(x_t)| &\leq \epsilon \max_{x \in [x_t, (i+1)\epsilon)} \left| f^{(t+1)}_{j,k}(x, \hat{\mathbf{w}}(x)) \right| \\ &\leq \epsilon L. \end{aligned}$$

Thus, equation (4.2.12) becomes

$$|\hat{w}_{j,k}((i+1)\epsilon) - w^{\epsilon}_{i+1,j,k}| \leq |\hat{w}_{j,k}(i\epsilon) - w^{\epsilon}_{i,j,k}| + \epsilon^2 K_1 + 2\epsilon K(\epsilon + e_i) + 2\epsilon L,$$

so

$$e_{i+1} = \sum_{(j,k) \in \mathcal{I}} |\hat{w}_{j,k}(i\epsilon) - w^{\epsilon}_{i,j,k}| \leq e_i + 2K|\mathcal{I}|\epsilon e_i + (K_1\epsilon + 2K\epsilon + 2L)|\mathcal{I}|\epsilon. \tag{4.2.13}$$

Adding this to the error obtained in (4.2.11), the error accumulated between time 0 and $\lfloor x_1/\epsilon \rfloor + 1$ is at most

$$\frac{\epsilon(K + K_1)}{2K} e^{2K|I|x_1} (1 + 2K|I|\epsilon) + |\mathcal{I}|(K_1\epsilon + 2K\epsilon + 2L)\epsilon.$$

This can now be used as the initial error for the second system of differential equations, and, inductively, we can bound the error over the entire process, since the number of transitions is finite. The arguments for bounding the error in the remaining intervals $[x_i, x_{i+1}]$, for $1 \leq i \leq m-1$ are exactly the same, and will therefore be omitted. As a consequence, for some $\epsilon' > 0$ sufficiently small, we have that

$$|w^{\epsilon}_{i,j,k} - \hat{w}_{j,k}(i\epsilon)| < \xi, \ i = 0, 1, \ldots, N_f$$

whenever $0 < \epsilon < \epsilon'$, establishing our result. ∎

The above result is now used to obtain bounds on the expected size of the output set of the algorithm. Recall that we are considering an application of Algorithm 3.1.1, where $d = 1$ and the parameters r and ℓ are fixed, and for which the initial probability p_0 is given. For some $\epsilon > 0$, the number of steps is $N = \lfloor x_m/\epsilon \rfloor$ and the probabilities $p_{i,j,k}$ are equal to $\epsilon \hat{p}_{j,k}(i\epsilon)$. We are also assuming that the input graph G is r-regular with girth larger than $4N + 4$ and that $\epsilon < 1/C$, so as to ensure that the probabilities $p_{i,j,k}$ are always between 0 and 1.

In the following, for each $0 < \epsilon < 1/C$, let $P_i(\epsilon)$ and $\bar{P}(\epsilon)$ be the random variables equivalent to the sets P_i and $\bar{P}$ in the instance of the algorithm where the probabilities $p_{i,j,k}$ are equal to $\epsilon \hat{p}_{j,k}(i\epsilon)$ and the number of steps is $N_f = N_f(\epsilon)$.

Theorem 4.2.4 *Given $\delta > 0$, there exists $\epsilon > 0$ such that*

$$\left| \mathbf{E}|\bar{P}(\epsilon)| - n \left(p_0(1-p_0)^r + \int_0^{\epsilon N_f} \sum_{(j,k)\in\mathcal{I}} \hat{p}_{j,k}(x)\hat{w}_{j,k}(x)\ dx \right) \right| \leq \delta n.$$

Proof We have to bound the expected size of the set $\bar{P}(\epsilon)$ returned by Algorithm 3.1.1. By the description of the algorithm, the set $\bar{P}(\epsilon)$ is clearly contained in the set $P_{N_f}(\epsilon) \setminus Q_0$, where Q_0 is the set of all vertices that are in $P_0(\epsilon)$ and which are adjacent to a vertex also in $P_0(\epsilon)$. Thus,

$$\begin{aligned} \mathbf{E}|\bar{P}(\epsilon)| &\leq \mathbf{E}|P_0(\epsilon)| - \mathbf{E}|Q_0| + \sum_{i=1}^{N_f} \mathbf{E}|P_i(\epsilon)| \\ &= p_0 n - p_0(1-(1-p_0)^r)n + \sum_{i=1}^{N_f} \mathbf{E}|P_i(\epsilon)| \\ &= p_0(1-p_0)^r n + \sum_{i=1}^{N_f} \mathbf{E}|P_i(\epsilon)|. \end{aligned} \tag{4.2.14}$$

On the other hand, a vertex that is added to P_i at step i but is not in $\bar{P}$ must be adjacent to another vertex that was added to R_i. Now, given a vertex in the graph, the probability that none of the vertices adjacent it are in R_i is at least $(1 - \max\{p_{i,j,k} : j,k\})^r$, since, to be in R_i, a vertex needs to have label i with some probability $p_{i,j,k}$ corresponding to the values j and k given by the colouring of its neighbours at time $i-1$. This also relies on the fact that the labels are assigned independently. As a consequence,

$$\mathbf{E}|\bar{P}(\epsilon)| \geq p_0(1-p_0)^r n + \sum_{i=1}^{N_f} \mathbf{E}|P_i(\epsilon)| \left(1 - \max\{p_{i,j,k} : j,k\}\right)^r. \tag{4.2.15}$$

Let $\delta > 0$ and let $h(x)$ be given by

$$h(x) = \sum_{(j,k)\in\mathcal{I}} \hat{p}_{j,k}(x)\hat{w}_{j,k}(x).$$

By definition, $h(x)$ is a bounded function with a finite number of discontinuities in the interval $[0, \epsilon N_f) \subseteq [0, x_m)$, so it is integrable over this interval. Let $\epsilon_1 > 0$ such that, for $0 < \epsilon < \epsilon_1$,

$$\left| \int_0^{\epsilon N_f} h(x)\ dx - \sum_{i=0}^{N_f} \epsilon h(i\epsilon) \right| < \delta/3. \tag{4.2.16}$$

By Theorem 4.2.2, there exists $\epsilon_2 > 0$ such that, for $0 < \epsilon < \epsilon_2$ and $w_{i,j,k}^{\epsilon} = \mathbf{P}\left(u \in W_i^{j,k}\right)$,

$$\left| w_{i,j,k}^{\epsilon} - \hat{w}_{j,k}(i\epsilon) \right| < \frac{\delta}{3|\mathcal{I}|Cx_m},\ i = 0, 1, \ldots, N_f.$$

As a consequence,

$$\begin{aligned}
\left| \sum_{i=1}^{N_f} \mathbf{E}|P_i(\epsilon)| - \sum_{i=0}^{N_f} \epsilon h(i\epsilon) \right| &= \left| \sum_{i=0}^{N_f} \sum_{(j,k)\in\mathcal{I}} \epsilon p_{i,j,k} w_{i,j,k}^{\epsilon} - \sum_{i=0}^{N_f} \epsilon h(i\epsilon) \right| \\
&= \left| \sum_{i=0}^{N_f} \sum_{(j,k)\in\mathcal{I}} \epsilon \hat{p}_{j,k}(i\epsilon) w_{i,j,k}^{\epsilon} - \sum_{i=0}^{N_f} \epsilon h(i\epsilon) \right| \\
&\leq C\epsilon \sum_{i=0}^{N_f} \sum_{(j,k)\in\mathcal{I}} |w_{i,j,k}^{\epsilon} - \hat{w}_{j,k}(i\epsilon)| \\
&\leq C\epsilon N_f |\mathcal{I}| \max_{i,j,k} |w_{i,j,k}^{\epsilon} - \hat{w}_{j,k}(i\epsilon)| \\
&\leq \delta/3.
\end{aligned} \tag{4.2.17}$$

We now choose $0 < \epsilon \leq \min\{\epsilon_1, \epsilon_2\}$ with the additional property that $(1 - C\epsilon)^r > 1 - \delta/3$. On the one hand, equation (4.2.14), combined with equations (4.2.16) and (4.2.17), implies that

$$\begin{aligned}
\mathbf{E}|\bar{P}(\epsilon)| &\leq p_0(1-p_0)^r n + \sum_{i=1}^{N_f} \mathbf{E}|P_i(\epsilon)| \\
&\leq np_0(1-p_0)^r + n\left(\sum_{i=1}^{N_f} \epsilon h(i\epsilon) + \delta/3 \right) \\
&\leq np_0(1-p_0)^r + n\left(\int_0^{\epsilon N_f} h(x)\ dx + 2\delta/3 \right) \\
&< n\left(p_0(1-p_0)^r + \int_0^{\epsilon N_f} h(x)\ dx + \delta \right).
\end{aligned} \tag{4.2.18}$$

On the other hand, equation (4.2.15), combined with equations (4.2.16) and (4.2.17), leads

to

$$\begin{aligned}\mathbf{E}|\bar{P}(\epsilon)| &\geq p_0(1-p_0)^r n + \sum_{i=1}^{N_f} \mathbf{E}|P_i(\epsilon)|\,(1-\max\{p_{i,j,k} : j,k\})^r \\ &\geq np_0(1-p_0)^r + n(1-C\epsilon)^r \left(\sum_{i=1}^{N_f} \epsilon h(i\epsilon) - \delta/3\right) \\ &\geq np_0(1-p_0)^r + n\,(1-C\epsilon)^r \left(\int_0^{\epsilon N_f} h(x)\;dx - 2\delta/3\right) \\ &\geq n\left(p_0(1-p_0)^r + \int_0^{\epsilon N_f} h(x)\;dx - \delta\right),\end{aligned} \tag{4.2.19}$$

concluding the proof of the theorem. ∎

In particular, if we define functions $\hat{p}_{j,k}$ and real numbers $0 < x_1 < \cdots < x_m$ satisfying the properties in Assumption 4.2.1, then, as $\epsilon > 0$ goes to zero, the solution to the system of differential equations (4.2.4) approximates the solution to the system of recurrence equations (4.2.5), which describes the performance of Algorithm 3.1.1 applied for $N = \lfloor x_m/\epsilon \rfloor$ with probabilities $p_{i,j,k} = \epsilon\hat{p}_{j,k}(i\epsilon)$. This framework will be used later to analyse the performance of locally greedy algorithms with particular choices of probabilities.

In the next section, we shall illustrate how this idea can be used to obtain the bounds on the size of a maximum independent set in a graph in a simple situation, for which the functions $\hat{p}_{j,k}$ are constants.

4.3 A simple algorithm for independent sets

As mentioned in the previous section, we shall use the framework introduced there to determine the performance of a greedy algorithm for independent sets, which was first analysed in [39]. Here, Algorithm 3.1.1 chooses white vertices independently with the same probability. Note the similarity between this algorithm and Algorithm 2.2.1, the simple algorithm introduced in Chapter 2 in the case of induced forests.

Let $r \geq 3$. If we set $d = 1$ and $\ell = 1$, the output $\bar{P}$ produced by Algorithm 3.1.1 is an independent set. Recall that this is because two adjacent elements can be added to P only if they are added in the same step, in which case they are not added to $\bar{P}$.

Let $p_0 = 0$ and define, for every $k \in \{0, \ldots, r\}$, functions $\hat{p}_k : \mathbb{R} \to [0,1]$ by

$$\hat{p}_k(x) = 1, \text{ for every } x.$$

This obviously satisfies $\hat{p}_k(x) < C$ whenever $C > 1$. Since these functions have no discontinuities, the algorithm will have only one phase.

Consider the functions $\mathbf{F} = (F_k)_{0\leq k\leq r}$ and $\mathbf{E} = (E_k)_{0\leq k\leq r}$ associated with Algorithm 3.1.1 with parameters $d = 1$ and $\ell = 1$, which are given in equation (3.3.4). In order to verify

that Assumption 4.2.1 holds in this case, we would have to define x_1, γ and M in advance. However, we shall first show that properties (P_1) and (P_2) hold for any choice of positive values γ, M and x_1, while (P_3) holds for any $x_1 > 0$ as long as M is sufficiently large and γ is sufficiently small. Using this, we will fix convenient constants γ, M and x_1.

For Property (P_1), we use the fact that the coefficients of $E_k(\mathbf{p}, \mathbf{w})$, seen as polynomials in the indeterminates $\mathbf{p}$, are rational functions in the indeterminates $\mathbf{w}$ whose denominators are powers of

$$s(\mathbf{w}) = \sum_{t=0}^{r-1} (r-t) w_t. \tag{4.3.1}$$

So, if $\mathbf{w}$ is such that $(x, \mathbf{w}) \in \Omega_{\gamma,M}$, for some $x \geq 0$ and $\gamma, M > 0$, then $s(\mathbf{w}) \geq \gamma r > 0$. In particular, these rational functions have no poles if $(x, \mathbf{w}) \in \Omega_{\gamma,M}$, as required.

Now, by equation (3.3.4), we conclude $f_k(x, \mathbf{w}) = F_k(\hat{\mathbf{p}}(x), \mathbf{y})$ is equal to

$$\begin{aligned} f_k(x, \mathbf{w}) = F_k(\hat{\mathbf{p}}(x), \mathbf{w}) &= -(r-k+1) w_k \\ &+ \frac{(r-k+1) w_{k-1} \delta_{k \geq 1} - (r-k) w_k}{s(\mathbf{w})} \sum_{k''=0}^{r-2} (r-k'')(r-k''-1) w_{k''}, \end{aligned}$$

with $s(\mathbf{w})$ as in (4.3.1). Since f_k is a rational function with no poles in the compact set $\Omega_{\gamma,M}$, we conclude that f_k is Lipschitz in $\Omega_{\gamma,M}$ for every $\gamma, M > 0$, and, in particular, is Lipschitz in

$$\Omega_{\gamma,M} \cap \left([0, x_1) \times \mathbb{R}^{|\mathcal{I}|} \right).$$

This verifies Property (P_2).

As for Property (P_3), consider the system of differential equations

$$\begin{aligned} &\frac{dy_k}{dx} = f_k(x, \mathbf{y}(x)), \ k \in \{0, \ldots, r\}, \\ &y_0(0) = 1, \ y_k(0) = 0 \text{ if } k \geq 1. \end{aligned}$$

Note that the values of the initial conditions come from equation (3.2.1) with $d = \ell = 1$ and $p_0 = 0$. It has non-negative solutions

$$\hat{w}_k(x) = \binom{r}{k} w(x) q(x)^{r-k} (1 - q(x))^k, \ k = 0, \ldots, r, \tag{4.3.2}$$

defined over the entire positive real line, where

$$w(x) = \frac{e^{-x}}{\left((r-1) - (r-2)e^{-x}\right)^{r(r-2)}}$$

and

$$q(x) = \frac{e^{-x}}{(r-1) - (r-2)e^{-x}}.$$

Clearly, if $M \geq x_1$ and $\gamma > 0$ is sufficiently small, $(x, \hat{\mathbf{w}}(x)) \in \Omega_{\gamma,M}$ for every $x \in [0, x_1)$. This verifies Property (P_3).

As promised, we shall now fix the constants x_1, M and γ.

Let $\delta' > 0$. Fix $M > 0$ so that

$$\left| \int_0^\infty \frac{e^{-x}}{((r-1)-(r-2)e^{-x})^{r(r-2)}}\, dx - \int_0^M \frac{e^{-x}}{((r-1)-(r-2)e^{-x})^{r(r-2)}}\, dx \right| < \delta'/2,$$

let $x_1 = M$ and let γ such that $\hat{w}_0(x) \geq \gamma > 0$ for $0 \leq x \leq M$. Note that $\hat{w}_0(x)$ is strictly positive, since the only term that could be equal to zero in the product in the right-hand side of equation (4.3.2) is $1 - q(x)$, which does not appear when $k = 0$.

Consider Algorithm 3.1.1 with parameters r and $d = \ell = 1$, and apply Theorem 4.2.4 with $m = 1$, $x_1 = x_m = M$ and $\delta = \delta'/2$. As a consequence, there exists $\epsilon > 0$ such that, if the algorithm is applied for $N_f = N = \lfloor M/\epsilon \rfloor$ steps and with probabilities $p_0 = 0$ and $p_{i,k} = \epsilon$ to an input graph with girth at least $4N + 4$, then, by our choice of M, the expected size of the independent set returned satisfies

$$\begin{aligned}
\mathbf{E}|\bar{P}(\epsilon)| &\geq n \left(\int_0^M \sum_{k=0}^{r} \hat{w}_k(x)\, dx - \delta'/2 \right) \\
&= n \left(\int_0^M e^{-x} / \left((r-1)-(r-2)e^{-x}\right)^{r(r-2)}\, dx - \delta'/2 \right) \\
&\geq n \left(\int_0^\infty e^{-x} / \left((r-1)-(r-2)e^{-x}\right)^{r(r-2)}\, dx - \delta' \right) \\
&\geq n \left(\frac{1}{2}\left(1 - (r-1)^{-2/(r-2)}\right) - \delta' \right).
\end{aligned}$$

By the first moment principle, any r-regular graph G on n vertices with girth at least $g = g(\epsilon)$ has an independent set of size at least

$$n \left(\frac{1}{2}\left(1 - (r-1)^{-2/(r-2)}\right) - \delta' \right),$$

which is precisely the bound given in [39]. In that paper, however, the analysis is simpler, since the variables considered are only $\sum_{k=0}^{r} w_{i,k}$ and $\sum_{k=0}^{r} q_{i,k}$ (denoted there by r_i and w_i, respectively). But this is to be expected, since the power of Algorithm 3.1.1, which lies in the possibility of defining the probability of a vertex being chosen according to its number of purple and yellow neighbours, is not used in this application.

4.4 Independent sets in r-regular graphs revisited

After seeing an application of the discussion in Section 4.2 to a simple example, our objective is now to define the functions $\hat{p}_{j,k}$ in a way that improves the performance of a locally greedy algorithm. In Section 4.1, we have described a prioritised algorithm for independent sets whose performance was previously analysed for a random regular graph. In light of this,

we now revisit this example in order to discuss some of the main ideas used by Wormald in [57] and [58] to analyse prioritised algorithms in the context of random regular graphs. We emphasise that the analysis presented here is not applicable for algorithms in which the input graph is fixed in advance, which happens to be the case for locally greedy algorithms. Moreover, the presentation of the results in this section is not rigorous, and the proofs of our later theorems do not depend on them. Nonetheless, this section will shed light on some of the definitions in later proofs.

In Section 4.1, we introduced an instance of Algorithm 3.1.1 that produces independent sets in r-regular graphs. As a reminder, this algorithm is repeated below.

Algorithm 4.1.1
Input: An r-regular graph G, a positive integer N, an initial probability p_0 and vectors of probabilities $\mathbf{p}_i = (p_{i,k} : 0 \leq k \leq r)$, $i = 1, \ldots, N$.

1. *In the first step, add each vertex in G to a set S_0 with probability p_0, at random, independently of all others. Let $G_0 = G \setminus (S_0 \cup N_0)$, where N_0 is the set of vertices in G adjacent with vertices in S_0.*

2. *At each step i, vertices of G_{i-1} are randomly and independently added to a set S_i, where a vertex with degree $r-k$ is added with probability $p_{i,k}$. Let $G_i = G_{i-1} \setminus (S_i \cup N_i)$, where N_i is the set of vertices in G_{i-1} adjacent with vertices in S_i. Repeat this iteratively for N steps.*

3. *Create a set $\bar{S} \subseteq S = \bigcup_{i=0}^{N} S_i$ by deleting any pair of adjacent vertices in S.*

Output: The set $\bar{S}$.

When this algorithm was first presented in Section 4.1, we observed that, since it adds a vertex to the independent set at step i with a probability depending on the degree of this vertex in G_{i-1}, one would expect its performance to be best if the vertices added to the independent set had minimum degree.

This motivated the discussion of a simplified algorithm, which we called degree-greedy, in which vertices are added to the independent set one at a time, and the vertex added at step i is chosen randomly amongst the vertices of minimum degree of G_{i-1}. In Section 4.1, we suggested that an application of this algorithm to a large typical graph would go through several phases, where, in Phase k, the basic operation consists of adding a vertex of degree $r-k$ to the independent set. Recall that vertices of smaller degree may be created while the algorithm is in this phase, which, by the definition of the algorithm, are all added to the independent set before another basic operation is performed. We now discuss some of the main ideas in the analysis an application of this algorithm to a random regular graph, which is rigorously presented in [57].

Let $Z_k(i)$ be the random variable denoting the number of vertices of degree $r-k$ in G_i in an application of the degree-greedy algorithm. Our choice of notation reflects the fact that, when the algorithm is in Phase k, the basic operation consists of processing a vertex of degree $r-k$.

Suppose that, after running this algorithm for $i-1$ steps, we obtain a graph G_{i-1} with minimum degree $\delta(G_{i-1}) = l$. Let v be the vertex of degree l in G_{i-1} selected for the independent set in the next step, and let $e_1, \ldots, e_l$ be the edges incident with v in G_{i-1}.

We shall make two independence assumptions in our next discussion. In some sense, these assumptions hold for random regular graphs, but we do not justify this here. In fact, the analysis of an application of this algorithm to a random regular graph involves generating the graph while the algorithm is being applied, so all the events are conditional on the part of the graph that has already been generated. Moreover, the events are not quite independent, and part of the analysis involves arguing that the error due to treating them as if they were independent is small. For a rigorous discussion, the reader is referred to [57].

(A_1) Let E_s denote the event that the other end of edge e_s has degree k_s, where $s \in \{1, \ldots, l\}$ and $k_s \in \{1, \ldots, r\}$. Then the events E_s are mutually independent.

(A_2) The probability of the vertex adjacent to v through e_s having degree k_s is equal to the proportion of the edges in G_{i-1} with an endpoint in a vertex of degree k_s.

Note that an analog of each of these types of independence has been verified in the case of Algorithm 3.1.1. The first is an analog of conditional independence of branches, given in Corollary 3.1.12, while the second is an analog of Corollary 3.1.13.

Using this independence, when v is added to the independent set, the probability that the other endpoint of a given edge incident with v has degree $r-j$ is

$$\frac{(r-j)Z_j(i-1)}{S(i-1)},$$

where

$$S(i-1) = \sum_{s=0}^{r-1}(r-s)Z_s(i-1).$$

As a consequence, the expected number of vertices of degree $r-j$ among the neighbours of v is

$$\frac{(r-l)(r-j)Z_j(i-1)}{S(i-1)}. \tag{4.4.1}$$

All these vertices are deleted from G_{i-1} when G_i is created, so they contribute to a decrease in Z_j.

Now, by the same argument, each time such a vertex w of degree $r-j$ is found to be adjacent to v, the expected number of vertices of degree $r-k$ (other than v) adjacent to w is

$$\frac{(r-j-1)(r-k)Z_k(i-1)}{S(i-1)},$$

hence the total number of vertices of degree $4-k$ affected due to being at distance two from v in G_{i-1} is

$$\sum_{j=0}^{r-2} \frac{(r-l)(r-j)Z_j(i-1)}{S(i-1)} \times \frac{(r-j-1)(r-k)Z_k(i-1)}{S(i-1)}.$$

If none of the vertices affected are adjacent to two neighbours of v, their degree changes from $r-k$ in G_{i-1} to $r-k-1$ in G_i. However, it could happen that one of the vertices affected, call it w, is adjacent to two neighbours of v, which would be the case if v and w were in a common cycle of length 4. Nevertheless, this can be ignored in this analysis. Note that this contributes to both a decrease in Z_k and an increase in Z_{k+1}. Therefore, when a vertex of degree l is chosen by the greedy algorithm, the expected change in the value of Z_k conditional on G_{i-1} is given by

$$\begin{aligned}\phi_k^{(l)}(Z_0(i-1),\ldots,Z_r(i-1)) &= -\delta_{k,l} - \frac{(r-l)(r-k)Z_k(i-1)}{S(i-1)} \\ &+ \sum_{j=0}^{r-2} \frac{(r-j-1)(r-j)[(r-k+1)\delta_{k\geq 1}Z_{k-1}(i-1) - (r-k)Z_k(i-1)]Z_j(i-1)}{S(i-1)^2}.\end{aligned} \tag{4.4.2}$$

In other words,

$$\mathbf{E}\left(Z_k(i) - Z_k(i-1) \middle| \, G_{i-1} \wedge (\delta(G_{i-1}) = l)\right) = \phi_k^{(l)}(Z_0(i-1),\ldots,Z_r(i-1)). \tag{4.4.3}$$

Let $i \in \{1,\ldots,N\}$ and let G_{i-1} be the graph produced by the algorithm after step $i-1$. Recall that, when the degree-greedy algorithm is in Phase k, the basic operation consists of processing vertices of degree $r-k$. However, vertices with many different degrees, all smaller than or equal to $r-k$, may be processed in Phase k, since vertices of smaller degrees can be created as a vertex is deleted from the graph. As a consequence, Wormald chose to determine the expected number of steps between two points in which the algorithm processes a vertex with the basic degree of the phase, and he used this to find the expected change of the value of each random variable Z_k between any two of these steps.

Once the expected changes in the variables are determined, the analysis of the algorithm for a random regular graph is concluded with the proof that, in an application of the algorithm, all the variables a.a.s. are sharply concentrated around their expected values, which is achieved through a general-purpose theorem (see [57], Theorem 5.1). This method has also been used to analyse other algorithms in random regular graphs and is now known as the *Differential Equation Method*.

It is important to mention that the analysis of the expected changes in these variables is rather involved due to the priority constraints. A different approach was given by the same author in a later work [58]. In this work, instead of adhering to the priorities, he considered an algorithm for which a probability vector $\mathbf{p} = \mathbf{p}(n,x) = (p_1,\ldots,p_{r-1})$ is prescribed for each G_i. The probability vector depends only on the number n of vertices in the original graph and on a parameter x, which is related with the current step i of the algorithm.

This new algorithm, at step i, first decides the degree of the vertex to be added to the independent set, with the degree being $r-j$ with some probability $p_{i,j}$ given by $\mathbf{p}$. The vertex is then randomly chosen among the vertices with the selected degree. The new algorithm is *deprioritised*, since the vertices are no longer prioritised according to degree, they just have different probabilities of being chosen. Despite some inherent complications, such as the need to ensure that G_{i-1} contains a vertex of degree j if the algorithm calls for one, the analysis becomes easier. The main reason for this is that the expected change in Z_k can now be determined one step at a time, being given by

$$\mathbf{E}(Z_k(i) - Z_k(i-1)|\ G_{i-1}) = \sum_{l=1}^{r-1} p_{i,l}\phi_k^{(l)}(Z_0(i-1), \ldots, Z_r(i-1)), \tag{4.4.4}$$

with $\phi_j^{(l)}(Z_0(i-1), \ldots, Z_r(i-1))$ given in (4.4.2).

For this deprioritised version to approximate the original prioritised algorithm, Wormald defines the probabilities in terms of the system of differential equations obtained in the prioritised case. However, his proof is technically independent of the analysis of the prioritised algorithm. As with the prioritised case, the fact that, in an application of the algorithm, all the variables a.a.s. are sharply concentrated around their expected values comes from a general-purpose theorem, namely Theorem 2 in [58].

We now suggest an alternative way to define probabilities for a deprioritised version of the algorithm. Here, we derive them directly from some simple properties that a deprioritised degree-greedy algorithm should satisfy. This approach has been previously used in the context of the bisection width of random d-regular graphs by Díaz, Serna and Wormald [20], and in the context of colouring of random regular graphs by Shi and Wormald [51].

Firstly, it is clear that

$$\sum_{l=1}^{r-1} p_{i,l} = 1, \quad \text{for every } i,$$

since, at every step, the algorithm has to choose the degree of the vertex to process. Also, if the algorithm is in phase t, all the vertices processed should have degree at most $r-t$, hence

$$p_{i,l} = 0 \text{ if the algorithm is in phase } t \text{ and } l < t.$$

Finally, during Phase t, the vertices of degree smaller than $r-t$ that are created should not be accumulated. In quantitative terms, this means that, if $k > t$, we must have

$$\mathbf{E}(Z_k(i) - Z_k(i-1)|\ G_{i-1}) = \sum_{l=1}^{r-1} p_{i,l}\phi_k^{(l)}(Z_0(i-1), \ldots, Z_r(i-1)) = 0.$$

So, if the algorithm is in phase t, the probabilities $(p_{i,t}, p_{i,t+1}, \ldots, p_{i,r-1})$ should satisfy the

system of equations

$$\begin{aligned} &\sum_{l=t}^{r-1} p_{i,l} = 1 \\ &\sum_{l=t}^{r-1} p_{i,l}\phi_k^{(l)} = 0, \text{ if } t < k \leq r-1, \end{aligned} \tag{4.4.5}$$

This is a linear system of $r-t$ equations on $r-t$ variables, which has a unique solution whenever the matrix

$$\begin{bmatrix} 1 & 1 & \dots & 1 \\ \phi_{t+1}^{(t+1)} & \phi_{t+1}^{(t+2)} & \dots & \phi_{t+1}^{(r-1)} \\ \phi_{t+2}^{(t+1)} & \phi_{t+2}^{(t+2)} & \dots & \phi_{t+2}^{(r-1)} \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ \phi_{r-1}^{(t+1)} & \phi_{r-1}^{(t+2)} & \dots & \phi_{r-1}^{(r-1)} \end{bmatrix} \tag{4.4.6}$$

has a nonzero determinant. Note that the solutions to this system are functions, since they are given in terms of the functions $\phi_k^{(t)}$.

Also, we remark that the probabilities defined here differ slightly from the probabilities defined in [58]. We illustrate this difference in terms of the greedy algorithm described above. Suppose that the algorithm is in Phase k and that $r-k > 3$. We have seen that at each step, the algorithm could process a vertex with degree $r-k$, a vertex with degree $r-k-1$ or even a vertex with smaller degree. It turns out that processing a vertex of the latter type is a very rare event, and the analysis in [58] was done as if none of these vertices were processed. In other words, the probabilities in [58] were given by the solution to the linear system derived from (4.4.5) by assuming that $p_{i,k} = 0$ for $k \geq t+2$ and by removing the equations

$$\sum_{l=t}^{r-1} p_{i,l}\phi_k^{(l)} = 0$$

for $k \geq t+2$.

The fact that we are using a different assignment of probabilities is somewhat more convenient for our later analysis. It also has the advantage of not requiring any previous knowledge about the process, and therefore can be easily generalised to other applications. However, both choices of probabilities lead to the same results, as will be discussed later.

4.5 A class of systems of differential equations

We shall later use the intuitive ideas of the previous section to devise a locally greedy algorithm whose performance approximates that of a prioritised algorithm in a random regular graph. The analysis of such an algorithm will depend on the solution of a special class of systems of differential equations, which will be presented in this section.

Note that, as in the case of the algorithm for independent sets in Section 4.3, we shall carry out this analysis using the framework of Section 4.2, and one of the ingredients of this framework, specified by Property (P_3) in Assumption 4.2.1, is the existence of solutions $\hat{\mathbf{w}}(x) = (\hat{w}_{j,k}(x))_{(j,k)\in\mathcal{I}}$ to a system of differential equations

$$\begin{aligned} &\frac{dy_{j,k}}{dx} = f_{j,k}^{(t)}(x, \mathbf{y}),\ (j,k) \in \mathcal{I},\ \text{in the interval } [x_{t-1}, x_t),\ t = 1, \ldots, m \\ &y_{j,k}(x_{t-1}) = \beta_{t-1,j,k},\ (j,k) \in \mathcal{I}, \end{aligned} \tag{4.5.1}$$

where the initial conditions $\beta_{t-1,j,k}$ are equal to the initial conditions $w_{0,j,k}(p_0)$, if $t = 1$, and are inductively defined as

$$\lim_{x \to x_{t-1}^-} \hat{w}_{j,k}(x)$$

if $t > 1$. The constants $0 = x_0 < x_1 < \cdots < x_m$ are the points of phase transition, while the functions $f_{j,k}^{(t)}$ represent the rate of change in the variable associated with $(j,k) \in \mathcal{I}$ when the basic operation realised by the algorithm has type t.

We shall now look at a particular class of systems of differential equations of this type. Let $\mathcal{I}$ be a finite set of indices containing an element i_0 and let $\mathcal{J} = \{j_1, \ldots, j_m\}$ be a subset of $\mathcal{I}$ not containing i_0, where $m \geq 1$. As the name suggests, the set $\mathcal{I}$ represents the set indexing the variables in the system of differential equations, while, in terms of the prioritised algorithm discussed in Section 4.4, the set $\mathcal{J}$ indexes the variables associated with each operation in the algorithm. For simplicity, the operation indexed by j_t will be called *the operation of type* t.

We also define a region

$$\Omega_{\gamma,\mu,M} = \{(x, \mathbf{y}) \in \mathbb{R}^{1+|\mathcal{I}|} : -\mu \leq x \leq M,\ y_{i_0} \geq \gamma \text{ and } -\mu \leq y_i < M,\ \text{for every } i \in \mathcal{I}\},$$

where γ, μ and M are positive constants. The difference between this region and the region $\Omega_{\gamma,M}$, defined in (4.2.3) with i_0 replacing the element $(0,0)$, is that some of the variables are now allowed to be negative. The reason for this choice is purely technical: it will be easier to analyse the system of differential conditions inside a region if the initial conditions lie in its interior. In the cases we shall analyse, the initial conditions always lie in the interior of $\Omega_{\gamma,\mu,M}$, but may lie on the boundary of $\Omega_{\gamma,M}$. Nevertheless, our analysis will also show that the solutions remain within $\Omega_{\gamma,M}$.

We consider, for each $i \in \mathcal{I}$ and each $t \in \{1, \ldots, m\}$, a function $\phi_i^{(t)} : \mathbb{R}^{1+|\mathcal{I}|} \to \mathbb{R}$. In terms of the discussion in Section 4.4, the functions $\phi_i^{(t)}$ play the role of the functions in

(4.4.3), that is, they represent the rate of change of the variable indexed by i if the algorithm realises an operation of type t. We assume that these functions have the property that, for every $\gamma > 0$ and $M > 1$, there is $\mu > 0$ with the property that $\phi_i^{(t)}$ is Lipschitz in $\Omega_{\gamma,\mu,M}$.

For $t \in \{1, 2, \ldots, m\}$, let $(\alpha_l^{(t)}(x, \mathbf{y}))_{l=1}^m$ be the vector denoting the proportion of operations of each type undertaken as the algorithm is in Phase t. Also based on the discussion in Section 4.4, we let $\alpha_k^{(t)} = 0$ if $k < t$, while $\alpha^{(t)}(x, \mathbf{y}) = (\alpha_l^{(t)}(x, \mathbf{y}))_{l=t}^m$ is a solution to the linear system given in (4.4.5), which is precisely

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ \phi_{j_{t+1}}^{(t)}(x, \mathbf{y}) & \phi_{j_{t+1}}^{(t+1)}(x, \mathbf{y}) & \cdots & \phi_{j_{t+1}}^{(m)}(x, \mathbf{y}) \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \phi_{j_m}^{(t)}(x, \mathbf{y}) & \phi_{j_m}^{(t+1)}(x, \mathbf{y}) & \cdots & \phi_{j_m}^{(m)}(x, \mathbf{y}) \end{bmatrix} \begin{bmatrix} \alpha_t^{(t)} \\ \alpha_{t+1}^{(t)} \\ \cdot \\ \cdot \\ \alpha_m^{(t)} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix}. \tag{4.5.2}$$

So, if $M^{(t)}(x, \mathbf{y})$ is the matrix in the above equation and $M_k^{(t)}$ is its submatrix obtained by removing the first row and the k^{th} column, Cramer's Rule establishes that

$$\alpha_k^{(t)}(x, \mathbf{y}) = \frac{(-1)^{k-t} \det M_{k-t+1}^{(t)}(x, \mathbf{y})}{\det M^{(t)}(x, \mathbf{y})}, \ k = t, \ldots, m, \tag{4.5.3}$$

provided that

$$\det M^{(t)}(x, \mathbf{y}) \neq 0. \tag{4.5.4}$$

Note that, in the case $t = m$, the vector $\alpha^{(t)}$ has a unique coordinate $\alpha_m^{(m)} = 1$.

For the present, we proceed as if condition (4.5.4) held for all points $(x, \mathbf{y}(x))$ of interest, which will be ensured subsequently. Then the vector $\alpha^{(1)}$ is well defined and we may consider the system of differential equations

$$\begin{aligned} &\frac{dy_i}{dx} = \sum_{l=1}^{m} \alpha_l^{(1)}(x, \mathbf{y}(x)) \phi_i^{(l)}(x, \mathbf{y}(x)), \ i \in \mathcal{I} \\ &y_{i_0}(0) = 1, \ y_i(0) = 0 \text{ if } i \neq i_0. \end{aligned} \tag{4.5.5}$$

Observe that the initial conditions given here correspond to the initial conditions in equation (4.5.1) in the case when $p_0 = 0$. Moreover, since the function $\phi_i^{(t)}$ represents the rate of change of the variable indexed by i as the algorithm realises operation t and the vector $\alpha^{(1)}$ defines the proportion of vertices of each type processed by the algorithm during Phase 1, we call (4.5.5) the *system associated with Phase 1.*

We would like to ensure that this system of differential equations has a solutions in an interval, which we call the *interval of definition of Phase 1.* In particular, following our suggestion that the components of $\alpha^{(1)}$ represent proportions, we need $0 \leq \alpha_k^{(1)}(x, \mathbf{y}(x)) \leq 1$ for every k, and for every x in the interval of definition of Phase 1. Also, because we wish to use the solution to this system in the context of Section 4.2, we look for solutions as in

Property (P_3) of Assumption 4.2.1. In particular, we also require that, for some positive constants γ and M, the point $(x, \mathbf{y}(x))$ lies in the region $\Omega_{\gamma,M}$ for any x in this interval, with $\Omega_{\gamma,M}$ defined as in (4.2.3).

As a consequence, we impose conditions on the functions $\phi_i^{(t)}$ to ensure the existence of a solution to (4.5.5) satisfying these additional conditions in an interval of definition $[0, x_1]$, where the point $x_1 > 0$, still to be defined, marks the end of Phase 1. Before stating these conditions, we fix constants $\gamma > 0$, $M > 1$ and $\mu > 0$ with the property that the functions $\phi_i^{(t)}(x, \mathbf{y})$ are Lipschitz in the region $\Omega_{\gamma,\mu,M}$ for every $i \in \mathcal{I}$ and every $t \in \{1, \dots, m\}$. These constants will be used in the definition of our conditions.

Our first assumption is that

$$|\det M^{(1)}(0, \mathbf{y}(0))| > \gamma. \tag{4.5.6}$$

This ensures that the condition given in (4.5.4) is fulfilled at $x = 0$, so that, because $\det M^{(1)}(x, \mathbf{y})$ is a continuous function, the vector $\alpha^1(x, \mathbf{y}(x))$ is defined in a neighbourhood of $x = 0$ for which $\mathbf{y}(x)$ is defined. Now, to show that $\mathbf{y}(x)$ is indeed defined in a neighbourhood of $x = 0$, we use Lemma 2.5.1. This result is applied to the system of differential equations (4.5.5), with $\Omega = \Omega_{\gamma,\mu,M}$, to find a solution $\mathbf{y}(x)$ that is uniquely extended arbitrarily close to the boundary of $\Omega_{\gamma,\mu,M}$. Note that, because the initial conditions are in the interior of $\Omega_{\gamma,\mu,M}$, the solution can be extended to x in an interval $[0, x']$ for some $x' > 0$. To ensure that x' may be chosen in a way that neither $\alpha_k^{(1)}$ nor y_k leave the interval $[0, 1]$, we assume that this system also satisfies the properties below. Assumption (a) in this list has already been made in (4.5.6) and is repeated here for future reference.

Assumption 4.5.1 ***(Assumptions for Phase 1)*** *Consider the sets*

$$A_1 = \left\{ k \in \{1, \dots, m\} : \frac{d\alpha_k^{(1)}}{dx} \equiv 0 \right\} \quad \textit{and} \quad Y_1 = \left\{ i \in \mathcal{I} : \frac{dy_i}{dx} \equiv 0 \right\}.$$

The system of differential equations (4.5.5) satisfies

(a) $|\det M^{(1)}(0, \mathbf{y}(0))| > \gamma$.

(b) $\alpha_1^{(1)}(0) > \gamma$, *and for each* $k \in \{2, \dots, m\} \setminus A_1$, *at least one of the following holds:*

(i) $\alpha_k^{(1)}(0) > \gamma$,

(ii) $0 \le \alpha_k^{(1)}(0)$, *the set* $A_{1,k} = \left\{ s \ge 1 \ : \ \dfrac{d^s \alpha_k^{(1)}}{dx^s}(0) \neq 0 \right\}$ *is not empty and is such that* $\dfrac{d^\nu \alpha_k^{(1)}}{dx^\nu}(0) > \gamma$, *where* $\nu = \min A_{1,k}$;

(c) for each $i \in \mathcal{I} \setminus Y_1$, *at least one of the following holds*

(i) $y_i(0) > \gamma$,

(ii) $0 \leq y_i(0)$, *the set* $Y_{1,i} = \left\{ s \geq 1 : \dfrac{d^s y_i}{dx^s}(0) \neq 0 \right\}$ *is not empty and is such that* $\dfrac{d^\nu y_i}{dx^\nu}(0) > \gamma$, *where* $\nu = \min Y_{1,i}$;

(d) $\dfrac{dy_{i_0}}{dx}(0) < -\gamma$.

In the above,

$$\frac{d^\nu \alpha_k^{(1)}}{dx^\nu} = \frac{d^\nu \alpha_k^{(1)}(x, \mathbf{y}(x))}{dx^\nu},$$

that is, $\mathbf{y}$ is also seen as a function of x in this derivative. Also note that the conditions involving $\alpha^{(1)}$ and the derivatives of the variables y_i are in fact conditions on the functions $\phi_i^{(k)}$ because of the equations (4.5.3) and (4.5.5).

Before proceeding with our discussion, we justify the above conditions. Recall that our objective is to find a solution to (4.5.5) in an interval $[0, x']$, $x' > 0$, such that neither $\alpha_k^{(1)}$ nor y_k leave the interval $[0, 1]$.

Indeed, condition (b) guarantees that, if $\alpha_k^{(1)}$ is not bounded away from zero at $x = 0$ (i.e., if (i) does not hold), then it either becomes positive in a neighbourhood of the point $x = 0$ (i.e., (ii) holds) or it is equal to 0 at every x for which a solution to (4.5.5) is defined (i.e., $k \in A_1$). Note that, by ensuring that, for every k, $\alpha_k^{(1)}(x, \mathbf{y}(x)) \geq 0$ in a neighbourhood of $x = 0$, the requirement $\alpha_k^{(1)}(x, \mathbf{y}(x)) \leq 1$ is automatically satisfied, since, by the definition of the vector $\alpha^{(1)}$, its coordinates add up to 1. Analogously, condition (c) ensures that the solutions y_i are nonnegative in a neighbourhood of $x = 0$. Finally, condition (d) guarantees that y_{i_0} is not greater than 1 in a neighbourhood of $x = 0$.

We now extend our solution to (4.5.5) to all $x \in [0, x_1]$, where x_1 is defined as the infimum of all $x > 0$ for which at least one of the following termination conditions hold:

$$\begin{aligned}
&1. \quad \text{for some } k \in \{1, \ldots, m\} \setminus A_1,\ \alpha_k^{(1)}(x, \mathbf{y}(x)) = 0;\\
&2. \quad \text{for some } i \in \mathcal{I} \setminus Y_1,\ y_i(x) = 0 \text{ or } y_i(x) = 1;\\
&3. \quad \det M^{(1)}(x, \mathbf{y}(x)) = 0;\\
&4. \quad \text{the solution is outside } \Omega_{\gamma,\mu,M}, \text{ or does not exist.}
\end{aligned} \tag{4.5.7}$$

Since system (4.5.5) is the system associated with Phase 1, the above x_1 is the point where Phase 1 *terminates*. If one of the above conditions holds for some x, we say that the condition is *active* at x.

We show that none of the above conditions are active arbitrarily close to $x = 0$. Clearly, condition 3 is not active due to part (a) in Assumption 4.5.1. Parts (c) and (d), together with the fact that the derivatives of the functions y_i are bounded in a neighbourhood of $x = 0$

(since the functions $\phi_i^{(1)}$ are Lipschitz in $\Omega_{\gamma,\mu,M}$), ensure that we may choose $x' > 0$ such that

$$0 \leq y_i(x) \leq 1, \text{ for } x \in [0, x'] \text{ and } i \in \mathcal{I}.$$

In particular, condition 2 is not active at x arbitrarily close to 0. This also implies that condition 4 is not active in a neighbourhood of $x = 0$.

To show that condition 1 is inactive in a neighbourhood of $x = 0$, we use part (b) of Assumption 4.5.1 in a similar way as part (c) was used to show that condition 2 is inactive. This can be done because the derivative of

$$\alpha_k^{(1)} = \frac{\det M_k^{(1)}(x, \mathbf{y})}{\det M^{(1)}(x, \mathbf{y})}$$

exists and is bounded in a neighbourhood of $x = 0$, since these determinants are polynomials in the Lipschitz functions $\phi_i^{(t)}$ and $\det M^{(1)}(x, \mathbf{y}(x))$ is bounded away from zero in a neighbourhood of $x = 0$. This proves that, under assumptions (a), (b), (c) and (d), we must have $x_1 > 0$.

If any termination condition other than

$$\alpha_k^{(1)}(x, \mathbf{y}(x)) = 0$$

is active at x_1 or if

$$\alpha_k^{(1)}(x_1, \mathbf{y}(x_1)) = 0 \text{ and } \frac{d\alpha_k^{(1)}}{dx}(x_1) \geq 0,$$

we define Phase 1 to be the *final phase*. The intuitive reason for this definition comes from the algorithm, as we now see. Clearly, if conditions 3 or 4 are active at x, then the vector $\alpha^{(1)}$ or the solutions y_i are not well defined at this point. Now, if termination condition 2 is active at x_1 for some $y_i \in \mathcal{I}$, where $i \notin Y_1$, then it must be that $y_i(x_1) = 0$ and, in terms of the algorithm, there are no vertices of type i in the graph. This may prevent us from choosing vertices in our graph with the proportion prescribed by the degree-greedy algorithm. Finally, if termination condition 1 is active for some k, the current phase of the algorithm has to finish, otherwise $\alpha_k^{(1)}$ may become negative, in which case the algorithm would have to process a negative proportion of vertices of type k. As we shall see, the fact that $k = 1$ is necessary for the next phase to start, as is the condition on the derivative.

Now, we suppose that Phase 1 is not the final phase. More generally, we inductively assume that Phase $t-1$ is not the final phase and that there exist constants $0 = x_0 < x_1 < \cdots < x_{t-1}$, and functions $y_i(x)$, $i \in \mathcal{I}$, over the interval $[0, x_{t-1}]$ satisfying, for $s \in \{1, \ldots, t-1\}$,

$$\frac{dy_i}{dx} = \sum_{l=s}^{m} \alpha_l^{(s)}(x, \mathbf{y}) \phi_k^{(l)}(x, \mathbf{y}), \text{ for } x \text{ in the interval } [x_{s-1}, x_s), \ i \in \mathcal{I},$$

$$y_i(x_{s-1}) = \beta_{s-1,i}, \ i \in \mathcal{I},$$

where, $\beta_{0,i_0} = 1$, $\beta_{0,i} = 0$, if $i \neq i_0$, and, for $s > 1$,

$$\beta_{s-1,i} = \lim_{x \to x_{s-1}^-} y_k(x).$$

We consider the system of differential equations

$$\begin{aligned} &\frac{dy_i}{dx} = \sum_{l=t}^{m} \alpha_l^{(t)}(x, \mathbf{y}) \phi_i^{(l)}(x, \mathbf{y}), \ i \in \mathcal{I} \\ &y_i(x_{t-1}) = \beta_{t-1,i}, \ i \in \mathcal{I}. \end{aligned} \tag{4.5.8}$$

We impose conditions to launch Phase t as with Assumption 4.5.1 in the case $t = 1$. Two cases have to be considered: the case when $t < m$ and the case $t = m$. If $t < m$, we suppose that the following conditions hold.

Assumption 4.5.2 ***(Assumptions for Phase t)*** *Consider the sets*

$$A_t = \left\{ k \in \{t, \dots, m\} : \frac{d\alpha_k^{(t)}}{dx} \equiv 0 \right\} \quad \textit{and} \quad Y_t = \left\{ i \in \mathcal{I} : \frac{dy_t}{dx} \equiv 0 \right\}.$$

The system of differential equations (4.5.8) satisfies

(a') $|\det M^{(t)}(x_{t-1}, \mathbf{y}(x_{t-1}))| > \gamma$.

(b') $\alpha_t^{(t)}(x_{t-1}) > \gamma$ *and, for each* $k \in \{t, \dots, m\} \setminus A_t$, *at least one of the following holds:*

(i) $\alpha_k^{(t)}(x_{t-1}) > \gamma$,

(ii) $0 \leq \alpha_k^{(t)}(x_{t-1})$, *the set* $A_{t,k} = \left\{ s \geq 1 \ : \ \frac{d^s \alpha_k^{(t)}}{dx^s}(x_{t-1}) \neq 0 \right\}$ *is not empty and is such that* $\frac{d^\nu \alpha_k^{(t)}}{dx^\nu}(x_{t-1}) > \gamma$, *where* $\nu = \min A_{t,k}$;

(c') *for each* $i \in \mathcal{I} \setminus Y_t$, *at least one of the following holds:*

(i) $y_i(x_{t-1}) > \gamma$,

(ii) $0 \leq y_i(x_{t-1})$, *the set* $Y_{t,i} = \left\{ s \geq 1 \ : \ \frac{d^s y_i}{dx^s}(x_{t-1}) \neq 0 \right\}$ *is not empty and is such that* $\frac{d^\nu y_i}{dx^\nu}(x_{t-1}) > \gamma$, *where* $\nu = Y_{t,i}$;

(d') *for every* $i \in \mathcal{I}$, $y_i(x_{t-1}) < 1 - \gamma$,

Parts (a'), (b') and (c') are analogous to (a), (b) and (c) in Assumption 4.5.1. Part (d'), however, differs from part (d) in the case $t = 1$, but still implies that $y_i(x) \leq 1$ in a neighbourhood of $x = x_{t-1}$ if a solution $\mathbf{y}(x)$ of (4.5.8) is defined in a neighbourhood of this point.

These assumptions again guarantee that a solution to

$$\frac{dy_i}{dx} = \sum_{l=t}^{m} \alpha_l^{(t)}(x, \mathbf{y})\phi_i^{(l)}(x, \mathbf{y}),\ i \in \mathcal{I},$$

where the initial conditions at x_{t-1} are equal to the values of the solution to the system of Phase $t-1$ at x_{t-1}, can be extended up to the point $x_t > x_{t-1}$, which we now define as the infimum of all $x > x_{t-1}$ for which at least one of the following termination conditions holds:

$$\begin{aligned} &1.\ \text{for some } k \in \{t, \dots, m\} \setminus A_t,\ \alpha_k^{(t)}(x, \mathbf{y}(x)) = 0; \\ &2.\ \text{for some } i \in \mathcal{I} \setminus Y_t,\ y_i(x) = 0 \text{ or } y_i(x) = 1; \\ &3.\ \det M^{(t)}(x, \mathbf{y}(x)) = 0; \\ &4.\ \text{the solution is outside } \Omega_{\gamma,\mu,M},\ \text{or does not exist.} \end{aligned} \tag{4.5.9}$$

The point $x = x_t$ is the point where Phase t terminates. If any termination condition other than

$$\alpha_k^{(t)}(x, \mathbf{y}(x)) = 0$$

is active at x_t, or if

$$\alpha_k^{(t)}(x_t, \mathbf{y}(x_t)) = 0 \text{ and } \frac{d\alpha_k^{(t)}}{dx}(x_t) \geq 0,$$

we define Phase t to be the final phase, otherwise the next phase can be defined by induction.

Now, if $t = m$, Phase t is again defined as the final phase of the algorithm. In this case, the conditions to launch Phase t are only parts (c') and (d') in Assumption 4.5.2, since the other two assumptions involve the definition of the vector $\alpha^{(t)}$, which for $t = m$ has a single component, which is equal to the constant 1. This phase continues up to a point x_m, the point where Phase m terminates, which we again define as the infimum of all $x > x_{m-1}$ for which at least one of the following termination conditions hold:

$$\begin{aligned} &1.\ \text{for some } i \in \mathcal{I},\ y_i(x) = 0 \text{ or } y_i(x) = 1; \\ &2.\ \text{the solution is outside } \Omega_{\gamma,\mu,M},\ \text{or does not exist.} \end{aligned} \tag{4.5.10}$$

So, we defined the point in which each phase terminates and the concept of final phase. If x_t is the point where Phase t terminates, and x_t is not the final phase, we say that x_t is a *point of phase transition.*

Let $b \in \{1, \dots, m\}$ denote the index of the final phase. In conclusion, if Assumption 4.5.1 hold at $x = 0$ and Assumption 4.5.2 holds at every point of phase transition, this discussion establishes the existence of a solution to the following systems of differential equations, defined for $t \in \{1, \dots, b\}$,

$$\begin{aligned} &\frac{dy_i}{dx} = \sum_{k=t}^{m} \alpha_k^{(t)}(x, \mathbf{y}(x))\phi_i^{(k)}(x, \mathbf{y}(x)),\ i \in \mathcal{I},\ \text{if } x \in [x_{t-1}, x_t) \\ &y_i(x_{t-1}) = \beta_{t-1,i},\ i \in \mathcal{I}, \end{aligned}$$

where $\beta_{0,i_0} = 1$, $\beta_{0,i} = 0$, if $i \neq i_0$, and, if $t > 0$,

$$\beta_{t,i} = \lim_{x \to x_t^-} y_i(x),$$

with the points $x_1, x_2, \ldots, x_b$ being defined through the termination conditions. Moreover, the solutions are within the region $\Omega_{\gamma,M}$ for every x in the interval $[0, x_b]$, while the term $\alpha_k^{(t)}(x, \mathbf{y}(x))$ lies in $[0, 1]$, for every $t \in \{1, \ldots, b\}$, every $k \in \{t, \ldots, m\}$ and every $x \in [x_{t-1}, x_t)$.

Chapter 5

Independent sets

In this chapter, we shall apply the framework of Section 4.2 to analyse the performance of a locally greedy algorithm for independent sets, as with the example presented in Section 4.3. This will lead to new lower bounds on the cardinality of a maximum independent set in regular graphs with large girth.

We again fix parameters $\ell = d = 1$ in the definition of Algorithm 3.1.1, since, as seen in Proposition 3.1.2, locally greedy algorithms produce independent sets in this case. Now, in order to apply the results of Section 4.2, we need to specify bounded nonnegative functions $\hat{p}_k$, $k = 0, \dots, r$, and constants $0 = x_0 < x_1 < \cdots < x_m$ such that the discontinuities of the functions $\hat{p}_k$ are restricted to the set $\{x_1, \dots, x_{m-1}\}$. Then, given $\epsilon > 0$ sufficiently small and $p_0 \in [0, 1]$, we may consider an application of Algorithm 3.1.1 for $N = \lfloor x_m/\epsilon \rfloor$ steps with initial probability p_0 and probabilities $p_{i,k} = \epsilon \hat{p}_k(i\epsilon)$, for $i \in \{0, \dots, N\}$ and $k \in \{0, \dots, r\}$.

Recall that the definition of the functions $\hat{p}_k$, $k = 0, \dots, r$, and of the constants $0 = x_0 < x_1 < \cdots < x_m$ in the example Section 4.3 was quite simple. Indeed, the functions there all had constant value, so that the ensuing algorithm runs on a single phase. In particular, there are no phase transitions and the only constants are x_0 and x_1. However, since we now aim for an algorithm with a degree-greedy behaviour, the definition of these quantities will be more complicated. In fact, it will be based on the solutions of a system of differential equations. We also observe that, for the functions $\hat{p}_k$, $k = 0, \dots, r$, and the points of phase transition $x_1 < x_2 < \cdots < x_{m-1}$ to be well defined, the solutions to these systems of differential equations are required to satisfy a series of properties, whose establishment in turn depends on a detailed analysis of the differential equations and on the numerical verification of some conditions that they satisfy. To make our argument more concise, we shall write the aspects related with the solutions to the differential equations as a property, which is shown to hold in Section 5.2. We then verify that, for our choice of parameters, there exist constants γ and M for which the properties of Assumption 4.2.1 are satisfied, so that the framework of Section 4.2 may be used to derive a lower bound on the independence number of regular graphs with

large girth through Theorem 4.2.4.

5.1 A general system of differential equations

Our goal is to devise an algorithm whose performance in a graph with large enough girth approximates the performance of the degree-greedy algorithm in a random regular graph. In Sections 4.1 and 4.4, we have seen that this algorithm performs different types of operations, with some operations having higher priority than others. In this chapter, we also define different types of operation in the context of locally greedy algorithms. Indeed, we shall say that the algorithm performs an *operation of type* t when it processes a white vertex with t yellow neighbours, which we call *a vertex of type* t.

Inspired by the description of the main ideas involved in the analysis of the performance of the degree-greedy algorithm in a random regular graph, given in Section 4.4, we first introduce a system of differential equations associated with locally greedy algorithms. To ensure a degree-greedy character, we pretend that, if only operations of type t are applied, the rate of change of the number of vertices of each type is given by equation (4.4.2), as in the case of random regular graphs. We only consider $t \in \{1, \dots, r-1\}$. This is because, on the one hand, white vertices with no yellow neighbours will only be processed in the initialization step. On the other hand, once the algorithm has produced an independent set, all the remaining white vertices with r yellow neighbours can be added to the set without affecting independence, and therefore there is no need of choosing any of them during the algorithm.

So, for $1 \leq t \leq r-1$ and $0 \leq k \leq r$, we consider the functions $\phi_k^{(t)} : \mathbb{R}^{r+2} \to \mathbb{R}$ given by

$$\begin{aligned} \phi_k^{(t)}(x, w_0, \dots, w_r) &= -\delta_{k,t} - (r-k)(r-t)w_k/s(w_0, \dots, w_r) \\ &\quad + (r-t)\frac{(r-k+1)w_{k-1}\delta_{k\geq 1} - (r-k)w_k}{s(w_0, \dots, w_r)^2} \sum_{k''=0}^{r-2} (r-k'')(r-k''-1)w_{k''}, \end{aligned} \tag{5.1.1}$$

where

$$s(w_0, \dots, w_r) = \sum_{k'=0}^{r-1} (r-k')w_{k'}. \tag{5.1.2}$$

Note that the index t refers to the type of operation performed by the algorithm, whereas k corresponds to the different types of vertices.

Now, as suggested in Section 4.4, the proportion $\alpha_k^{(t)}$ of operations of type k while the algorithm is in Phase t, for a fixed $t \in \{1, \dots, r-1\}$, should satisfy the linear system in (4.4.5). If we use our values for $\phi_k^{(t)}$, this linear system is equal to

$$M^{(t)}\, u = v, \tag{5.1.3}$$

where $M^{(t)}$ is the matrix given by

$$\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ (r-t)a_{t+1} & (r-t-1)a_{t+1}-1 & (r-t-2)a_{t+1} & \cdots & a_{t+1} \\ (r-t)a_{t+2} & (r-t-1)a_{t+2} & (r-t-2)a_{t+2}-1 & \cdots & a_{t+2} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ (r-t)a_{r-1} & (r-t-1)a_{r-1} & (r-t-2)a_{r-1} & \cdots & a_{r-1}-1 \end{bmatrix} \tag{5.1.4}$$

and the vectors u and v are given by

$$u = \begin{bmatrix} \alpha_t^{(t)} \\ \alpha_{t+1}^{(t)} \\ \cdot \\ \cdot \\ \alpha_m^{(t)} \end{bmatrix} \quad \text{and} \quad v = \begin{bmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix}. \tag{5.1.5}$$

Here,

$$a_k = a_k(x, \mathbf{w}) = -(r-k)w_k/s + [(r-k+1)w_{k-1} - (r-k)w_k]\lambda/s^2, \ k = 1, \ldots, r-1,$$

with $s = s(x, \mathbf{w})$ defined as in equation (5.1.2) and

$$\lambda = \lambda(x, \mathbf{w}) = \sum_{k=0}^{r-2} (r-k)(r-k-1)w_k.$$

It will be proved in Proposition B.1.6 that the matrix $M^{(t)}$ in this system has determinant

$$\det M^{(t)} = (-1)^{r-t+1} \left(1 + \sum_{l=t+1}^{r-1} (l-t)a_k \right).$$

When this determinant is nonzero, the solution to this linear system is

$$\alpha_k^{(t)}(x, \mathbf{w}) = \begin{cases} \dfrac{1 - \sum_{l=t+1}^{r-1}(r-l)a_l}{1 + \sum_{l=t+1}^{r-1}(l-t)a_l}, & \text{if } k = t, \\ \dfrac{(r-t)a_k}{1 + \sum_{l=t+1}^{r-1}(l-t)a_l}, & \text{if } k > t. \end{cases} \tag{5.1.6}$$

In the case $t = r-1$, the solution is given by

$$\alpha_{r-1}^{(r-1)} = \alpha_{r-1}^{(r-1)}(x, \mathbf{w}) = 1.$$

Once again, we extrapolate the discussion of Section 4.4, and, inspired by equation (4.4.4), we force the rate of change of the variable w_k during Phase t to equal

$$\sum_{l=t}^{r-1} \alpha_l^{(t)}(x, \mathbf{w}) \phi_k^{(l)}(x, \mathbf{w}).$$

As yet, there is no information about the points of phase transition, but, if we fix constants $0 = x_0 < \cdots < x_{r-1}$, a phase $t \in \{1, \ldots, r-1\}$, and initial values $\mu_t = (\mu_{t,k})_{k=0}^{r} \in \mathbb{R}^{r+1}$, we may define the following system of differential equations, called *the basic system of differential equations*;

$$\begin{aligned} &\frac{dw_k}{dx} = F_k^{(t)}(x, \mathbf{w}(x)) \text{ for } x \in [x_{t-1}, x_t],\ k = 0, \ldots, r, \\ &w_k(x_t) = \mu_{t,k}, \end{aligned} \tag{5.1.7}$$

where

$$F_k^{(t)}(x, w_0, \ldots, w_r) = \begin{cases} \sum_{l=1}^{r-1} \alpha_l^{(t)}(x, \mathbf{w}) \phi_k^{(l)}(x, \mathbf{w}), & t \leq r-2, \\ \phi_k^{(r-1)}(x, \mathbf{y}), & t = r-1, \end{cases} \tag{5.1.8}$$

with $\phi_k^{(l)}$ given by equation (5.1.1) and $\alpha_k^{(t)}$ given by equation (5.1.6), if $k \geq t$, and by

$$\alpha_k^{(t)}(x, \mathbf{y}) = 0, \text{ if } k < t.$$

Our objective is to assign probabilities to a locally greedy algorithm in such a way that the performance of the algorithm is described by the solutions to the system of differential equations in (5.1.7). To do this, we define the following property associated with the system of differential equations. The intuitive reasons behind each of the assumptions will be discussed following its statement.

Definition 5.1.1 ***(Property*** $(b, p_0', \mathbf{x})$***)*** *Let $b \in \{1, \ldots, r-1\}$ and $p_0' \in (0, 1]$. With respect to a vector $\mathbf{x}$ of fixed functions*

$$x_1, \ldots, x_b : [0, p_0'] \to [0, 1],$$

where $0 < x_1(p_0) < x_2(p_0) < \cdots < x_b(p_0)$, for every p_0, we say that the basic system of differential equations (5.1.7) satisfies Property $(b, p_0', \mathbf{x})$ if the following holds. For $p_0 \in [0, p_0']$, there are bounded functions $\mathbf{y}^{p_0}(x) = (y_k^{p_0}(x))_{k=0}^{r}$ defined for x in the interval $[0, x_b(p_0))$ such that, for $t \in \{1, \ldots, b\}$,

$$\begin{aligned} &\frac{dy_k^{p_0}}{dx} = F_k^{(t)}(x, \mathbf{y}^{p_0}(x)) \text{ for } x \in [x_{t-1}(p_0), x_t(p_0)),\ k = 0, \ldots, r, \\ &y_k^{p_0}(x_{t-1}(p_0)) = \beta_{t-1,k}^{p_0}. \end{aligned} \tag{5.1.9}$$

Here, the functions $F_k^{(t)}$ are defined as in (5.1.8) and $\beta_{t,k}^{p_0}$ is given as follows. If $t = 1$, it is equal to the initial condition $w_{0,k}(p_0)$ given by equation (3.2.1) with $\ell = d = 1$. If $t > 1$, it is inductively defined as

$$\beta_{t,k}^{p_0} = \lim_{x \to x_{t-1}(p_0)^-} y_k^{p_0}(x).$$

The following additional properties hold.

1. *For x in the interval $[x_{t-1}(p_0), x_t(p_0))$, we have $0 \leq \alpha_k^{(t)}(x, \mathbf{y}^{p_0}(x)) \leq 1$ for every k.*

2. *Given $p_0 \in (0, p_0']$ and $\xi > 0$, there exists $\gamma_\xi^{p_0} > 0$ such that $\gamma_\xi^{p_0} \leq y_k^{p_0}(x) \leq 1 - \gamma_\xi^{p_0}$ for every $x \in [0, x_b(p_0) - \xi]$ and every $k \in \{0, \ldots, r\}$. In particular, $(x, \mathbf{y}(x)) \in \Omega_{\gamma_\xi^{p_0}, M^{p_0}}$ for every $x \in [0, x_b(p_0) - \xi]$, where M^{p_0} is an upper bound on the functions $y_k^{p_0}$.*

3. *For each $t \in \{1, \ldots, b\}$, $\lim_{p_0 \to 0^+} x_t(p_0) = x_t(0)$.*

In future applications of this definition, the number $b \in \{1, \ldots, r-1\}$ will represent the number of phases in the algorithm, while the functions $x_t = x_t(p_0)$ determine the points of transition from Phase t to Phase $t+1$ for different values of the initial probability p_0.

Before proceeding, we briefly discuss the necessity of items 1, 2 and 3 in the previous definition. It is clear that the solutions to the system of differential equations should satisfy item 1, since the vector $\alpha_k^{(t)}$ plays the role of a vector of probabilities. The second item is useful because, in terms of the algorithm, having y_k bounded away from zero implies that the graph has a positive proportion of vertices of type k, which rules out the possibility that the algorithm calls for more vertices of type k in one step than there actually are vertices of this type in the graph. The third item allows us to obtain a bound on the size of the independent set that is essentially the limit, as p_0 goes to zero, of the bounds obtained for each $p_0 > 0$.

The following result shows that, if the system of differential equations (5.1.7) satisfies Property $(b, p_0', \mathbf{x})$ for a fixed assignment of b, p_0' and $\mathbf{x}$, then it can be used to devise and analyse a locally greedy algorithm for independent sets, producing a lower bound on the independence ratio of r-regular graphs with girth sufficiently large. We observe, however, that the quality of the bound provided depends on the choice of b, p_0' and $\mathbf{x}$. We shall see how to define these quantities in Section 5.2.

Theorem 5.1.2 *Let $b \in \{1, \ldots, r-1\}$ and $p_0' \in (0, 1]$, and consider functions*

$$x_1, \ldots, x_b : [0, p_0'] \to [0, 1]$$

for which the basic system of differential equations defined in equation (5.1.7) satisfies Property $(b, p_0', \mathbf{x})$. Then, for $\delta > 0$ and $r \in \mathbb{N}$, there exists $g > 0$ such that every r-regular graph G on n vertices with girth greater than or equal to g satisfies $\alpha(G) \geq (x_b(0) - \delta)n$.

Proof Let $\delta' > 0$. By item 3 in the definition of Property $(b, p_0', \mathbf{x})$, we may choose p_0 with the property that $|x_b(p_0) - x_b(0)| < \delta'/4$. If we fix $\xi = \delta'/4$, we have

$$|x_b(p_0) - \xi - x_b(0)| < \frac{\delta'}{2}.$$

We are now ready to define the functions $\hat{p}_k : \mathbb{R} \to [0, x_b(p_0) - \xi]$ by

$$\hat{p}_k(x) = \begin{cases} \dfrac{\alpha_k^{(t)}(x, \mathbf{y}^{p_0}(x))}{y_k^{p_0}(x)}, & \text{if } t \leq k \text{ and } x \in [x_{t-1}(p_0), x_t(p_0)),\ t = 1, \ldots, \min\{b, r-2\}, \\ \dfrac{1}{y_{r-1}^{p_0}(x)}, & \text{if } k = b = r-1 \text{ and } x \in [x_{r-2}(p_0), x_{r-1}(p_0)], \\ 0, & \text{in all other cases.} \end{cases}$$

Items 1 and 2 in the definition of Property $(b, p_0', \mathbf{x})$, which are satisfied by the solutions $\mathbf{y}^{p_0}(x)$ of (5.1.9), ensure that, if we fix a constant

$$C > \frac{1}{\gamma_\xi^{p_0}},$$

these functions satisfy $0 \leq \hat{p}_k(x) < C$, for every $x \in [0, x_b(p_0) - \xi]$.

With $p_{i,k} = \epsilon \hat{p}_k(\epsilon i)$, the function associated with the algorithm $\mathbf{F} = (F_k)_{k=0}^r$, which was obtained in (3.3.5), induces, for $t \in \{1, \ldots, b\}$, the system of differential equations

$$\begin{aligned} &\frac{dw_k}{dx} = f_k^{(t)}(x, \mathbf{w}(x)) \text{ for } x \text{ in the interval } [x_{t-1}(p_0), x_t(p_0)),\ k = 0, \ldots, r, \\ &w_k(x_{t-1}(p_0)) = \beta_{t-1,k}^{p_0},\ k = 0, \ldots, r, \end{aligned} \tag{5.1.10}$$

where

$$f_k^{(t)}(x, w_0, \ldots, w_r) = F_k(\hat{\mathbf{p}}(x), \mathbf{w}) = \sum_{l=1}^{r-1} \alpha_l^{(t)}(x, \mathbf{y}^{p_0}(x)) \phi_k^{(l)}(x, \mathbf{w}).$$

Here, $\mathbf{y}^{p_0} = \mathbf{y}^{p_0}(x)$ is the solution of (5.1.9). The initial conditions $\beta_{t-1,k}^{p_0}$ are defined as follows. If $t = 1$, they are given by $w_{0,k}(p_0)$, defined in equation (3.2.1) with $\ell = d = 1$. If $t > 1$, they are given inductively by

$$\beta_{t,k}^{p_0} = \lim_{x \to x_{t-1}(p_0)^-} w_k(x).$$

To use our results of Section 4.2, we now verify that the properties in Assumption 4.2.1 are satisfied. To remind the reader of these assumptions, we restate them in terms of the current notation. Recall that these assumptions are given in terms of the functions $\mathbf{F} = (F_k)_{k=0}^r$ and $\mathbf{E} = (E_k)_{k=0}^r$ defined through the system of recurrence equations associated with the algorithm in (3.2.2) and calculated in (3.3.5).

Assumption 4.2.1 *The functions $\mathbf{F} = (F_k)_{k=0}^r$ and $\mathbf{E} = (E_k)_{k=0}^r$ satisfy the following properties.*

(P_1) The coefficients of the polynomials E_k, which are rational functions in the variables $\mathbf{w}$, do not have poles in the region $\Omega_{\gamma_\xi^{p_0}, M^{p_0}}$.

(P_2) The functions $f_k^{(t)}(x, \mathbf{w}) = F_k(\hat{\mathbf{p}}(x), \mathbf{w})$ *are Lipschitz in the region*

$$\Omega_{\gamma_\xi^{p_0}, M^{p_0}} \cap \left([x_{t-1}(p_0), x_t(p_0)) \times \mathbb{R}^{r+1}\right),$$

for each $t \in \{1, \ldots, b\}$ *and* $k \in \{0, \ldots, r\}$.

(P_3) There exist functions $\hat{\mathbf{w}}(x) = (\hat{w}_k(x))_{k=0}^r$ *defined for* x *in the interval* $[0, x_b(p_0))$ *such that* $(x, \hat{\mathbf{w}}(x)) \in \Omega_{\gamma_\xi^{p_0}, M^{p_0}}$ *for every* x *and that, for* $t = 1, \ldots, b$,

$$\begin{aligned} &\frac{d\hat{w}_k}{dx} = f_k^{(t)}(x, \hat{\mathbf{w}}),\ k = 0, \ldots, r,\ \text{in the interval } [x_{t-1}(p_0), x_t(p_0)), \\ &\hat{w}_k(x_{t-1}(p_0)) = \beta_{t-1,k},\ k = 0, \ldots, r \end{aligned} \tag{4.2.11}$$

where $\beta_{t-1,k}$ *is equal to the initial condition* $w_{0,k}(p_0)$, *if* $t = 1$, *and is inductively defined as*

$$\lim_{x \to x_{t-1}(p_0)^-} \hat{w}_k(x),\ \text{if } t > 1.$$

The solution $\hat{\mathbf{w}}(\mathbf{x})$ of Property (P_3) is precisely the solution to (5.1.10), which clearly coincides with the solution to (5.1.9) provided by Property $(b, p_0', \mathbf{x})$.

We verify the remaining properties. Property (P_1) holds as in the case of the greedy algorithm (see Section 4.3), since it depends only on the definition of $E_{j,k}$, and not on the particular assignment of probabilities.

For Property (P_2), note that each $f_k^{(t)}$ is Lipschitz in the region

$$\Omega_{\gamma_\xi^{p_0}, M^{p_0}} \cap ([x_{t-1}(p_0), x_t(p_0)) \times \mathbb{R}^r),$$

as the functions $\phi_k^{(t)}$ are Lipschitz in the region $\Omega_{\gamma_\xi^{p_0}, M^{p_0}}$ and the terms $\alpha_l^{(t)}(x, \mathbf{y}^{p_0}(x))$ are continuous and bounded in $[x_{t-1}(p_0), x_t(p_0))$. Thus, all the properties in Assumption 4.2.1 hold in this case. (To simplify the notation, the interval $[x_{b-1}(p_0), x_b(p_0))$ appears in the above, even if the interval $[x_{b-1}(p_0), x_b(p_0) - \xi]$ is really meant.)

We are now ready to define an instance of the locally greedy algorithm to which the results of Section 4.2 may be applied. Indeed, consider Algorithm 3.1.1 with parameters $d = \ell = 1$ and r, and initial probability p_0. For $0 < \epsilon < 1/C$, suppose that the number of steps is

$$N = \left\lfloor \frac{x_b(p_0) - \xi}{\epsilon} \right\rfloor$$

and the probabilities are $p_{i,k} = \epsilon \hat{p}_k(i\epsilon)$.

We first argue that, if $\epsilon > 0$ is sufficiently small, the final point N_f in the statement of Theorem 4.2.4, which is defined in (4.2.6), is equal to N. Recall that, with initial conditions determined by p_0, the solutions to the system of differential equations (4.2.4) and to the system of recurrence equations (4.2.5) are arbitrarily close to each other for sufficiently small ϵ by Theorem 4.2.2, as long as both solutions lie in the region $\Omega_{\gamma_\xi^{p_0}, M^{p_0}}$. By item 2 in the

definition of Property $(b, p_0', \mathbf{x})$, the point $(x, y^{p_0}(x))$ is in the interior of the compact region $\Omega_{\gamma_\xi^{p_0}, M^{p_0}}$ for every x in the interval $[0, x_b(p_0) - \xi]$, where $\mathbf{y} = \mathbf{y}(x)$ is a solution to (4.2.11). As a consequence, for some $\epsilon_1 > 0$ small enough, the solution $\mathbf{w}_i^\epsilon$ to (4.2.5) cannot leave $\Omega_{\gamma_\xi^{p_0}, M^{p_0}}$ for $i \leq N = \lfloor (x_b(p_0) - \xi)/\epsilon \rfloor$ whenever $0 < \epsilon < \epsilon_1$. This establishes that

$$N_f = N_f(\epsilon) = \left\lfloor \frac{x_b(p_0) - \xi}{\epsilon} \right\rfloor,$$

as desired.

We may now use Theorem 4.2.4 with $M = M^{p_0}$, $\gamma = \gamma_\xi^{p_0}$, $m = b$, $x_m = x_b(p_0) - \xi$ and $\delta = \delta'/2$. This yields $\epsilon > 0$, also satisfying $\epsilon < \epsilon_1$, such that, if Algorithm 3.1.1 is applied with probabilities fixed by this value of ϵ to a graph G with girth larger than $4N + 4$, then the expected size of the independent set returned satisfies

$$\begin{aligned}
\mathbf{E}|\bar{P}(\epsilon)| &\geq n \left(p_0(1-p_0)^r + \int_0^{x_b(p_0)-\xi} h(x, \hat{\mathbf{w}}(x))\, dx - \delta'/2 \right) \\
&= n \left(p_0(1-p_0)^r + \int_0^{x_b(p_0)-\xi} dx - \delta'/2 \right) \\
&= n(p_0(1-p_0)^r + x_b(p_0) - \xi - \delta'/2),
\end{aligned}$$

where $h(x)$ denotes the function $\sum_{k=0}^r \hat{p}_k(x)\hat{w}_k(x)$. The fact that $h(x, \hat{\mathbf{w}}(x)) = 1$ follows immediately from the fact that the systems of differential equation (5.1.9) and (5.1.10) have the same solutions for a fixed p_0.

As a consequence, by the first moment principle, any r-regular graph G with girth at least $g = g(\epsilon)$ has an independent set of size at least

$$n \left(p_0(1-p_0)^r + x_b(p_0) - \xi - \delta'/2 \right) \geq n(x_b(0) - \delta').$$

This concludes the proof of our theorem. ∎

5.2 An appropriate choice of parameters

In the previous section, we have obtained a bound on the independence number of a graph with sufficiently large girth, which depends on the solution to the basic system of differential equations defined in (5.1.7), provided that this system satisfies Property $(b, p_0', \mathbf{x})$, for a positive integer $b \in \{1, \ldots, r-1\}$, a constant $p_0' \in (0, 1]$ and functions

$$x_1, \ldots, x_b : [0, p_0'] \to [0, 1].$$

Our aim in this section is to define these quantities in such a way that Property $(b, p_0', \mathbf{x})$ is satisfied. Moreover, they should be defined so that Theorem 5.1.2 produces "good" bounds,

which, in our case, will be numbers that are equal to the bounds obtained in [56] for independent sets in random regular graphs.

Let $r \geq 3$ be fixed. Recall that Property $(b, p_0', \mathbf{x})$ involves solutions, satisfying a series of properties, to the system of differential equations

$$\begin{aligned} &\frac{dy_k^{p_0}}{dx} = F_k^{(t)}(x, \mathbf{y}^{p_0}(x)) \text{ for } x \in [x_{t-1}(p_0), x_t(p_0)),\ k = 0, \dots, r, \\ &y_k^{p_0}(x_{t-1}(p_0)) = \beta_{t-1,k}^{p_0},\ k = 0, \dots, r. \end{aligned} \tag{5.2.1}$$

Here, the functions $F_k^{(t)}$ are defined as in (5.1.8) and $\beta_{t,k}^{p_0}$ is given as follows. If $t = 1$, it is equal to the initial condition $w_{0,k}(p_0)$ given by equation (3.2.1) with $\ell = d = 1$. If $t > 1$, it is inductively defined as

$$\beta_{t,k}^{p_0} = \lim_{x \to x_{t-1}(p_0)^-} y_k^{p_0}(x).$$

The remainder of this section is devoted to determining parameters b, p_0' and $\mathbf{x} = (x_1, \dots, x_b)$ so as to have x_b as large as possible. This will be done in two parts. The first part consists of analysing the system of differential equations (5.2.1) when the initial conditions are given with $p_0 = 0$. This is a system of differential equations in the class introduced in Section 4.5, and, in particular, we can prove the existence of a solution satisfying a series of properties by verifying that the conditions defined in Section 4.5 are satisfied. For $\gamma > 0$ sufficiently small, we shall obtain $b = b(\gamma) \in \{1, \dots, r-1\}$ and constants $x_1^0 = x_1^0(\gamma), \dots, x_b^0 = x_b^0(\gamma)$, where, in the terminology of Section 4.5, b is the number of phases in the algorithm until a termination condition of the final phase is met, while x_t^0 is the point of termination of Phase t.

In the second part, we derive solutions to (5.2.1) with initial conditions given with respect to a small positive integer p_0 by analysing the sensitivity of this system with respect to its initial conditions. More specifically, we show that, for any value of γ for which the conditions of the first part hold, we may find $p_0' > 0$ and functions $x_1, \dots, x_b : [0, p_0'] \to [0, 1]$ for which the system of differential equations (5.2.1) satisfies Property $(b, p_0', \mathbf{x})$. Moreover $x_t(0) = x_t^0$, for every $t \in \{1, \dots, b\}$.

First part As already mentioned, we shall find a solution to the system of differential equations (5.2.1) in the particular case in which the initial conditions are defined with $p_0 = 0$, which will be done by means of Section 4.5.

We first fix the value of the parameters specified in that section. For $r \geq 3$, the finite set $\mathcal{I}$ indexing the variables in the differential equations is

$$\mathcal{I} = \{(j, k) \in \mathbb{Z}^2 : j = 0,\ 0 \leq k \leq r\},$$

with the element i_0 being equal to $(0, 0)$. The set $\mathcal{J}$ indexing the variables associated with each operation in the algorithm is

$$\mathcal{J} = \{(0, 1), (0, 2), \dots, (0, r-1)\}.$$

Since, in the present application to independent sets, the coordinate j is always equal to 0, it will henceforth be dropped from the notation.

As in the definition of the functions $F_k^{(t)}$ in (5.1.8), we let the functions $\phi_{0,k}^{(t)} = \phi_k^{(t)}$ be given by

$$\begin{aligned}\phi_k^{(t)}(x, w_0, \ldots, w_r) &= -\delta_{k,t} - (r-k)(r-t)w_k/s(w_0, \ldots, w_r) \\ &+ (r-t)\frac{(r-k+1)w_{k-1}\delta_{k\geq 1} - (r-k)w_k}{s(w_0, \ldots, w_r)^2} \sum_{k''=0}^{r-2} (r-k'')(r-k''-1)w_{k''},\end{aligned} \tag{5.2.2}$$

which are precisely the functions $\phi_k^{(t)}$ defined in (5.1.1).

We shall show that, with this choice of $\phi_k^{(t)}$, there exist constants M, γ and μ for which Assumption 4.5.1 is verified at the point $x = 0$ and, whenever Phase $t-1$ is not the final phase for some $t > 1$, Assumption 4.5.2 is verified at $x = x_{t-1}$, the point where Phase $t-1$ terminates. If these assumptions hold, the discussion in Section 4.5 implies that there exists a solution $\mathbf{y}(x)$ to the following systems of differential equations, defined for $t \in \{1, \ldots, b\}$,

$$\begin{aligned}\frac{dy_k}{dx} &= \sum_{l=t}^{r-1} \alpha_l^{(t)}(x, \mathbf{y}(x))\phi_k^{(l)}(x, \mathbf{y}(x)), \ k \in \{0, \ldots, r\}, \text{ if } x \in [x_{t-1}, x_t) \\ y_k(x_{t-1}) &= \beta_{t-1,k}, \ k \in \{0, \ldots, r\},\end{aligned}$$

where $\beta_{0,0} = 1$, $\beta_{0,k} = 0$, if $k > 0$, and, if $t > 0$,

$$\beta_{t,k} = \lim_{x \to x_t^-} y_k(x).$$

The points $x_1, x_2, \ldots, x_b$ are obtained through the termination conditions of each phase, which are defined in Section 4.5, while b is the index of the final phase. Moreover, the solutions lie inside the region $\Omega_{\gamma,M}$ for every x in the interval $[0, x_b]$, and the term $\alpha_k^{(t)}(x, \mathbf{y}(x))$ lies in $[0, 1]$, for every $t \in \{1, \ldots, b\}$, every $k \in \{t, \ldots, r-1\}$ and every $x \in [x_{t-1}, x_t)$.

Let $M > 1$. In the discussion that follows, we shall show that there exists $\gamma > 0$ such that Assumption 4.5.1 and Assumption 4.5.2 hold if

$$0 < \mu < \frac{r\gamma}{\sum_{k=1}^{r-1}(r-k)} = \frac{2\gamma}{r-1}.$$

For this choice of μ, the term $\displaystyle\sum_{k=0}^{r-1}(r-k)y_k$ is always positive in $\Omega_{\gamma,\mu,M}$, so the functions $\phi_k^{(t)}$ are rational functions with no poles in the compact set $\Omega_{\gamma,\mu,M}$, and hence are Lipschitz in this set.

We start with some fixed value $0 < \gamma < 1$, which may be redefined as a smaller value if necessary. This can be done because, whenever the assumptions are satisfied for $\gamma > 0$, they are also satisfied for any γ' in the interval $(0, \gamma)$, so the action of decreasing γ at some point of the argument does not affect the assumptions verified prior to this.

We first verify that Assumption 4.5.1 is satisfied at the point $x = 0$. Recall that we are looking for a solution to

$$\begin{aligned} &\frac{dy_k}{dx} = \sum_{l=1}^{r-1} \alpha_l^{(1)}(x, \mathbf{y}(x))\phi_k^{(l)}(x, \mathbf{y}(x)),\ 0 \le k \le r, \\ &y_0(0) = 1,\ y_k(0) = 0 \text{ if } k > 0, \end{aligned} \tag{5.2.3}$$

where the vector $\alpha^{(1)} = (\alpha_l^{(1)})_{l=1}^{r-1}$ satisfies the linear system given in (4.5.2), which, for our values of $\phi_k^{(t)}$, is equal to

$$M^{(1)}\, u = v, \tag{5.2.4}$$

where

$$u = \begin{bmatrix} \alpha_1^{(1)} \\ \alpha_2^{(1)} \\ \alpha_3^{(1)} \\ \cdot \\ \cdot \\ \alpha_{r-1}^{(1)} \end{bmatrix} \text{ and } v = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix},$$

while $M^{(1)} = M(r-1, a_2, \ldots, a_{r-1})$, where the matrix on the right-hand side is defined by

$$M(n, c_1, \ldots, c_{n-1}) = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ n\,c_1 & (n-1)c_1 - 1 & (n-2)c_1 & \cdots & c_1 \\ n\,c_2 & (n-1)c_2 & (n-2)c_2 - 1 & \cdots & c_2 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ n\,c_{n-1} & (n-1)c_{n-1} & (n-2)c_{n-1} & \cdots & c_{n-1} - 1 \end{bmatrix}. \tag{5.2.5}$$

Here, for $k = 1, \ldots, r-1$

$$a_k = a_k(x, \mathbf{y}) = -(r-k)y_k/s + [(r-k+1)y_{k-1} - (r-k)y_k]\lambda/s^2, \tag{5.2.6}$$

with

$$s = s(x, \mathbf{y}) = \sum_{k=0}^{r-1} (r-k)y_k$$

and

$$\lambda = \lambda(x, \mathbf{y}) = \sum_{k=0}^{r-2} (r-k)(r-k-1)y_k. \tag{5.2.7}$$

We repeat the statement of Assumption 4.5.1 with our current notation to help with the verification of each of its properties.

Assumption 4.5.2 ***(Assumptions for Phase 1)*** *Consider the sets*

$$A_1 = \left\{ k \in \{1, \dots, r-1\} : \frac{d\alpha_k^{(1)}}{dx} \equiv 0 \right\} \quad \textit{and} \quad Y_1 = \left\{ k \in \{0, \dots, r\} : \frac{dy_k}{dx} \equiv 0 \right\}.$$

The system of differential equations (5.2.3) satisfies

(a) $|\det M^{(1)}(0, \mathbf{y}(0))| > \gamma$.

(b) $\alpha_1^{(1)}(0) > \gamma$, *and for each* $k \in \{2, \dots, r-1\} \setminus A_1$, *at least one of the following holds:*

(i) $\alpha_k^{(1)}(0) > \gamma$,

(ii) $0 \le \alpha_k^{(1)}(0)$, *the set* $A_{1,k} = \left\{ s \ge 1 \, : \, \frac{d^s \alpha_k^{(1)}}{dx^s}(0) \ne 0 \right\}$ *is not empty and is such that* $\frac{d^\nu \alpha_k^{(1)}}{dx^\nu}(0) > \gamma$, *where* $\nu = \min A_{1,k}$;

(c) for each $k \in \{0, \dots, r\} \setminus Y_1$, *at least one of the following holds*

(i) $y_k(0) > \gamma$,

(ii) $0 \le y_k(0)$, *the set* $Y_{1,k} = \left\{ s \ge 1 \, : \, \frac{d^s y_k}{dx^s}(0) \ne 0 \right\}$ *is not empty and is such that* $\frac{d^\nu y_k}{dx^\nu}(0) > \gamma$, *where* $\nu = \min Y_{1,k}$;

(d) $\frac{dy_0}{dx}(0) < -\gamma$.

It is a consequence of Proposition B.1.6 that

$$\det M^{(1)}(x, \mathbf{y}) = (-1)^r \left(1 + \sum_{k=2}^{r-1} (k-1) a_k \right).$$

By the definition of a_k in (5.2.6) and by the initial conditions of the system of differential equations (5.2.3), the elements $a_k(0, \mathbf{y}(0))$ are nonnegative and we have

$$|\det M^{(1)}(0, \mathbf{y}(0))| = 1 + \sum_{k=2}^{r-1} (k-1) a_k(0, \mathbf{y}(0)) \ge 1 > \gamma,$$

so that part (a) of Assumption 4.5.1 is satisfied.

Now, whenever $\det M^{(1)}(x, \mathbf{y}) \ne 0$, the solution to the linear system (5.2.4) is

$$\alpha_k^{(1)}(x, \mathbf{y}) = \begin{cases} \dfrac{1 - \sum_{l=2}^{r-1} (r-l) a_l}{1 + \sum_{l=2}^{r-1} (l-1) a_l}, & \text{if } k = 1, \\ \dfrac{(r-1) a_k}{1 + \sum_{l=2}^{r-1} (l-1) a_l}, & \text{if } k > 1. \end{cases} \tag{5.2.8}$$

Recall that $\alpha^{(1)}$ has been defined as the solution to (5.2.4) precisely to ensure that, for $2 \leq k \leq r-1$, the differential equations in (5.2.3) satisfy

$$\frac{dy_k}{dx} = \sum_{l=1}^{r-1} \alpha_l^{(1)} \phi_k^{(l)} \equiv 0.$$

In particular, $\{2, \ldots, r-1\} \subseteq Y_1$, with Y_1 defined as in Assumption 4.5.1. Moreover, the initial conditions of the differential equations (5.2.3) tell us that $y_k(0) = 0$ for $k \in \{2, \ldots, r-1\}$. As a consequence,

$$y_k(x) = 0 \text{ for every } k \in \{2, \ldots, r-1\}, \tag{5.2.9}$$

for every x in the interval of definition of a solution to the system of differential equations in (5.2.3).

By equation (5.2.6), this implies that

$$a_2(x, \mathbf{y}(x)) = (r-1) y_1(x) \lambda(x, \mathbf{y}(x)) / s(x, \mathbf{y}(x))^2 \tag{5.2.10}$$

and that

$$a_k(x, \mathbf{y}(x)) = 0 \text{ for every } x, \text{ if } 3 \leq k \leq r-1. \tag{5.2.11}$$

In particular $\{3, \ldots, r-1\} \subseteq A_1$, where A_1 is defined as in Assumption 4.5.1.

As $\{2, \ldots, r-1\} \subseteq Y_1$, the only cases to be considered in the verification of part (c) of the assumption are $k \in \{0, 1, r\}$. Since $y_0(0) = 1 > \gamma$, item (i) holds for $k = 0$. The case $k = r$ is also simple, since, by equation (5.2.3),

$$\frac{dy_r(x)}{dx} = \sum_{l=1}^{r-1} \alpha_l^{(1)} \phi_{r-1}^{(l)} = \sum_{l=1}^{r-1} \frac{\alpha_l^{(1)} (r-l) \lambda y_{r-1}}{s^2} \equiv 0,$$

given that $y_{r-1} \equiv 0$ by (5.2.9). This shows that $r \in Y_1$. We observe that the reason for which the case $k = r$ has to be treated separately from the case involving values of k with $k > 1$ is a consequence of our option, in Section 5.1, of not processing white vertices with r yellow neighbours during the algorithm, and instead just adding them to the independent set when the algorithm ends.

For the case $k = 1$, note that equation (5.2.6) leads to $\alpha_1^{(1)}(0, \mathbf{y}(0)) = 1$ and to $\alpha_l^{(1)}(0, \mathbf{y}(0)) = 0$, if $l > 1$. By equation (5.2.3), we get

$$\begin{aligned} \frac{dy_1}{dx}(0) &= \sum_{l=1}^{r-1} \alpha_l^{(1)}(0, \mathbf{y}(0)) \phi_1^{(l)}(0, \mathbf{y}(0)) \\ &= \alpha_1^{(1)}(0, \mathbf{y}(0)) \phi_1^{(1)}(0, \mathbf{y}(0)) \\ &= \frac{r(r-1) y_0(0) r(r-1) y_0(0)}{r^2 y_0(0)^2} - 1 = (r-1)^2 - 1 > \gamma, \end{aligned} \tag{5.2.12}$$

concluding the verification of part (c).

We may also derive part (b) of Assumption 4.5.1. If $k > 2$, we know from (5.2.11) that $k \in A_1$ and, if $k = 1$, we have already seen that $\alpha_1^{(1)}(0, \mathbf{y}(0)) = 1 > \gamma$. So, it remains to show that this assumption is satisfied in the case $k = 2$. By equations (5.2.8) and (5.2.10),

$$\alpha_2^{(1)}(x, \mathbf{y}) = \frac{(r-1)a_2}{1 + \sum_{l=2}^{r-1}(l-1)a_l} = \frac{(r-1)^2 y_1 \lambda}{s^2 + s^2 \sum_{l=2}^{r-1}(l-1)a_l}.$$

In particular, $\alpha_2^{(1)}(0, \mathbf{y}(0)) = 0$. Moreover, because equations (5.2.7) and (5.2.12) imply $\lambda(0, \mathbf{y}(0)) \geq r(r-1)\gamma > 0$ and $\dfrac{dy_1}{dx}(0) > 0$, respectively, we have

$$\frac{d\alpha_2^{(1)}}{dx}(0) > 0.$$

So, either item (ii) in (b) holds for the first derivative of $\alpha_2^{(1)}$ with our current choice of γ, or it holds by redefining γ as a smaller positive constant. This concludes the verification of part (b).

To conclude the verification of Assumption 4.5.1, we note that equations (5.2.3) and (5.2.8) lead to

$$\begin{aligned}\frac{dy_0(x)}{dx}(0) &= \sum_{t=1}^{r-1} \alpha_t^{(1)}(0, \mathbf{y}(0))\phi_0^{(t)}(0, \mathbf{y}(0)) = \phi_0^{(1)}(0, \mathbf{y}(0)) \\ &= -\frac{r^2 y_0(0)}{s(0, \mathbf{y}(0))} - \frac{r^2 y_0(0) r(r-1) y_0^{(0)}}{s(0, \mathbf{y}(0))^2} = -r - r(r-1) = -r^2 < -\gamma,\end{aligned}$$

implying part (d), as required. By the discussion in Section 4.5, we may extend our solution to (5.2.3) to all $x \in [0, x_1]$, where x_1 is defined as the infimum of all $x > 0$ for which at least one of the following termination conditions hold:

1. for some $k \in \{1, \ldots, r-1\} \setminus A_1$, $\alpha_k^{(1)}(x, \mathbf{y}(x)) = 0$;
2. for some $k \in \{0, \ldots, r\} \setminus Y_1$, $y_k(x) = 0$ or $y_k(x) = 1$;
3. $\det M^{(1)}(x, \mathbf{y}(x)) = 0$;
4. the solution is outside $\Omega_{\gamma,\mu,M}$, or does not exist.

(5.2.13)

Also, if any termination condition other than

$$\alpha_k^{(1)}(x, \mathbf{y}(x)) = 0$$

is active at x_1, or if

$$\alpha_k^{(1)}(x_1, \mathbf{y}(x_1)) = 0 \text{ and } \frac{d\alpha_k^{(1)}}{dx}(x_1) \geq 0,$$

Phase 1 was defined to be the *final phase*. We observe that the termination condition $y_k(x) = 1$ will never be active in our discussion. This is because the function $\sum_{k=0}^{r} y_k(x)$ is a sum of

nonnegative functions which is equal to 1 at $x = 0$, and the derivative of this function is strictly negative.

If Phase 1 is not the final phase, in particular condition 2 above does not hold, so we may redefine $\gamma > 0$ as a smaller positive number, if necessary, to ensure that

$$\begin{aligned} &\text{(i)} \quad \gamma < y_k(x_1) < 1 - \gamma \text{ if } k \notin Y_1; \\ &\text{(ii)} \quad \frac{d\alpha_1^{(1)}}{dx}(x_1) \leq -\gamma. \end{aligned} \tag{5.2.14}$$

So, we must have

$$y_k(x_1) > \gamma, \text{ if } k \leq 1, \text{ and } y_k(x_1) = 0, \text{ if } k > 1, \tag{5.2.15}$$

where the second fact comes from (5.2.9).

We now show that, provided that Phase $t-1$ is not the final phase, Assumption 4.5.2 holds at the point $x = x_{t-1}$, the point in which Phase $t-1$ terminates. In other words, let $2 \leq t \leq r-1$ and inductively assume that x_{t-1} has been defined and that

$$\alpha_{t-1}^{(t-1)}(x_{t-1}, \mathbf{y}(x_{t-1})) = 0$$

is the only active termination condition of Phase $t-1$ at x_{t-1} and that

$$\frac{d\alpha_{t-1}^{(t-1)}}{dx}(x_{t-1}) < 0.$$

We may also inductively assume that $\gamma > 0$ has been defined so that

$$\begin{aligned} &\text{(i)} \quad \gamma < y_k(x_{t-1}) < 1 - \gamma, \text{ if } k \notin Y_{t-1}; \\ &\text{(ii)} \quad \frac{d\alpha_{t-1}^{(t-1)}}{dx}(x_{t-1}) \leq -\gamma, \end{aligned} \tag{5.2.16}$$

and that, as in equation (5.2.15),

$$y_k(x_{t-1}) > \gamma, \text{ if } k \leq t-1, \text{ and } y_k(x_{t-1}) = 0, \text{ if } k > t-1. \tag{5.2.17}$$

In Phase t, the system of differential equations of interest is

$$\frac{dy_k}{dx} = \sum_{l=t}^{r-1} \alpha_l^{(t)}(x, \mathbf{y}(x)) \phi_k^{(l)}(x, \mathbf{y}(x)), \ 0 \leq k \leq r, \tag{5.2.18}$$

with initial conditions at $x = x_{t-1}$ given by the values of each y_k at the termination of Phase $t-1$. Recall that the terms $\alpha_k^{(t)}$ in the previous equations are the solutions to the linear system (4.5.2), where the matrix $M^{(t)}(x, \mathbf{y})$ is given by $M(r-t, a_{t+1}, \ldots, a_{r-1})$, as defined in (5.2.5), and with a_k defined in (5.2.6).

We remind the reader of the statement of Assumption 4.5.2 under the current notation. We also observe that, although the several of the assumptions of Phase t are the same as the assumptions of Phase 1, we opted to verify them separately to clarify the inductive character of the argument.

Assumption 4.5.3 ***(Assumptions for Phase t)*** *Consider the sets*

$$A_t = \left\{ k \in \{t, \dots, m\} : \frac{d\alpha_k^{(t)}}{dx} \equiv 0 \right\} \quad \text{and} \quad Y_t = \left\{ k \in \{0, \dots, r\} : \frac{dy_k}{dx} \equiv 0 \right\}.$$

The system of differential equations (5.2.18) satisfies

(a') $|\det M^{(t)}(x_{t-1}, \mathbf{y}(x_{t-1}))| > \gamma$.

(b') $\alpha_t^{(t)}(x_{t-1}) > \gamma$ *and, for each* $k \in \{t, \dots, b\} \setminus A_t$, *at least one of the following holds:*

(i) $\alpha_k^{(t)}(x_{t-1}) > \gamma$,

(ii) $0 \le \alpha_k^{(t)}(x_{t-1})$, *the set* $A_{t,k} = \left\{ s \ge 1 \,:\, \frac{d^s \alpha_k^{(t)}}{dx^s}(x_{t-1}) \ne 0 \right\}$ *is not empty and is such that* $\frac{d^\nu \alpha_k^{(t)}}{dx^\nu}(x_{t-1}) > \gamma$, *where* $\nu = \min A_{t,k}$;

(c') *for each* $k \in \{0, \dots, r\} \setminus Y_t$, *at least one of the following holds:*

(i) $y_k(x_{t-1}) > \gamma$,

(ii) $0 \le y_k(x_{t-1})$, *the set* $Y_{t,k} = \left\{ s \ge 1 \,:\, \frac{d^s y_k}{dx^s}(x_{t-1}) \ne 0 \right\}$ *is not empty and is such that* $\frac{d^\nu y_k}{dx^\nu}(x_{t-1}) > \gamma$, *where* $\nu = \min Y_{t,k}$;

(d') *for every* $k \in \{0, \dots, r\}$, $y_k(x_{t-1}) < 1 - \gamma$.

As in the verification of the assumptions at the start of Phase 1, we can use Proposition B.1.6 to obtain

$$|\det M^{(t)}(x_{t-1}, \mathbf{y}(x_{t-1}))| = \left| (-1)^{r-t+1} \left(1 + \sum_{k=t+1}^{r-1} (k-t) a_k(x_{t-1}, \mathbf{y}(x_{t-1})) \right) \right| \ge 1,$$

which establishes part (a') of Assumption 4.5.2, and

$$\alpha_k^{(t)}(x, \mathbf{y}) = \begin{cases} \dfrac{1 - \sum_{l=t+1}^{r-1} (r-l) a_l}{1 + \sum_{l=t+1}^{r-1} (l-t) a_l}, & \text{if } k = t, \\ \dfrac{(r-t) a_k}{1 + \sum_{l=t+1}^{r-1} (l-t) a_l}, & \text{if } k > t. \end{cases} \tag{5.2.19}$$

Again, the definition of $\alpha^{(t)}$ is such that

$$\frac{dy_k}{dx} \equiv 0 \text{ if } t+1 \le k \le r-1,$$

so $\{t+1, \dots, r-1\} \subseteq Y_t$, where Y_t is defined in Assumption 4.5.2. By the inductive hypothesis in (5.2.17), this also implies that

$$y_k(x) = 0 \text{ for } k \in \{t+1, \dots, r-1\}, \tag{5.2.20}$$

for every x in the interval of definition of the solution to (5.2.18).

We now look at part (b') of Assumption 4.5.2. By equations (5.2.6) and (5.2.20), this implies that

$$a_{t+1}(x, \mathbf{y}(x)) = (r-t)y_t(x)\lambda(x, \mathbf{y}(x))/s(x, \mathbf{y}(x))^2 \tag{5.2.21}$$

and that

$$a_k(x, \mathbf{y}(x)) = 0 \text{ for every } x, \text{ if } t+2 \leq k \leq r-1. \tag{5.2.22}$$

In particular $\{t+2, \ldots, r-1\} \subseteq A_t$.

Moreover, equation (5.2.19) leads to

$$\alpha_t^{(t)}(x_{t-1}, \mathbf{y}(x_{t-1})) = 1 > \gamma \text{ and } \alpha_k^{(t)}(x_{t-1}, \mathbf{y}(x_{t-1})) = 0 \text{ if } k > t,$$

which establishes part (b') for $k = t$. So, in the verification of part (b') of this assumption, only the case $k = t+1$ remains. Note that this case will follow if we can show that the nonzero derivative of smallest order of $\alpha_{t+1}^{(t)}$ is positive, which, by the definition of $\alpha_{t+1}^{(t)}$, amounts to showing that the nonzero derivative of smallest order of y_t is positive at $x = x_{t-1}$.

This will be derived from our verification of part (c'). We have already seen that $\{t+1, \ldots, r-1\} \subseteq Y_t$. Also, if $k \leq t-1$, we have $y_k(x_{t-1}) > \gamma$ by the choice of γ in (5.2.16), so item (i) in part (c') of the assumption is satisfied. Now, we look at the case $k = t$. We shall show that

$$\frac{dy_t}{dx}(x_{t-1}) = -\alpha_{t-1}^{(t-1)}(x_{t-1}, \mathbf{y}(x_{t-1})) = 0 \text{ and } \frac{d^2y_t}{dx^2}(x_{t-1}) > 0.$$

This implies that the case $k = t$ satisfies part (c') due to item (ii), possibly by decreasing the value of γ. For the first part, note that, by equation (5.2.18) and by the fact that $\alpha_k^{(t)}(x_{t-1}) = 0$ for $k > t$,

$$\begin{aligned}
\frac{dy_t}{dx}(x_{t-1}) &= \sum_{k=t}^{r-1} \alpha_k^{(t)}(x_{t-1}, \mathbf{y}(x_{t-1}))\phi_t^{(k)}(x_{t-1}, \mathbf{y}(x_{t-1})) = \phi_t^{(t)}(x_{t-1}, \mathbf{y}(x_{t-1})) \\
&= \frac{(r-t+1)(r-t)y_{t-1}(x_{t-1})\lambda(x_{t-1}, \mathbf{y}(x_{t-1}))}{s(x_{t-1}, \mathbf{y}(x_{t-1}))^2} - 1 \\
&= -\alpha_{t-1}^{(t-1)}(x_{t-1}, \mathbf{y}(x_{t-1})) = 0.
\end{aligned}$$

The last equality comes from our assumption on the active termination condition of Phase $t-1$. For the second part, we have

$$\frac{d^2y_t}{dx^2} = \sum_{k=t}^{r-1} \frac{d\alpha_k^{(t)}}{dx}\phi_t^{(k)} + \alpha_k^{(t)}\frac{d\phi_t^{(k)}}{dx}.$$

Since, at $x = x_{k-1}$, we have already seen that $\alpha_k^{(t)} = 0$ for $k > t$ and that $\dfrac{d\alpha_k^{(t)}}{dx} = 0$ for

$k > t+1$, this becomes

$$\frac{d^2y_t}{dx^2}(x_{t-1}) = \frac{d\alpha_t^{(t)}}{dx}(x_{t-1})\phi_t^{(t)}(x_{t-1},\mathbf{y}(x_{t-1})) + \frac{d\alpha_{t+1}^{(t)}}{dx}(x_{t-1})\phi_t^{(t+1)}(x_{t-1},\mathbf{y}(x_{t-1})) + \alpha_t^{(t)}(x_{t-1},\mathbf{y}(x_{t-1}))\frac{d\phi_t^{(t)}}{dx}(x_{t-1},\mathbf{y}(x_{t-1})).$$

On the one hand,

$$\phi_t^{(t)}(x_{t-1},\mathbf{y}(x_{t-1})) = -\alpha_{t-1}^{(t-1)}(x_{t-1},\mathbf{y}(x_{t-1})) = 0.$$

On the other hand,

$$\frac{d\alpha_{t+1}^{(t)}}{dx} = \frac{d}{dx}\frac{(r-t)a_{t+1}}{1+\sum_{l=t+1}^{r-1}(l-t)a_l}.$$

By (5.2.21), if we calculate the derivative of the right-hand side using the quotient rule, every term multiplies the function y_t or its first derivative, and they both vanish at the point x_{t-1}. This implies

$$\begin{aligned}\frac{d^2y_t}{dx^2}(x_{t-1}) &= \alpha_t^{(t)}(x_{t-1},\mathbf{y}(x_{t-1}))\frac{d\phi_t^{(t)}}{dx}(x_{t-1},\mathbf{y}(x_{t-1}))\\ &= \frac{d\phi_t^{(t)}}{dx}(x_{t-1},\mathbf{y}(x_{t-1}))\\ &= \frac{d}{dx}\frac{(r-t)(r-t+1)y_{t-1}(x_{t-1})\lambda(x_{t-1},\mathbf{y}(x_{t-1})}{s(x_{t-1},\mathbf{y}(x_{t-1}))^2}\\ &= -\left(1+\frac{y_{t-1}(x_{t-1})\lambda(x_{t-1},\mathbf{y}(x_{t-1})}{s(x_{t-1},\mathbf{y}(x_{t-1}))^2}\right)\frac{d\alpha_{t-1}^{(t-1)}}{dx}(x_{t-1}) > 0,\end{aligned}$$

as required. In the previous calculations, we used the definitions of $\alpha^{(t)}$ and $\alpha^{(t-1)}$ given in (5.2.19) and the definition of $\phi^{(t)}$ in (5.2.2). We also use the termination condition of Phase $t-1$ and the fact that y_t and its first derivative vanish at x_{t-1}. Note that, as mentioned above, this also verifies the case $k = t+1$ of part (b') in the assumption.

To conclude the verification of part (c') of Assumption 4.5.2, we look at the case $k = r$. Note that

$$\frac{dy_r}{dx} = \sum_{l=t}^{r-1}\frac{\alpha_l^{(t)}(r-l)\lambda y_{r-1}}{s^2} \equiv 0.$$

If $t < r-1$, then y_{r-1} is identically zero and $r \in Y_t$. If $t = r-1$, our result in the case $k = t$ implies that the function y_r and its derivatives of order one and two vanish at $x = x_{t-1}$, but its third derivative is positive at this point.

Finally, part (d') is an immediate consequence of the induction hypothesis assumed about γ at the end of Phase $t-1$, given by equations (5.2.16) and (5.2.17). This concludes the verification of Assumption 4.5.2 at the point x_{t-1}. If Phase t is not the final phase, we may

again redefine $\gamma > 0$ as a smaller positive number, if necessary, to ensure that

$$\text{(i)} \quad \gamma < y_k(x_t) < 1 - \gamma, \text{ if } k \notin Y_t;$$

$$\text{(ii)} \quad \frac{d\alpha_t^{(t)}}{dx}(x_t) \leq -\gamma.$$

Moreover, in the above calculations, we established that

$$y_k(x_t) > \gamma, \text{ if } k \leq t, \text{ and } y_k(x_t) = 0, \text{ if } k > t,$$

validating the induction hypothesis given in equations (5.2.16) and (5.2.17).

Let b denote the index of the final phase, as defined in Section 4.5. Since γ is redefined a finite number of times in the previous discussion, always being assigned a positive value, there exists $\gamma > 0$ for which the assumptions hold. As a consequence, by the discussion in Section 4.5, there is a solution to the following systems of differential equations, defined for $t \in \{1, \ldots, b\}$;

$$\frac{dy_k}{dx} = \sum_{l=t}^{r-1} \alpha_l^{(t)}(x, \mathbf{y}(x))\phi_k^{(l)}(x, \mathbf{y}(x)), \ 0 \leq k \leq r, \text{ if } x \in [x_{t-1}, x_t),$$
$$y_k(x_{t-1}) = \beta_{t-1,k}, 0 \leq k \leq r,$$

where $\beta_{0,0} = 1$, $\beta_{0,k} = 0$, if $k > 0$, and, if $t > 0$,

$$\beta_{t,k} = \lim_{x \to x_t^-} y_k(x),$$

with the points $x_1, x_2, \ldots, x_b$ being defined through the termination conditions. Furthermore, the solutions are within the region $\Omega_{\gamma,M}$ for every x in the interval $[0, x_b]$, while $\alpha_k^{(t)}(x, \mathbf{y}(x))$ lies in $[0, 1]$, for every $t \in \{1, \ldots, b\}$, every $k \in \{t, \ldots, r-1\}$ and every $x \in [x_{t-1}, x_t)$.

We shall call the constants defined above $x_0^0 = 0 < x_1^0 < \cdots < x_b^0$. These constants, the number of phases b and the functions $\mathbf{y}^0(x) = (y_k^0(x))_{k=0}^r$ will appear in Property $(b, p_0', \mathbf{x})$ with $p_0 = 0$, for some constant $p_0' > 0$ and functions $x_1, \ldots, x_b$. Before showing that these can determined, we discuss the dependency of b and $x_1^0, \ldots, x_b^0$ on γ. On the one hand, it is clear that, by choosing a smaller value of γ, there is no change in the values of $x_1^0, \ldots, x_{b-1}^0$, since all the conditions and termination conditions would be verified in the same way. However, if the only active termination condition at x_b^0 is $y_0(x_b^0) = \gamma$, it may be the case that, by decreasing γ, the solution of the differential equation could be extended beyond x_b^0. It is even conceivable that a different termination condition would then become active, which could potentially originate a new phase. However, if $b(\gamma)$ denotes the index of a final phase for a particular choice of γ, the limit

$$b = \lim_{\gamma \to 0^+} b(\gamma), \tag{5.2.23}$$

is well defined, since $b(\gamma)$ is bounded above by $r-1$ and is non-decreasing as γ decreases. Moreover, since $b(\gamma)$ is integer-valued, we know that this limit is achieved for γ sufficiently small. For this value of b, we may also define

$$\lambda(r) = \lim_{\gamma \to 0^+} x_b^0(\gamma), \tag{5.2.24}$$

since $x_b^0(\gamma)$ is again non-decreasing as γ decreases, and it is easy to see that 1 is an upper bound on its value.

Second part In order to determine the functions $x_1, \ldots, x_b$ for which the basic system of differential equations satisfies Property $(b, p_0', \mathbf{x})$, we discuss the dependency of the system of differential equations (5.2.1) upon its initial conditions, as we now look at this system with initial conditions, at $x = 0$, given by $\beta_{0,k} = w_{0,k}(p_0)$, where p_0 is now a small positive real. Recall that $\mathbf{w}_0$ is defined in (3.2.1) with parameters $d = 1$ and $\ell = 1$.

The following result will be especially useful. A proof of this result is standard, by the method of successive approximations (see [33], Theorem 2, Chapter 2, or [34], Section 3.22), and therefore is omitted.

Lemma 5.2.1 *Suppose that* $\mathbf{y}$ *satisfies the equations* $\dfrac{dw_i}{dx} = f_i(x, \mathbf{w})$ *in a bounded open set* $\Omega \subseteq \mathbb{R}^{n+1}$ *with initial conditions* $\mathbf{y}(0) = \mathbf{y}_0 = \mathbf{y}_0(\epsilon)$. *Let* $\mathbf{z}$ *denote another solution, with initial conditions* $\mathbf{z}(0) = \mathbf{z}_0 = \mathbf{z}_0(\epsilon)$. *Suppose that the functions* f_i *are Lipschitz on* Ω *and* $|\mathbf{y}_0(\epsilon) - \mathbf{z}_0(\epsilon)| \to 0$ *as* $\epsilon \to 0$. *Let*

$$x_1 = \inf\{x \,:\, (x, \mathbf{y}(x)) \notin \Omega \text{ or } (x, \mathbf{z}(x)) \notin \Omega\}.$$

Then $|\mathbf{y}(x) - \mathbf{z}(x)| \to 0$ *uniformly for* $x \in [0, x_1)$.

Let $p_0 > 0$ and consider the system of differential equations

$$\begin{aligned} \frac{dy_k^{(p_0)}}{dx} &= F_k^{(1)}(x, \mathbf{y}^{(p_0)}(x)) \\ y_k^{(p_0)}(0) &= w_{0,k}(p_0), \end{aligned} \tag{5.2.25}$$

given in Definition 5.1.1 with this value of p_0 and $t = 1$. Also consider the region $\Omega_{\gamma,\mu,M}$, where γ, μ and M are defined as in the proof of existence of a solution in the case $p_0 = 0$.

Note that, as functions of p_0, the initial conditions $w_{0,k}(p_0)$ are continuous. Moreover, it is clear that, if p_0 is sufficiently small, they are interior points of $\Omega_{\gamma,\mu,M}$, since $w_{0,k}(0)$ lies in the interior of this region.

So, we may again uniquely extend the solution $\mathbf{y}^{p_0}$ to (5.2.25) arbitrarily close to the boundary of $\Omega_{\gamma,\mu,M}$ through Lemma 2.5.1, as we did in the case $p_0 = 0$.

To show that this solution can be extended so as to satisfy the properties of Property $(b, p_0', \mathbf{x})$, we now verify that, if $p_0 > 0$ is sufficiently small, Assumption 4.5.1, with the

probability p_0 changed from 0 to a small positive constant, also holds for $\mathbf{y}^{p_0}$ at $x = 0$. The continuity of the initial conditions immediately implies that this is the case for part (a), for item (i) of parts (b) and (c), and for part (d).

So, the only parts of Assumption 4.5.1 that require additional discussion involve the definition of the indices in Y_1 and A_1 or item (ii) in parts (b) and (c). In the analysis of the case $p_0 = 0$, these assumptions only appear in the verification of the facts that

1. $$\frac{d\alpha_k^{(1)}}{dx} \equiv 0 \text{ if } k > 2,$$

2. $$\alpha_2^{(1)}(0, \mathbf{y}^0(0)) = 0 \text{ and } \frac{d\alpha_2^{(1)}}{dx}(0) > \gamma,$$

3. $$\frac{dy_k^0}{dx} \equiv 0, \text{ if } k > 1,$$

4. $$y_1^0(0) = 0 \text{ and } \frac{dy_1^0}{dx} > \gamma.$$

For the items 3 and 4, note that, because $\alpha^{(1)}$ is defined so as to imply that

$$\frac{dy_k^{p_0}}{dx} = \sum_{l=1}^{r-1} \alpha_l^{(1)} \phi_k^{(l)} \equiv 0 \tag{5.2.26}$$

for $1 < k \leq r-1$, we still have $\{2, \ldots, r-1\} \subseteq Y_1$. Also,

$$\frac{dy_r^{p_0}}{dx} = \sum_{l=1}^{r-1} \frac{\alpha_l^{(1)}(r-l)\lambda y_{r-1}^{p_0}}{s^2},$$

with

$$s = s(x, \mathbf{y}) = \sum_{k=0}^{r-1} (r-k) y_k$$

and

$$\lambda = \lambda(x, \mathbf{y}) = \sum_{k=0}^{r-2} (r-k)(r-k-1) y_k.$$

So, this derivative is positive at $x = 0$ when the initial conditions are given with $p_0 > 0$. More generally, we must have $y_r^{p_0}(x) \geq 0$ for every x for which $s(x, \mathbf{y}^{p_0}(x)) > 0$, and $y_{r-1}^{p_0}(x)$ and $\lambda(x, \mathbf{y}^{p_0}(x))$ are nonnegative. Finally, for $p_0 > 0$, we have $y_1^{p_0} > 0$, while, by continuity of the initial conditions, we still have

$$\frac{dy_1^{p_0}}{dx} > \gamma.$$

We now look at the conditions involving $\alpha_k^{(1)}$ for $k \geq 2$. From equation (5.2.8), we know that, for these values of k,

$$\alpha_k^{(1)} = \frac{(r-1)a_k}{1 + \sum_{l=2}^{r-1}(l-1)a_l}, \tag{5.2.27}$$

where

$$a_k = a_k(x, \mathbf{y}) = -(r-k)y_k/s + [(r-k+1)y_{k-1} - (r-k)y_k]\lambda/s^2.$$

By definition, for $(x, \mathbf{y}) \in \Omega_{\gamma,M}$, we have the bounds

$$r\gamma = s(r) \leq s(x, \mathbf{y}) \leq S(r) = \frac{r(r-1)}{2}$$

and

$$(r-1)(r-2)\gamma = \lambda(r) \leq \lambda(x, \mathbf{y}) \leq \Lambda(r) = \sum_{k=0}^{r-2}(r-k)(r-k-1).$$

On the one hand, this implies that, for $(x, \mathbf{y}) \in \Omega_{\gamma,M}$,

$$a_k = a_k(x, \mathbf{y}^{p_0}) \geq \frac{(r-k+1)\lambda(r)}{S(r)^2} y_{k-1}^{p_0} - \frac{(r-k)(S(r)+\Lambda(r))}{s(r)^2} y_k^{p_0}. \tag{5.2.28}$$

On the other hand, for $k \geq 1$,

$$y_k(0) = w_{0,k} = (1-p_0)^{r+1}\binom{r}{k}(1-p_0)^{(r-1)(r-k)}\left(1-(1-p_0)^{r-1}\right)^k,$$

which is a polynomial in the variable p_0 with minimum degree k whose coefficient of order k is positive. So, for $k \geq 2$, the lower bound on $a_k(0, \mathbf{y}^{p_0}(0))$ provided by equation (5.2.28) is a polynomial in p_0 with minimum degree $k-1$ whose coefficient of order k is positive. In particular, if p_0 is sufficiently small, $a_k(0, \mathbf{y}^{p_0}(0)) > 0$ for every such k.

However, for $k > 2$, we also know that the values of y_k are constant over the interval of definition of the solution to (5.2.25), since, by equation (5.2.26), their derivatives are identically zero in this interval. Hence, these lower bound also imply that $a_k(x, \mathbf{y}^{p_0}(x))) > 0$ for every $k > 2$ and every x in the interval of definition of a solution to (5.2.25). We conclude that $\alpha_k^{(1)}(x, \mathbf{y}^{p_0}(x))) > 0$ in this interval. Finally, in the case $k = 2$, we have, for $p_0 > 0$ sufficiently small, that

$$\alpha_2^{(1)}(0, y^{p_0}(0)) > 0 \text{ and } \frac{d\alpha_2^{(1)}}{dx}(0) > \gamma,$$

where the first part follows from the previous argument and the second part is a consequence of the continuity of the initial conditions.

This concludes the verification of Assumption 4.5.1 at $x = 0$, with the initial probability $p_0 = 0$ being replaced by a small positive value. As a consequence, by the same argument as in the case $p_0 = 0$, a solution to (5.2.25) can be extended up to $x_1^{p_0}$, the infimum over all $x > 0$ for which at least one of the termination conditions in (4.5.7) holds for $(x, \mathbf{y}^{p_0}(x))$.

We now argue that we may choose p_0' so that, for $0 < p_0 \leq p_0'$, any active termination condition at $(x_1^{p_0}, \mathbf{y}^{p_0}(x_1^{p_0}))$ is also an active termination condition at $(x_1^0, \mathbf{y}^0(x_1^0))$ in the case $p_0 = 0$.

Firstly, any of the termination conditions that appear at $(x_1^0, y^0(x_1^0))$, yet are not active at this point, cannot be active at $(x_1^{p_0}, y^{p_0}(x_1^{p_0}))$ for p_0 sufficiently small, because, by Lemma 5.2.1, the functions $\mathbf{y}^{p_0}(x)$ and $\mathbf{y}^0(x)$ are arbitrarily close to each other for x in the interval $[0, \min\{x_1^0, x_1^{p_0}\}]$.

Secondly, note that there could be termination conditions arising in the case $p_0 > 0$ which were not present for $p_0 = 0$. However, these conditions are only related to indices that are in the sets A_1 and Y_1 with respect to $p_0 = 0$, but are not in these sets when $p_0 > 0$. In our particular case, this happens for the indices $3, \ldots, r-1$ in A_1 and for the index r in Y_1. So, we have to show that, for p_0 sufficiently small, Phase 1 does not terminate because

$$\alpha_k^{(1)}(x_1^{p_0}, \mathbf{y}^{p_0}(x_1^{p_0})) = 0$$

for some $k \in \{3, \ldots, r-1\}$ or because $y_r^{p_0}(x_1^{p_0}) = 0$. For $\alpha^{(1)}$, this is an easy consequence of equations (5.2.27) and (5.2.28), while the derivative of $y_r^{p_0}$ has been shown to be nonnegative.

As a consequence, we may indeed choose p_0' so that, for $0 < p_0 \leq p_0'$, any active termination condition at $(x_1^{p_0}, \mathbf{y}^{p_0}(x_1^{p_0}))$ is also active at $(x_1^0, \mathbf{y}^0(x_1^0))$ in the case $p_0 = 0$. In particular, if Phase 1 is not the final phase with respect to $p_0 = 0$, the only active termination condition at $(x_1^{p_0}, \mathbf{y}^{p_0}(x_1^{p_0}))$ is

$$\alpha_1^{(1)}(x_1^{p_0}, \mathbf{y}^{p_0}(x_1^{p_0})) = 0 \text{ and } \frac{d\alpha_1^{(1)}}{dx}(x_1^{p_0}) > 0.$$

We may now apply Lemma 5.2.1 again to deduce that

$$\lim_{p_0 \to 0^+} x_1^{p_0} = x_1^0.$$

Note that the above discussion also implies that, for x in the interval $[0, x_1^{p_0}]$, we must have $y_k^{p_0}(x) > 0$, since, for $p_0 > 0$, all the components of $\mathbf{y}^{p_0}(0)$ are strictly positive, and, for $p_0 \leq p_0'$, the termination conditions $y_k^{p_0}(x) = 0$ do not hold for x in the interval $[0, x_1^{p_0})$, for every $\xi > 0$, and could only hold at $x = x_1^{p_0}$ if Phase 1 were the final phase.

The behaviour of $\mathbf{y}^{p_0}$ in the next phases can be analysed similarly. That is, we again derive it from the information we have about the solution in the case $p_0 = 0$, using Lemma 2.5.1 and Lemma 5.2.1. Most of the details may be verified in exactly the same way, and therefore are omitted. However, one case is different. Indeed, if x_{t-1} is the point of termination of Phase $t-1$, and this phase is not final, then we have shown in the case $p_0 = 0$ that

$$\frac{d^\nu y_t^0}{dx^\nu}(x_{t-1}) = 0, \text{ if } \nu \leq 1 \text{ and } \frac{d^2 y_t^0}{dx^2}(x_{t-1}) > \gamma.$$

In the case $p_0 > 0$, since the derivative of $y_t^{p_0}$ can be inductively shown to vanish in the intervals $[x_{s-1}^{p_0}, x_s^{p_0}]$ for $s \leq t-1$, we have

$$y_t^{p_0}(x_{t-1}) = y_t^{p_0}(0) > 0.$$

Recall that

$$\frac{dy_t^{p_0}}{dx} = \sum_{k=t}^{r-1} \alpha_k^{(t)} \phi_t^{(k)}. \tag{5.2.29}$$

Now, the definition of $\alpha^{(t)}$ in equation (5.2.19) implies that

$$\begin{aligned}\alpha_t^{(t)}(x_{t-1}^{p_0}, \mathbf{y}^{p_0}(x_{t-1}^{p_0})) &= \frac{1 - \sum_{l=t+1}^{r-1}(r-l)a_l(x_{t-1}^{p_0}, y^{p_0}(x_{t-1}^{p_0}))}{1 + \sum_{l=t+1}^{r-1}(l-t)a_l(x_{t-1}, y(x_{t-1}))} \\ &= \frac{(r-t)a_t(x_{t-1}^{p_0}, y^{p_0}(x_{t-1}))}{1 + \sum_{l=t+1}^{r-1}(l-t)a_l(x_{t-1}^{p_0}, y(x_{t-1}^{p_0}))},\end{aligned}$$

where the last step uses the fact that the quantity

$$1 - \sum_{l=t}^{r-1}(r-l)a_l(x_{t-1}^{p_0}, y^{p_0}(x_{t-1}^{p_0}))$$

is equal to the numerator of $\alpha_{t-1}^{(t-1)}(x_{t-1}^{p_0}, \mathbf{y}^{p_0}(x_{t-1}^{p_0}))$, which is equal to zero by the termination condition of Phase $t-1$. By the definitions of $\phi_k^{(t)}$ and of a_k, given in (5.2.2) and (5.2.6), respectively, we have

$$\phi_t^{(t)} = -1 + (r-t)a_t \text{ and } \phi_t^{(k)} = (r-k)a_t.$$

So, if we multiply the right-hand side of equation (5.2.29) by the positive quantity

$$1 + \sum_{l=t+1}^{r-1}(l-t)a_l,$$

we obtain

$$\begin{aligned}&(r-t)a_t\left(-1 + (r-t)a_t\right) + \sum_{k=t+1}^{r-1}(r-t)a_k(r-k)a_t \\ &= -(r-t)a_t\left(1 - \sum_{k=t}^{r-1}(r-k)a_k\right).\end{aligned}$$

But the term within brackets is precisely the numerator of $\alpha_{t-1}^{(t-1)}$, hence the first derivative of $y_t^{p_0}$ vanishes at $x_{t-1}^{p_0}$, as claimed. Finally, by continuity, we may again choose p_0 sufficiently small so as to have

$$\frac{d^2 y_t^0}{dx^2}(x_{t-1}) > \gamma,$$

concluding the verification of our property. Applying the argument used in the case $t = 1$, we may again choose p_0' so that, for $0 < p_0 \leq p_0'$, any active termination condition at $(x_{t-1}^{p_0}, \mathbf{y}^{p_0}(x_{t-1}^{p_0}))$ is also an active termination condition at $(x_{t-1}^0, \mathbf{y}^0(x_{t-1}^0))$ in the case $p_0 = 0$.

In particular, for p_0' sufficiently small and $p_0 \in [0, p_0']$, we find functions $\mathbf{y}^{p_0}(x) = (y_k^{p_0}(x))_{k=0}^r$ such that, for $t \in \{1, \ldots, b\}$,

$$\frac{dy_k^{p_0}}{dx} = F_k^{(t)}(x, \mathbf{y}^{p_0}(x)) \text{ for } x \in [x_{t-1}^{p_0}, x_t^{p_0}),\ k = 0, \ldots, r,$$
$$y_k^{p_0}(x_{t-1}^{p_0}) = \beta_{t-1,k}^{p_0},$$

where the initial conditions $\beta_{0,k}^{p_0}$ are defined as in Definition 5.1.1, and the constants $x_0^{p_0} = 0 < \cdots < x_b^{p_0}$ are defined through the termination conditions.

Now, recall that, for $p_0 > 0$, all the components of $\mathbf{y}^{p_0}(0)$ are strictly positive. Moreover, for $p_0 \leq p_0'$, none of the termination condition $y_k^{p_0}(x) = 0$, $k \in \{0, \ldots, r\}$, hold for x in the interval $[0, x_b^{p_0}]$, except possibly at $x = x_b^{p_0}$. Thus, for any $\xi > 0$,

$$\gamma_\xi^{p_0} = \inf\{y_k(x) : x \in [0, x_b^{p_0} - \xi],\ k \in \{0, \ldots, r\}\} > 0.$$

With the values of b, p_0' and $x_t^{p_0}$ defined as above, we are now ready to define functions

$$x_1, \ldots, x_b : [0, p_0'] \to [0, 1]$$

by setting

$$x_t(p_0) = x_t^{p_0}, \text{ for every } t \in \{1, \ldots, b\} \text{ and every } p_0 \in [0, p_0'].$$

By the work done above, it is clear that, for these values of b and p_0', and these functions $x_1, \ldots, x_b$, Property $(b, p_0', \mathbf{x})$ is satisfied.

5.3 A lower bound on the independence ratio

We shall now combine the results of the previous two sections to give new lower bounds $\lambda(r)$ on the independence ratio of r-regular graphs with large girth.

In the previous section, we proved that there exist a positive integer $b \in \{1, \ldots, r-1\}$, a constant $p_0' \in (0, 1]$ and functions

$$x_1, \ldots, x_b : [0, p_0'] \to [0, 1]$$

for which the system of differential equations in (5.1.7) satisfies Property $(b, p_0', \mathbf{x})$. To obtain these values, we first showed that exists $\gamma > 0$ for which the solution of (5.1.7) with initial conditions given with $p_0 = 0$ could be extended until a point $x_b(0)$, with the points of phase transition $x_1(0), \ldots, x_{b-1}(0)$ being determined by means of the termination conditions. We then proved that, with this value of γ, we could find a constant $p_0' > 0$ and functions $x_1, \ldots, x_b : [0, p_0'] \to [0, 1]$ for which Property $(b, p_0', \mathbf{x})$ holds.

For each $r \geq 3$, consider the positive integer $b \in \{1, \ldots, r-1\}$ defined in (5.2.23) and the constant $\lambda(r)$ defined in (5.2.24). We have the following theorem.

Theorem 5.3.1 *Let $\delta > 0$ and $r \geq 3$. Then there exists $g > 0$ such that every r-regular graph G on n vertices with girth greater than or equal to g satisfies $\alpha(G) \geq (\lambda(r) - \delta)n$.*

Proof Let $r \geq 3$ and let $\delta' > 0$. Let $\gamma > 0$ such that, in the definition of $\lambda(r)$, we have

$$|x_b(0) - \lambda(r)| \leq \frac{\delta'}{2}.$$

Consider $p_0' \in (0, 1]$ and the functions

$$x_1, \ldots, x_b : [0, p_0'] \to [0, 1],$$

as defined in the previous section, for which the system of differential equations in (5.1.7) satisfies Property $(b, p_0', \mathbf{x})$. By Theorem 5.1.2, there exists $g > 0$ such that every r-regular graph G on n vertices with girth greater than or equal to g satisfies

$$\alpha(G) \geq \left(x_b(0) - \frac{\delta'}{2}\right) n \geq (\lambda(r) - \delta')n,$$

as required. ∎

Now, it is easy to see that being an independent set is a vertex monotone property, as defined in Chapter 2. In particular, Lemma 2.1.2 and Lemma 2.1.4 immediately lead to the following results.

Corollary 5.3.2 *Let $\delta > 0$ and $r \geq 3$. Then there exists $g > 0$ such that every graph G on n vertices with maximum degree r and girth greater than or equal to g satisfies $\alpha(G) \geq (\lambda(r) - \delta)n$.*

Corollary 5.3.3 *Let $\delta > 0$ and $r \geq 3$. Then a random r-regular graph G on n vertices a.a.s satisfies $\alpha(G) \geq (\lambda(r) - \delta)n$.*

In the case of random r-regular graphs, the numbers $\lambda(r)$ coincide with the bounds given by Wormald in [56]. The work done here provides a new proof of the validity of these bounds, the first proof of this result that does not rely on sharp concentration.

Now, as seen in the previous section, the numbers $\lambda(r)$ are defined in terms of the system of differential equations (5.1.7). We "solved" the system numerically, without using careful error bounds, but just apparent good convergence as the step size was made smaller. The points of phase transition were determined by the termination conditions, as in Section 5.2. However, since the points of phase transition are not determined exactly, we also verified that, if we slightly perturb the points in which there is transition from one differential equation to the next, the overall change in the solutions was very small. Furthermore, we observed that, in a small interval in which a phase transition seems to occur, the values given by the numerical calculations suggest that the remaining conditions are "far" from being satisfied,

Table 5.3.1: Lower and upper bounds on $\alpha(G)/n$.

r	$\lambda_1(r)$	$\lambda_2(r)$	$\lambda(r)$	$\Lambda(r)$
3	0.4139	0.3750	0.4328	0.4554
4	0.3510	0.3333	0.3901	0.4163
5	0.3085	0.3016	0.3566	0.3844
6	0.2771	0.2764	0.3296	0.3580
7	0.2528	0.2558	0.3071	0.3357
8	0.2332	0.2386	0.2880	0.3165
9	0.2169	0.2240	0.2716	0.2999
10	0.2032	0.2113	0.2573	0.2852

hence we have numerically verified that the appropriate termination condition is active at the end of the phase, which ensures that the next phase does start, as defined in the framework of Section 4.5.

The numerical lower bounds on $\lambda(r)$ are given in Table 5.3.1 for some values of r. As a remark, we observe that, for all values of r for which a numerical bound was calculated, the final phase of the process was Phase $r-2$. The numbers $\gamma_1(r)$ and $\gamma_2(r)$ in this table approximate the lower bounds obtained by Shearer [50] and Lauer and Wormald [39], respectively, while the numbers $\Lambda(r)$ are the best upper bounds known to date, due to McKay [44].

An alternative numerical approach to obtain rigorous lower bounds is as follows. Since our system of differential equations is well-behaved, in the sense that the derivatives may be easily bounded (with the exception of a small interval close to the end of the process, for which the function $s = s(x, \mathbf{w})$ in the denominator may approximate zero), we could solve this system using interval arithmetic. That is, we could use a numerical method that, at every step, gives upper and lower bounds on the actual value of the solution at that point. At the same time, this would allow us to obtain upper and lower bounds on the values of the functions involved in the termination conditions. which allows us to identify an interval in which a phase transition has to occur. So, if our bounds show that all but one of the termination conditions cannot be satisfied in this interval, we can undoubtedly determine which termination condition became active first and, therefore, we may decide whether the final phase of the algorithm was reached or the next phase begins. This has been applied by Shi and Wormald in [51] to the solution of a system of differential equations arising in the analysis of an algorithm for colouring random regular graphs. In this work, they also had to determine points of phase transition. For more information concerning interval arithmetic, the reader is referred to [43]. A similar approach has been investigated in preparation for this thesis, but the bounds obtained were not very accurate.

Chapter 6

Induced forests

In Chapter 2, we have found lower bounds on the maximum number of vertices inducing a forest in regular graphs with large girth. The simple algorithm presented there has been generalised by the class of locally greedy algorithms. Indeed, a locally greedy algorithm coincides with Algorithm 2.2.1 if we set its parameters to be $d = 1$ and $\ell = 2$, and we fix the probabilities as follows. The initial probability is a constant $p_0 \in [0, 1]$ and, for some constant $p \in [0, 1]$, the remaining probabilities are $p_{i,0,k} = 0$ and $p_{i,1,k} = p$, for every value of k. We observe that this algorithm can be analysed using the framework of Chapter 4 just as the simple locally greedy for independent sets was analysed in Section 4.3, producing the lower bounds stated in Theorem 2.1.1.

These lower bounds will be improved in this chapter. The idea here is to define the probabilities assigned to the locally greedy algorithm in a cleverer way, which strongly resembles the work done for independent sets in Chapter 5.

6.1 A generalised system of differential equations

As with independent sets in Chapter 5, we shall classify the operations performed by the algorithm as different types and then rank them according to their benefit to the algorithm. We then look for probabilities in such a way that the ensuing locally greedy algorithm gives priority to operations with higher rank.

The *type* of an operation performed by this algorithm will again be defined in terms of the number of purple and yellow neighbours of a vertex added by the algorithm to the set of purple vertices, so that, for an r-regular graph, the operations are labelled by the set

$$\mathcal{I} = \{(j, k) \,:\, 0 \leq j \leq 1,\ 0 \leq k \leq r - j\}. \tag{6.1.1}$$

However, not all the possible types of operations will be performed by the algorithm. To understand why this is the case, we go back to the definition of Algorithm 3.1.1. Recall that there is an initialization step in which all vertices are chosen with some fixed probability p_0.

The chosen vertices are coloured purple, while the remaining vertices are coloured white if they have at most one purple neighbour and yellow otherwise. In a subsequent step, call it Step i, we choose each white vertex v with a given probability $p_{i,j,k}$, where j is the number of yellow neighbours of v and k is the number of purple neighbours of v. Again, the chosen vertices become purple, while white vertices turn yellow if their number of purple neighbours becomes larger than 1.

In terms of induced forests, the set P of purple vertices is precisely the set of vertices inducing the forest, while the set of yellow vertices corresponds to the set of vertices that will not be added to this induced forest because their addition may create a cycle in the graph. However, if all the purple neighbours of a yellow vertex lie in different components of $G[P]$, the addition of this vertex would not create a cycle, but just merge two different components of $G[P]$. Unfortunately, our algorithms cannot identify whether this is the case, and so, as a result, even yellow vertices merging two components are not allowed to be added to the induced forest and, although these extra components might lead to nonminimal forests, there is nothing that can be done about this.

Given that the extra components are wasteful, priority should be given to white vertices adjacent to a purple vertex. In fact, we shall require that only white vertices with a purple neighbour are added to the set of purple vertices after the initial step, that is, $p_{i,0,k}$ is set to be 0 for every $k \in \{0, \dots, r\}$. Note that, by not choosing white vertices that are not adjacent to purple neighbours, we may reach a point in which there are still white vertices, but none of them are adjacent to purple. For instance, this is expected to happen in the simple algorithm for induced forests in Chapter 2, and we analysed the "leftover" white vertices separately. However, we shall assign probabilities $p_{i,j,k}$ to a local greedy algorithm in a way such that this does not seem to occur, as indicated by the analysis of the algorithm that we will carry out.

Moreover, white vertices that do not have any white neighbours are also not selected, since they can be dealt with at the end of the algorithm. This is because, on the one hand, they will not become yellow due to the addition of a white vertex to P, since none of their neighbours are white. On the other hand, they can all be added to the induced forest at the end of the algorithm without creating cycles, since they have at most one purple neighbour and are not adjacent to each other.

In light of the above, we say that the algorithm performs an *operation of type* t when it chooses a white vertex adjacent to one purple vertex and to t yellow vertices. A vertex will be assigned higher priority if it has more yellow neighbours, since the addition of such a vertex to the the set of purple vertices is expected to create fewer new yellow vertices.

For $0 \leq t \leq r-2$ and $(j,k) \in \mathcal{I}$, with $\mathcal{I}$ defined as in (6.1.1), we introduce functions

$\phi_{j,k}^{(t)} : \mathbb{R}^{2r+1} \to \mathbb{R}$ given by

$$\begin{aligned}\phi_{j,k}{}^{(t)}(x, \mathbf{w}) &= -\delta_{k,t}\delta_{j,1} - (r-j-k)(r-t-1)w_{j,k}/s(\mathbf{w}) \\ &+ (r-t-1)\frac{(r-j-k+1)w_{j,k-1}\delta_{k\geq 1} - (r-j-k)w_{j,k}}{s(\mathbf{w})^2}\lambda(\mathbf{w}) \\ &+ \delta_{j,1}(r-j-k+1)(r-t-1)w_{j-1,k}/s(\mathbf{w}),\end{aligned} \tag{6.1.2}$$

where

$$s(\mathbf{w}) = s(w_{0,0}, \dots, w_{0,r}, w_{1,0}, \dots, w_{1,r-1}) = \sum_{j''=0}^{1}\sum_{k''=0}^{r-j}(r-j''-k'')w_{j'',k''} \tag{6.1.3}$$

and

$$\lambda(\mathbf{w}) = \sum_{k''=0}^{r-3}(r-k''-1)(r-k''-2)w_{1,k''}. \tag{6.1.4}$$

We note that, as in the case of independent sets, the formula for $\phi_{j,k}^{(t)}$ may be deduced by looking at the rate of change of the number of vertices with j yellow neighbours and k purple neighbours when the algorithm performs operation of type t. Also, since the value of t ranges from 0 to $r-2$, and the first phase applies operations of type 0, it is convenient to refer to it as Phase 0.

Now, the proportion $\alpha_k^{(t)}$ of operations of type k performed by the algorithm while in Phase t, for a fixed $t \in \{0, \dots, r-2\}$, is given by the solution to the linear system (4.4.6) with our choice of functions $\phi_{j,k}^{(t)}$, so that

$$\alpha_k^{(t)}(x, \mathbf{w}) = \begin{cases} \dfrac{1 - \sum_{l=t+1}^{r-2}(r-l-1)a_l}{1 + \sum_{l=t+1}^{r-2}(l-t)a_l}, & \text{if } k = t \\ \dfrac{(r-t-1)a_k}{1 + \sum_{l=t+1}^{r-2}(l-t)a_l}, & \text{if } k > t. \end{cases} \tag{6.1.5}$$

Here,

$$a_k = a_k(x, \mathbf{w}) = -(r-k-1)w_{1,k}/s + (r-k)w_{0,k}/s + [(r-k)w_{1,k-1} - (r-k-1)w_{1,k}]\lambda/s^2,$$

with $s = s(x, \mathbf{w})$ and $\lambda = \lambda(x, \mathbf{w})$ defined as in (6.1.3) and (6.1.4), respectively.

We now define a basic system of differential equations, as in the case of independent sets. For a fixed set of constants $0 = x_{-1} < x_0 < \cdots < x_{r-2}$, which represent the points of phase transition in the algorithm, a phase $t \in \{0, \dots, r-2\}$ and a set of initial conditions $\mu_t = (\mu_{t,j,k})_{(j,k)\in\mathcal{I}} \in \mathbb{R}^{|\mathcal{I}|+1}$, we consider the basic system of differential equations

$$\begin{aligned}&\frac{dw_{j,k}}{dx} = F_{j,k}^{(t)}(x, \mathbf{w}(x)) \text{ for } x \in [x_{t-1}, x_t),\ (j,k) \in \mathcal{I}, \\ &w_{j,k}(x_t) = \mu_{t,j,k},\end{aligned} \tag{6.1.6}$$

where

$$F_{j,k}^{(t)}(x,\mathbf{w}) = \begin{cases} \sum_{l=0}^{r-2} \alpha_l^{(t)} \phi_{j,k}^{(l)} & t \le r-3 \\ \phi_k^{(r-2)}(x,\mathbf{w}), & t = r-2 \end{cases} \tag{6.1.7}$$

with functions $\phi_{j,k}^{(l)}$ defined as in (6.1.2) and $\alpha_k^{(t)}$ given by equation (6.1.5), if $k \ge t$, and by

$$\alpha_k^{(t)}(x,\mathbf{y}) = 0, \text{ if } k < t.$$

Now, to devise a locally greedy algorithm for induced forests whose performance approximates the solution to the system of differential equations in (6.1.6), we shall again use Property $(b, p_0', \mathbf{x})$.

Definition 6.1.1 ***(Property*** $(b, p_0', \mathbf{x})$***)*** *Let $b \in \{0,\dots,r-2\}$ and $p_0' \in (0,1]$. With respect to a vector $\mathbf{x}$ of fixed functions*

$$x_0,\dots,x_b : [0,p_0'] \to [0,1],$$

where $0 < x_0(p_0) < x_1(p_0) < \cdots < x_b(p_0)$, for every p_0, we say that the system of differential equations (6.1.6) satisfies Property $(b,p_0',\mathbf{x})$ if the following holds. For $p_0 \in [0,p_0']$, there are bounded functions $\mathbf{y}^{p_0}(x) = (y_{j,k}^{p_0}(x))_{(j,k)\in\mathcal{I}}$ defined for x in the interval $[0, x_b(p_0))$ satisfying, for $t \in \{0,\dots,b\}$,

$$\begin{aligned} &\frac{dy_{j,k}^{p_0}}{dx} = F_{j,k}^{(t)}(x,\mathbf{y}^{p_0}(x)),\ (j,k)\in\mathcal{I},\ \text{for } x \in [x_{t-1}(p_0), x_t(p_0)), \\ &y_{j,k}^{p_0}(x_{t-1}(p_0)) = \beta_{t-1,j,k}^{p_0},\ (j,k)\in\mathcal{I}. \end{aligned} \tag{6.1.8}$$

Here, the functions $F_{j,k}^{(t)}$ are defined as in (6.1.7) and $\beta_{t,j,k}^{p_0}$ is defined as follows. If $t=-1$, it is equal to the initial condition $w_{0,j,k}(p_0)$ defined in equation (3.2.1) with $d=1$ and $\ell=2$. If $t \ge 0$,

$$\beta_{t,j,k}^{p_0} = \lim_{x \to x_t(p_0)^-} y_{j,k}^{p_0}(x).$$

Moreover, the following additional properties hold.

1. *For x in the interval $[x_{t-1}(p_0), x_t(p_0))$, we have $0 \le \alpha_k^{(t)}(x,\mathbf{y}^{p_0}(x)) \le 1$ for every $k \in \{t,\dots,r-2\}$.*

2. *Given $p_0 \in (0,p_0']$ and $\xi > 0$, there exists $\gamma_\epsilon^{p_0} > 0$ such that*

$$\gamma_\epsilon^{p_0} < y_{j,k}^{p_0}(x) < 1 - \gamma_\epsilon^{p_0}$$

for every $x \in [0, x_b(p_0) - \xi]$ and every $(j,k) \in \mathcal{I}$. In particular, $(x,\mathbf{y}(x)) \in \Omega_{\gamma_\epsilon^{p_0}, M^{p_0}}$ for every $x \in [0, x_b(p_0)-\xi]$, where M^{p_0} is an upper bound on the functions $y_{j,k}^{p_0}$.

3. *For each* $t \in \{0, \ldots, r-2\}$, $\lim_{p_0 \to 0^+} x_t(p_0) = x_t(0)$, *and, for every* $\xi > 0$, $|y^{p_0}_{j,k} - y^0_{j,k}| \to 0$ *uniformly in the interval* $[0, x^0_b - \xi]$ *as* $p_0 \to 0^+$.

The following theorem establishes that, if the system of differential equations in (6.1.6) satisfies Property $(b, p_0', \mathbf{x})$ for some b, p_0' and $\mathbf{x}$, we may derive bounds on the minimum number of vertices in a largest induced forest of a regular graph with sufficiently large girth, where the bounds depend on the values of b, p_0' and $\mathbf{x}$. A "good" choice of parameters will be defined in Section 6.2.

Theorem 6.1.2 *Let* $b \in \{0, \ldots, r-2\}$ *and* $p_0' \in (0, 1]$, *and consider functions*

$$x_0, \ldots, x_b : [0, p_0'] \to [0, 1]$$

for which the basic system of differential equations defined in equation (6.1.6) satisfies Property $(b, p_0', \mathbf{x})$, *and let* $\mathbf{y} = \mathbf{y}(x)$ *be the solution to this system. Then, for* $\delta > 0$ *and* $r \geq 3$, *there exists* $g > 0$ *such that every* r*-regular graph* G *on* n *vertices with girth greater than or equal to* g *satisfies*

$$\tau(G) \geq (x_b(0) + y_{0,r}(x_b(0)) + y_{1,r-1}(x_b(0)) - \delta)n,$$

where $\mathbf{y} = y(x)$ *is a solution to the system of differential equations (6.1.8) with initial conditions given with* $p_0 = 0$.

Proof Let $\delta' > 0$. By item 3 in the definition of Property $(b, p_0', \mathbf{x})$, we may choose p_0 with the property that $|x_b(p_0) - x_b(0)| < \delta'/8$. Also, if we fix $\xi = \delta'/8$, the constant p_0 may be chosen as to further satisfy

$$|y^{p_0}_{0,r}(x_b(p_0) - \xi) - y_{0,r}(x_b(0))| < \frac{\delta'}{8} \text{ and } |y^{p_0}_{1,r-1}(x_b(p_0) - \xi) - y_{1,r-1}(x_b(0))| < \frac{\delta'}{8},$$

by item 3 of Property $(b, p_0', \mathbf{x})$ and by continuity of the functions $y_{0,r}$ and $y_{1,r-1}$.

As a consequence, we have

$$|x_b(p_0) - \xi + y_{0,r}(x_b(p_0) - \xi) + y_{1,r-1}(x_b(p_0) - \xi) - x_b(0) - y_{0,r}(x_b(0)) - y_{1,r-1}(x_b(0))| < \frac{\delta'}{2}.$$

As in the case of independent sets, we may now define $\hat{p}_k : \mathbb{R} \to [0, x_b(p_0) - \xi]$ by

$$\hat{p}_{j,k}(x) = \begin{cases} \dfrac{\alpha^{(t)}_k(x, \mathbf{y}(x))}{y^{p_0}_{1,k}(x)}, & \text{if } j = 1,\ t \leq k,\ x \in [x_{t-1}(p_0), x_t(p_0)),\ t = 0, \ldots, \min\{b, r-3\} \\ \dfrac{1}{y^{p_0}_{1,r-2}(x)}, & \text{if } j = 1,\ k = b = r-2 \text{ and } x \in [x_{r-3}(p_0), x_{r-2}(p_0)), \\ 0, & \text{in all other cases.} \end{cases}$$

Items 1 and 2 in the definition of Property$(b, p_0', \mathbf{x})$ imply that, if the constant

$$C > \frac{1}{\gamma^{p_0}_\xi}$$

is fixed, we have $0 \leq \hat{p}_{j,k}(x) < C$, for every $x \in [0, x_b(p_0) - \xi]$.

The functions $\mathbf{F} = (F_{j,k})_{(j,k)\in\mathcal{I}}$ and $\mathbf{E} = (E_{j,k})_{(j,k)\in\mathcal{I}}$ associated with the algorithm in the case $d = 1$ and $\ell = 2$ were obtained in (3.4.4). If $p_{i,j,k} = \epsilon\hat{p}_{j,k}(\epsilon i)$, they induce, in the sense of Section 4.2, the following system of differential equations, for $t \in \{0, \ldots, b\}$;

$$\begin{aligned} &\frac{dw_{j,k}}{dx} = f_{j,k}^{(t)}(x, \mathbf{w}(x)),\ (j,k) \in \mathcal{I},\ \text{for } x \in [x_{t-1}(p_0), x_t(p_0)), \\ &w_{j,k}(x_{t-1}(p_0)) = \beta_{t,j,k},\ (j,k) \in \mathcal{I}, \end{aligned} \tag{6.1.9}$$

where

$$f_{j,k}^{(t)}(x, \mathbf{w}) = \sum_{l=0}^{r-2} \alpha_l^{(t)}(x, \mathbf{y}^{p_0}(x))\phi_{j,k}^{(l)}(x, \mathbf{w}),$$

and the initial conditions $\beta_{t,j,k}^{p_0}$ are defined as follows. If $t = -1$, it is equal to the initial condition $w_{0,j,k}(p_0)$ defined in equation (3.2.1) with $d = 1$ and $\ell = 2$. If $t \geq 0$,

$$\beta_{t,j,k}^{p_0} = \lim_{x \to x_t^{p_0-}} y_{j,k}^{p_0}(x).$$

Note that the function $\mathbf{y}^{p_0} = \mathbf{y}^{p_0}(x)$ in the above equation is the solution of (6.1.8).

Recall that the equivalent theorem in the case of independent sets has been obtained by applying the results of Section 4.2. In order to use a similar argument here, we need to check the validity of the properties in Assumption 4.2.1.

The fact that Property (P_3) is satisfied follows immediately from the solution to (6.1.9), which coincides with the solution $\mathbf{y}^{p_0}$ to (6.1.8) given by Property $(b, p_0', \mathbf{x})$.

We look at the remaining properties of Assumption 4.2.1. For Property (P_1), note that the denominators of the coefficients of each polynomial $E_{j,k}$ are powers of the term $s(x, \mathbf{w})$ defined in (6.1.3), and that $s(x, \mathbf{w})$ has no zeroes in the region $\Omega_{\gamma_\xi^{p_0}, M^{p_0}}$. So, the coefficients of $E_{j,k}$ do not have poles in $\Omega_{\gamma_\xi^{p_0}, M^{p_0}}$, and Property (P_1) is satisfied.

Property (P_2) also holds, since the functions $\phi_{j,k}^{(t)}$ are Lipschitz in the region $\Omega_{\gamma_\xi^{p_0}, M^{p_0}}$ and the terms $\alpha_l^{(t)}(x, \mathbf{y}^{p_0}(x))$ are continuous and bounded in $[x_{t-1}(p_0), x_t(p_0))$, and therefore each $f_{j,k}^{(t)}$ is Lipschitz in the region

$$\Omega_{\gamma_\xi^{p_0}, M^{p_0}} \cap ([x_{t-1}(p_0), x_t(p_0)) \times \mathbb{R}^{|\mathcal{I}|}).$$

This concludes the verification of Assumption 4.2.1 in this case. To simplify the notation, we again refer to the interval $[x_{b-1}(p_0), x_b(p_0))$ even if the interval $[x_{b-1}(p_0), x_b(p_0) - \xi]$ is really meant.

We now consider a locally greedy algorithm with parameters $d = 1$ and $\ell = 2$ and r, and with initial probability p_0. For $0 < \epsilon < 1/C$, suppose that the number of steps is

$$N = \left\lfloor \frac{x_b(p_0) - \xi}{\epsilon} \right\rfloor$$

and the remaining probabilities are $p_{i,j,k} = \epsilon \hat{p}_{j,k}(i\epsilon)$.

With the same argument as in the proof of Theorem 5.1.2, we may show that there is $\epsilon_1 > 0$ for which the solution $\mathbf{w}_i^\epsilon$ to (4.2.5) cannot leave $\Omega_{\gamma_\xi^{p_0}, M^{p_0}}$ for $i \leq N = \lfloor (x_b(p_0) - \xi)/\epsilon \rfloor$ if $0 < \epsilon < \epsilon_1$, so

$$N_f = N_f(\epsilon) = \left\lfloor \frac{x_b(p_0) - \xi}{\epsilon} \right\rfloor.$$

We may now apply Theorem 4.2.4 with $M = M^{p_0}$, $\gamma = \gamma_\xi^{p_0}$, $m = b$, $x_m = x_b(p_0) - \xi$ and $\delta = \delta'/2$. This establishes the existence of $0 < \epsilon < \epsilon_1$ such that, if a locally greedy algorithm is applied, with the probabilities given by this value of ϵ, to an r-regular graph G with girth larger than $4N + 4$, then the expected size of the induced forest returned satisfies

$$\begin{aligned}
\mathbf{E}|\bar{P}(\epsilon)| &\geq n \left(p_0(1-p_0)^r + \int_0^{x_b(p_0)-\xi} h(x, \hat{\mathbf{w}}(x))\, dx - \delta'/2 \right) \\
&= n \left(p_0(1-p_0)^r + \int_0^{x_b(p_0)-\xi} dx - \delta'/2 \right) \\
&= n(p_0(1-p_0)^r + x_b(p_0) - \xi - \delta'/2),
\end{aligned}$$

where $h(x)$ denotes the function $\sum_{(j,k)\in\mathcal{I}} \hat{p}_{j,k}(x)\hat{w}_{j,k}(x)$. The fact that $h(x, \hat{\mathbf{w}}(x)) = 1$ follows immediately from the fact that the systems of differential equation (6.1.8) and (6.1.9) have the same solutions for a fixed p_0. Now, we may obtain a larger induced forest by adding, to the set $\bar{P}$, any vertices that are in $W_N^{0,r}$ and $W_N^{1,r-1}$, since their addition does not create any cycles. Now, the expected number of vertices in each of these sets are given by $n\hat{w}_{0,r}(x_b(p_0) - \xi)$ and $n\hat{w}_{1,r-1}(x_b(p_0) - \xi)$, respectively. As a consequence, by the first moment principle, every n-vertex, r-regular graph G with girth at least

$$g = g(\epsilon) = \left\lfloor \frac{4(x_b(p_0) - \xi)}{\epsilon} \right\rfloor + 4$$

has an induced forest with at least

$$n\left(p_0(1-p_0)^r + x_b(p_0) - \xi + \hat{w}_{0,r}(x_b(p_0) - \xi) + \hat{w}_{1,r-1}(x_b(p_0) - \xi) - \delta'/2\right)$$

vertices. By our choice of p_0, this leads to

$$\alpha(G) \geq n\left(x_b(0) + y_{0,r}(x_b(0)) + y_{1,r-1}(x_b(0)) - \delta'\right),$$

concluding the proof of the theorem. ∎

6.2 An appropriate choice of parameters

In this section, we shall how to find a positive integer $b \in \{1, \ldots, r-1\}$, a constant $p_0' \in (0, 1]$ and functions

$$x_1, \ldots, x_b : [0, p_0'] \to [0, 1]$$

so as to obtain a good bound for induced forests using Theorem 6.1.2. The discussion here is completely analogous to the work done in Section 5.2. However, the system of differential equations under consideration is now given in (6.1.6).

As before, the first part of our discussion analyses the system of differential equations in (6.1.8) when the initial conditions are given with $p_0 = 0$, while, in the second part, we derive solutions to (6.1.8) with initial conditions given with respect to a small positive integer p_0 by analysing the sensitivity of this system with respect to its initial conditions.

First part We show that, if the initial conditions of the differential equations (6.1.8) are defined with $p_0 = 0$, a solution can be obtained through the framework of Section 4.5. For a fixed $r \geq 3$, let

$$\mathcal{I} = \{(j,k) \in \mathbb{Z}^2 \,:\, 0 \leq j \leq 1,\ 0 \leq k \leq r-j\},$$

let i_0 be the element $(0,0)$ and let the index set $\mathcal{J}$ of the variables associated with each operation in the algorithm be given by

$$\{(1,0),(1,1),\ldots,(1,r-2)\}.$$

As in the definition of the functions $F^{(t)}_{j,k}$ in (6.1.7), we let the functions $\phi^{(t)}_{j,k}$ be given by

$$\begin{aligned}\phi^{(t)}_{j,k}(x,\mathbf{w}) &= -\delta_{k,t}\delta_{j,1} - (r-j-k)(r-t-1)w_{j,k}/s(\mathbf{w}) \\ &\quad + (r-t-1)\frac{(r-j-k+1)w_{j,k-1}\delta_{k\geq 1} - (r-j-k)w_{j,k}}{s(\mathbf{w})^2}\lambda(\mathbf{w}) \\ &\quad + \delta_{j,1}(r-j-k+1)(r-t-1)w_{j-1,k}/s(\mathbf{w}),\end{aligned} \tag{6.2.1}$$

which are the functions $\phi^{(t)}_{j,k}$ defined in (6.1.2), where

$$s = s(x,\mathbf{y}) = \sum_{j=0}^{1}\sum_{k=0}^{r-1-j}(r-j-k)y_{j,k} \tag{6.2.2}$$

and

$$\lambda = \lambda(x,\mathbf{y}) = \sum_{k=0}^{r-3}(r-k-1)(r-k-2)y_{1,k}. \tag{6.2.3}$$

Since, in this case, the basic operation of the first phase is $j = 1$ and $k = 0$, and the operations undertaken by the locally greedy algorithm are determined by the value of k, we shall call the first phase Phase 0 to simplify the use of the indices. Also, the endpoints of the phases will be denoted $0 = x_{-1} < x_0 < \cdots < x_{r-2}$ so that Phase k terminates at the point x_k.

As in the case of independent sets, we have to show that there exist constants M, γ and μ for which the assumptions in Section 4.5 hold, that is, for which Assumption 4.5.1 holds at $x = 0$ and Assumption 4.5.2 holds at the point $x = x_{t-1}$, whenever x_{t-1} is the termination point of Phase $t-1$, and Phase $t-1$ is not the final phase.

Let $M > 1$, let $0 < \gamma < 1$ be some fixed value, which may be redefined as a smaller constant during the argument, and let

$$0 < \mu < \frac{r\gamma}{\sum_{(j,k)\in\mathcal{I}\setminus\{(0,0)\}}(r-j-k)}$$

be defined in terms of γ. This choice of μ ensures that the functions $\phi_{j,k}^{(t)}$ are Lipschitz in $\Omega_{\gamma,\mu,M}$, since the term

$$s(x,\mathbf{y}) = \sum_{j''=0}^{1}\sum_{k''=0}^{r-j}(r-j''-k'')y_{j'',k''},$$

which is the only term to appear as a denominator of the rational functions $\phi_{j,k}^{(t)}(x,\mathbf{w})$, is positive in $\Omega_{\gamma,\mu,M}$.

Our first step is to verify that Assumption 4.5.1 is satisfied at the point $x = 0$. Recall that we are looking for a solution to

$$\begin{aligned}&\frac{dy_{j,k}}{dx} = \sum_{l=0}^{r-2}\alpha_l^{(0)}(x,\mathbf{y}(x))\phi_{j,k}^{(l)}(x,\mathbf{y}(x)),\ (j,k)\in\mathcal{I},\\ &y_{0,0}(0) = 1,\ y_{j,k}(0) = 0 \text{ if } (j,k)\neq(0,0),\end{aligned} \tag{6.2.4}$$

where the vector $\alpha^{(0)} = (\alpha_l^{(0)})_{l=0}^{r-2}$ satisfies the linear system given by

$$M^{(0)}\,u = v, \tag{6.2.5}$$

where

$$u = \begin{bmatrix}\alpha_0^{(0)}\\ \alpha_1^{(0)}\\ \alpha_2^{(0)}\\ \cdot\\ \cdot\\ \alpha_{r-2}^{(0)}\end{bmatrix} \text{ and } v = \begin{bmatrix}1\\0\\0\\ \cdot\\ \cdot\\ 0\end{bmatrix},$$

and the matrix $M^{(0)}(x,\mathbf{y}(x))$ is equal to $M(r-1,a_1,\ldots,a_{r-2})$, as defined in (5.2.5). For our values of $\phi_{j,k}^{(t)}$, the elements a_k are defined as

$$\begin{aligned}a_k = a_k(x,\mathbf{y}) &= -(r-k-1)y_{1,k}/s + (r-k)y_{0,k}/s\\ &\quad + [(r-k)y_{1,k-1} - (r-k-1)y_{1,k}]\lambda/s^2,\ \text{ for } k = 0,\ldots,r-2,\end{aligned} \tag{6.2.6}$$

with $s = s(x,\mathbf{y})$ and $\lambda(x,\mathbf{y})$ defined as in (6.2.2) and (6.2.3), respectively.

In our current notation, Assumption 4.5.1 is given as follows.

Assumption 4.5.2 ***(Assumptions for Phase 0)***

Consider the sets

$$A_0 = \left\{ k \in \{1, \dots, r-2\} : \frac{d\alpha_k^{(0)}}{dx} \equiv 0 \right\} \quad \textit{and} \quad Y_0 = \left\{ (j,k) \in \mathcal{I} : \frac{dy_{j,k}}{dx} \equiv 0 \right\}.$$

The system of differential equations (6.2.4) satisfies

(a) $|\det M^{(0)}(0, \mathbf{y}(0))| > \gamma$.

(b) $\alpha_0^{(0)}(0) > \gamma$, *and for each* $k \in \{1, \dots, r-2\} \setminus A_0$, *at least one of the following holds:*

(i) $\alpha_k^{(0)}(0) > \gamma$,

(ii) $0 \le \alpha_k^{(0)}(0)$, *the set* $A_{0,k} = \left\{ s \ge 1 \, : \, \frac{d^s \alpha_k^{(0)}}{dx^s}(0) \ne 0 \right\}$ *is not empty and is such that* $\frac{d^\nu \alpha_k^{(0)}}{dx^\nu}(0) > \gamma$, *where* $\nu = \min A_{0,k}$;

(c) for each $(j,k) \in \mathcal{I} \setminus Y_0$, *at least one of the following holds*

(i) $y_{j,k}(0) > \gamma$,

(ii) $0 \le y_{j,k}(0)$, *the set* $Y_{0,j,k} = \left\{ s \ge 1 \, : \, \frac{d^s y_{j,k}}{dx^s}(0) \ne 0 \right\}$ *is not empty and is such that* $\frac{d^\nu y_{j,k}}{dx^\nu}(0) > \gamma$, *where* $\nu = \min\{s \, : \, s \in Y_{0,j,k}\}$;

(d) $\frac{dy_{0,0}}{dx}(0) < -\gamma$.

Again using Proposition B.1.6,

$$|\det M^{(0)}(0, \mathbf{y}(0))| = \left| (-1)^r \left(1 + \sum_{l=1}^{r-2} l \, a_l(0, \mathbf{y}(0)) \right) \right| = 1 > \gamma,$$

since $y_{j,k}(0) = 0$ if $(j,k) \ne (0,0)$. This establishes part (a) of the previous assumption. Moreover,

$$\alpha_k^{(0)}(x, \mathbf{y}) = \begin{cases} \dfrac{1 - \sum_{l=1}^{r-2} (r-l-1) a_l}{1 + \sum_{l=1}^{r-2} l \, a_l}, & \text{if } k = 0 \\ \dfrac{(r-1) a_k}{1 + \sum_{l=1}^{r-2} l \, a_l}, & \text{if } k > 0. \end{cases} \tag{6.2.7}$$

We now verify parts (b) and (c) of Assumption 4.5.1. The former is true for $l = 0$, since

$$\alpha_0^{(0)}(0, \mathbf{y}(0)) = 1 > \gamma,$$

and the latter holds for $(0,0)$ and $(1,k)$, $1 \leq k \leq r-2$, since, by the initial conditions, $y_{0,0}(0) = 1 > \gamma$, while, by our choice of $\alpha^{(0)}$,

$$\frac{dy_{1,k}}{dx} \equiv 0,$$

hence $\{(1,1),\ldots,(1,r-2)\} \subseteq Y_0$, with Y_0 defined in the statement of Assumption 4.5.1. So,

$$y_{1,k}(x) = 0 \text{ for every } k \in \{1,\ldots,r-2\}. \tag{6.2.8}$$

Because of this, equation (6.2.6) leads to

$$a_1(x,\mathbf{y}(x)) = (r-1)y_{0,1}(x)/s(x,\mathbf{y}(x)) + (r-1)y_{1,0}(x)\lambda(x,\mathbf{y}(x))/s(x,\mathbf{y}(x))^2 \tag{6.2.9}$$

and to

$$a_k(x,\mathbf{y}(x)) = (r-k)y_{0,k}(x)/s(x,\mathbf{y}(x)), \text{ if } 2 \leq k \leq r-2. \tag{6.2.10}$$

We now look at part (c) of the assumption for the element $(1,0)$. By the initial conditions in (6.2.4), we have $y_{1,0}(0) = 0$. By equations (6.2.7) and (6.2.6), we also have $\alpha_0^{(0)}(0,\mathbf{y}(0)) = 1$ and $\alpha_k^{(0)}(0,0) = 0$ if $k > 0$. So,

$$\begin{aligned}\frac{dy_{1,0}}{dx}(0) &= \sum_{l=0}^{r-2} \alpha_l^{(0)}(0)\phi_{1,0}^{(l)}(0,\mathbf{y}(0)) \\ &= \phi_{1,0}^{(0)}(0,\mathbf{y}(0)) = \frac{r(r-1)y_{0,0}(0)}{s(0)} - 1 = r-2 > \gamma,\end{aligned}$$

so that $(1,0)$ satisfies item (ii) in part (c) of Assumption 4.5.1 with $\nu = 1$.

We now show that part (b) holds for $\alpha_k^{(0)}$, whenever $k \in \{1,\ldots,r-2\}$, while part (c) holds for $y_{0,k}$ if $k \geq 1$. The initial conditions in (6.2.4), together with equations (6.2.6) and (6.2.7), implies that all these quantities are equal to zero at the point $x = 0$. The following result will be useful for this purpose.

Claim 6.2.1 *For $1 \leq a \leq r$,*

$$\begin{aligned}&\frac{d^{2a-1}y_{0,k}}{dx^{2a-1}}(0) = 0, \text{ for every } k \geq a, \\ &\frac{d^{2a}y_{0,a}}{dx^{2a}}(0) = \frac{(2a-1)!!\ r!\ (r-1)^a (D\lambda(0,\mathbf{y}(0)))^a}{(r-a)!\ s(0,\mathbf{y}(0))^{2a}}, \text{ and } \frac{d^{2a}y_{0,k}}{dx^{2a}}(0) = 0, \text{ if } k > a,\end{aligned} \tag{6.2.11}$$

where $D\lambda$ denotes the derivative $\dfrac{d\lambda(x,\mathbf{y}(x))}{dx}$ and $(2a-1)!!$ denotes the product of all odd numbers smaller than or equal to $2a-1$, respectively.

Before proving this claim, we argue that it fulfils our objective. Clearly, it implies part (c) of Assumption 4.5.1 for $y_{0,k}$, $k \in \{1,\ldots,r\}$, since it shows that its $2k$-th derivative is positive at $x = 0$, hence larger than γ, possibly by redefining γ as a smaller positive constant,

while all derivatives of smaller order vanish at this point. By (6.2.6), the same fact is true for a_k, $k \in \{1, \ldots, r-1\}$, hence the same holds for $\alpha_k^{(0)}$ by (6.2.7), leading to part (b) for these values of k, and therefore concluding the verification of part (b). To conclude the verification of part (c), we look at the case $(1, r-1)$. By (6.2.4) and the above claim, it is easy to see that the nonzero derivative of smallest order of $y_{1,r-1}$, which is the derivative of order $2r-1$, is nonnegative.

We now prove the above claim by induction on a.

Proof of the Claim For $a = 1$,

$$\frac{d^{2a-1}y_{0,k}}{dx^{2a-1}}(0) = \frac{dy_{0,k}}{dx}(0) = \sum_{l=0}^{r-2} \alpha_l^{(0)}(0, \mathbf{y}(0))\phi_{0,k}^{(l)}(0, \mathbf{y}(0)) = \phi_{0,k}^{(0)}(0, \mathbf{y}(0)),$$

since $\alpha_l^{(0)}(0, \mathbf{y}(0)) = \delta_{0,l}$. Now,

$$\begin{aligned}\phi_{0,k}^{(l)}(x, \mathbf{y}) &= -\frac{(r-k)(r-l-1)y_{0,k}}{s(0, \mathbf{y}(0))} \\ &+ \frac{(r-k+1)(r-l-1)y_{0,k-1}\lambda(0, \mathbf{y}(0))}{s(0, \mathbf{y}(0))^2} - \frac{(r-k)(r-l-1)y_{0,k}\lambda(0, \mathbf{y}(0))}{s(0, \mathbf{y}(0))^2},\end{aligned} \tag{6.2.12}$$

so every summand of $\phi_{0,k}^{(0)}$ has a multiplicative factor of y_l for some $l \geq k$, or of λ, all of which are zero at $x = 0$. This establishes the first part.

For the second part, we have

$$\frac{d^2y_{0,k}}{dx^2}(0) = \frac{d}{dx}\sum_{l=0}^{r-2} \alpha_l^{(0)}(0, \mathbf{y}(0))\phi_{0,k}^{(l)}(0, \mathbf{y}(0)).$$

Consider expressing the derivative of the right-hand side of this equations using the sum and product rules. We have already seen that $\phi_{0,k}^{(l)}(0, \mathbf{y}(0)) = 0$ for $k \geq 1$ and $l \geq 0$, so the only possible nonzero summands involve derivatives of $\phi_{0,k}^{(l)}$. We also have $\alpha_l^{(0)}(0, \mathbf{y}(0)) = \delta_{0,l}$, so only the term multiplying $\dfrac{d\phi_{0,k}^{(0)}}{dx}(0)$ may be nonzero.

Now, we look at the definition of $\phi_{0,k}^{(0)}$ in (6.2.12). Consider its derivative obtained through the product and quotient rules. Every summand in this derivative that multiplies one of λ, $y_{0,l}$ or $\dfrac{dy_{0,l}}{dx}$ for some $l \geq 1$ vanishes at $x = 0$, since these three elements vanish at $x = 0$. Note that, for $k \geq 2$, all summands have this property, so the derivative of $\phi_{0,k}^{(0)}$ vanishes at $x = 0$ for $k \geq 2$. For $k = 1$, this same observation leads to

$$\frac{d^2y_{0,1}}{dx^2}(0) = \frac{d\phi_{0,1}^{(0)}}{dx} = \frac{r(r-1)y_{0,0}(0)D\lambda(0)}{s(0, \mathbf{y}(0))^2},$$

which concludes the base of induction.

A very similar argument is used for the induction step. Assume that the result is true for $1 \leq a-1 < r$ and consider, for $k \geq a$,

$$\frac{d^{2a-1}y_{0,k}}{dx^{2a-1}}(0) = \frac{d^{2a-2}}{dx^{2a-2}} \sum_{l=0}^{r-2} \alpha_l^{(0)}(0, \mathbf{y}(0))\phi_{0,k}^{(l)}(0, \mathbf{y}(0)).$$

In the definition of $\phi_{0,k}^{(l)}$, all the summands multiply $y_{0,k}$ or $y_{0,k-1}$. By induction, these two functions and their derivatives of order smaller than or equal to $2a-2$ vanish at $x = 0$, provided that $k > a$. In the case $k = a$, the only possibly non-vanishing summand in the expression obtained by calculating the right-hand side of the above equation using the sum, product and quotient rules multiplies $\dfrac{d^{2a-2}y_{a-1}}{dx^{2a-2}}$. Now, this term must come from the $(2a-2)$-nd derivative of $\phi_{0,a}^{(0)}$ at $x = 0$, and, as a consequence, it multiplies $\lambda(0, \mathbf{y}(0))$ in this derivative, which is equal to zero.

We may use the same argument to deduce that $\dfrac{d^{2a}y_{0,k}}{dx^{2a}}(0)$ vanishes if $k > a$. If $k = a$, the only nonzero term in the derivative of the right-hand side of

$$\frac{d^{2a}y_{0,k}}{dx^{2a}}(0) = \frac{d^{2a-1}}{dx^{2a-1}} \sum_{l=0}^{r-2} \alpha_l^{(0)}(0, \mathbf{y}(0))\phi_{0,k}^{(l)}(0, \mathbf{y}(0))$$

is

$$\begin{aligned}
\frac{d^{2a-1}\phi_{0,k}}{dx^{2a-1}}(0) &= (2a-1)\frac{(r-a+1)(r-1)D\lambda(0)}{s(0, \mathbf{y}(0))^2}\frac{d^{2a-2}y_{a-1}}{dx^{2a-2}}(0) \\
&= \frac{(2a-1)(r-a+1)(r-1)D\lambda(0)}{s(0, \mathbf{y}(0))^2}\frac{(2a-3)!!r!(r-1)^{a-1}D\lambda(0)^{a-1}}{(r-a+1)!s(0, \mathbf{y}(0))^{2a-2}} \\
&= \frac{(2a-1)!!\ r!\ (r-1)^a(D\lambda(0, \mathbf{y}(0)))^a}{(r-a)!\ s(0, \mathbf{y}(0))^{2a}},
\end{aligned}$$

concluding the proof of our claim. □

Now, to prove part (d) of Assumption 4.5.2, observe that

$$\begin{aligned}
\frac{dy_{0,0}}{dx}(0) &= \sum_{l=0}^{r-2} \alpha_l^{(0)}(0, \mathbf{y}(0))\phi_{0,0}^{(l)}(0, \mathbf{y}(0)) \\
&= \phi_{0,0}^{(0)}(0, \mathbf{y}(0)) \\
&= -r(r-1)y_{0,0}(0)/s^2 = -(r-1)/r < -\gamma,
\end{aligned}$$

as required (once again, $\gamma > 0$ is redefined if necessary).

As a consequence, Assumption 4.5.1 holds at $x = 0$. By the discussion in Section 4.5, we may extend our solution to (6.2.4) to all $x \in [0, x_0]$, where x_0 is defined as the infimum of all

$x > 0$ for which at least one of the following termination conditions hold:

$$\begin{aligned} &1.\quad \text{for some } k \in \{0,\dots,r-2\} \setminus A_0,\ \alpha_k^{(0)}(x,\mathbf{y}(x)) = 0;\\ &2.\quad \text{for some } (j,k) \in \mathcal{I} \setminus Y_0,\ y_{j,k}(x) = 0 \text{ or } y_{j,k}(x) = 1;\\ &3.\quad \det M^{(0)}(x,\mathbf{y}(x)) = 0;\\ &4.\quad \text{the solution is outside } \Omega_{\gamma,\mu,M},\ \text{or does not exist.} \end{aligned} \tag{6.2.13}$$

Also, if any termination condition other than

$$\alpha_k^{(0)}(x,\mathbf{y}(x)) = 0$$

is active at x_0, or if

$$\alpha_k^{(0)}(x_0,\mathbf{y}(x_0)) = 0 \text{ and } \frac{d\alpha_k^{(0)}}{dx}(x_0) \geq 0,$$

Phase 0 was defined to be the *final phase*.

In particular, if Phase 0 is not the final phase, we may redefine $\gamma > 0$ as a smaller positive number, if necessary, to ensure that

$$\begin{aligned} &\text{(i)}\quad \gamma < y_{j,k}(x_0) < 1-\gamma,\ \text{if } (j,k) \notin Y_0;\\ &\text{(ii)}\quad \frac{d\alpha_0^{(0)}(x,\mathbf{y}(x))}{dx}(x_0) \leq -\gamma. \end{aligned} \tag{6.2.14}$$

So, we must have

$$\begin{aligned} y_{j,k}(x_0) &= 0, && \text{if } (j,k) = (1,k) \text{ and } k \geq 1,\\ \gamma < y_{j,k}(x_0) &< 1-\gamma, && \text{otherwise,} \end{aligned} \tag{6.2.15}$$

where the second fact comes from (6.2.8).

We now show that, provided that Phase $t-1$ is not the final phase, Assumption 4.5.2 holds at the point $x = x_{t-1}$, the point in which Phase $t-1$ terminates. In other words, let $1 \leq t \leq r-2$ and inductively assume that x_{t-1} has been defined, that

$$\alpha_{t-1}^{(t-1)}(x_{t-1},\mathbf{y}(x_{t-1})) = 0$$

is the only active termination condition of Phase $t-1$ at x_{t-1} and that

$$\frac{d\alpha_{t-1}^{(t-1)}}{dx}(x_{t-1}) < 0.$$

We may also inductively assume that $\gamma > 0$ has been defined so that

$$\begin{aligned} &\text{(i)}\quad \gamma < y_{j,k}(x_{t-1}) < 1-\gamma,\ \text{if } (j,k) \notin Y_{t-1};\\ &\text{(ii)}\quad \frac{d\alpha_{t-1}^{(t-1)}}{dx}(x_{t-1}) \leq -\gamma, \end{aligned} \tag{6.2.16}$$

and that, as in equation (6.2.15),

$$\begin{array}{ll} y_{j,k}(x_{t-1}) = 0, & \text{if } (j,k) = (1,k) \text{ and } k \geq t, \\ \gamma < y_{j,k}(x_0) < 1-\gamma, & \text{otherwise.} \end{array} \tag{6.2.17}$$

In Phase t, the system of differential equations of interest is

$$\frac{dy_{j,k}}{dx} = \sum_{l=t}^{r-2} \alpha_l^{(t)}(x, \mathbf{y}(x)) \phi_{j,k}^{(l)}(x, \mathbf{y}(x)), \ (j,k) \in \mathcal{I}, \tag{6.2.18}$$

with initial conditions at $x = x_{t-1}$ given by the values of each $y_{j,k}$ at the termination of Phase $t-1$. Recall that the terms $\alpha_k^{(t)}$ in the previous equations are the solutions to the linear system (4.5.2), where the matrix $M^{(t)}(x, \mathbf{y})$ is given by $M(r-t-1, a_{t+1}, \ldots, a_{r-2})$, defined in (5.2.5), with a_k defined in (6.2.6).

Recall that the statement of Assumption 4.5.2 is as follows.

Assumption 4.5.3 ***(Assumptions for Phase t)*** *Consider the sets*

$$A_t = \left\{ k \in \{t, \ldots, r-2\} : \frac{d\alpha_k^{(t)}}{dx} \equiv 0 \right\} \quad \textit{and} \quad Y_t = \left\{ (j,k) \in \mathcal{I} : \frac{dy_{j,k}}{dx} \equiv 0 \right\}.$$

The system of differential equations (6.2.18) satisfies

(a') $|\det M^{(t)}(x_{t-1}, \mathbf{y}(x_{t-1}))| > \gamma$.

(b') $\alpha_t^{(t)}(x_{t-1}) > \gamma$ *and, for each* $k \in \{t, \ldots, r-2\} \setminus A_t$, *at least one of the following holds:*

(i) $\alpha_k^{(t)}(x_{t-1}) > \gamma$,

(ii) $0 \leq \alpha_k^{(t)}(x_{t-1})$, *the set* $A_{t,k} = \left\{ s \geq 1 \ : \ \frac{d^s \alpha_k^{(t)}}{dx^s}(x_{t-1}) \neq 0 \right\}$ *is not empty and is such that* $\frac{d^\nu \alpha_k^{(t)}}{dx^\nu}(x_{t-1}) > \gamma$, *where* $\nu = \min A_{t,k}$;

(c') *for each* $(j,k) \in \mathcal{I} \setminus Y_t$, *at least one of the following holds:*

(i) $y_{j,k}(x_{t-1}) > \gamma$,

(ii) $0 \leq y_{j,k}(x_{t-1})$, *the set* $Y_{t,j,k} = \left\{ s \geq 1 \ : \ \frac{d^s y_{j,k}}{dx^s}(x_{t-1}) \neq 0 \right\}$ *is not empty and is such that* $\frac{d^\nu y_{j,k}}{dx^\nu}(x_{t-1}) > \gamma$, *where* $\nu = \min Y_{t,j,k}$;

(d') *for every* $(j,k) \in \mathcal{I}$, $y_{j,k}(x_{t-1}) < 1-\gamma$,

Using the formula in Proposition B.1.6, we obtain

$$|\det M^{(t)}(x_{t-1}, \mathbf{y}(x_{t-1}))| = 1 + \sum_{k=t+1}^{r-2} (k-t) a_k(x_{t-1}, \mathbf{y}(x_{t-1})) \geq 1 > \gamma,$$

establishing part (a') of the assumption. Also,

$$\alpha_k^{(t)}(x,\mathbf{y}(x)) = \begin{cases} \dfrac{1-\sum_{l=t+1}^{r-2}(r-l-1)a_l}{1+\sum_{l=t+1}^{r-2}(l-t)a_l}, & \text{if } k=t \\ \dfrac{(r-t-1)a_k}{1+\sum_{l=t+1}^{r-2}(l-t)a_l}, & \text{if } k>t. \end{cases} \tag{6.2.19}$$

The definition of $\alpha^{(t)}$ is such that

$$\frac{dy_{1,k}}{dx} \equiv 0 \text{ if } k > t,$$

so $\{(1,t+1),\ldots,(1,r-2)\} \subseteq Y_t$, with Y_t defined in the statement of Assumption 4.5.2, and

$$y_{1,k}(x) = 0 \text{ for every } k \in \{t+1,\ldots,r-2\}, \tag{6.2.20}$$

for every x in the interval of definition of the solution to (6.2.18). Using (6.2.6), this implies that

$$a_{t+1}(x,\mathbf{y}(x)) = (r-t-1)y_{0,t}(x)/s(x,\mathbf{y}(x)) + (r-t-1)y_{1,t-1}(x)\lambda(x,\mathbf{y}(x))/s(x,\mathbf{y}(x))^2 \tag{6.2.21}$$

and that

$$a_k(x,\mathbf{y}(x)) = (r-k)y_{0,k}/s(x,\mathbf{y}(x)), \text{ if } t+1 < k \leq r-2. \tag{6.2.22}$$

We now verify parts (b') and (c') of Assumption 4.5.2. For the latter, we have already seen that $\{(1,t+1),\ldots,(1,r-2)\} \subseteq Y_t$. Now, by our choice of γ in equation (6.2.16), we have that $y_{j,k}(x_{t-1}) > \gamma$ for the case $j=0$, the case $j=1$ and $k<t$, and the case $j=1$ and $k=r-1$. Hence, the only remaining case to verify for part (c') is $j=1$ and $k=t$.

Before analysing this case, we look at part (b') of the assumption. By equations (6.2.21) and (6.2.22), we have

$$a_k(x_{t-1},\mathbf{y}(x_{t-1})) \geq (r-k)y_{0,k}(x_{t-1})/s(x_{t-1},\mathbf{y}(x_{t-1})),$$

which is positive by 6.2.15, for every $k \geq t$. So, by (6.2.19), item (i) of part (b') of Assumption 4.5.2 holds for $\alpha_k^{(t)}$, $k=t+1,\ldots,r-2$, possibly by decreasing the value of γ. To conclude the verification of part (b'), we consider the case $k=t$. The numerator of $\alpha_t^{(t)}$ in equation (6.2.19) is given by

$$1-\sum_{l=t+1}^{r-2}(r-l-1)a_l = \left(1-\sum_{l=t}^{r-2}(r-l-1)a_l\right) + (r-t-1)a_t,$$

where the first summand on the right-hand side is the numerator $\alpha_{t-1}^{(t-1)}$, which vanishes at $x=x_{t-1}$ by the termination condition of Phase $t-1$. Since

$$a_t(x_{t-1},\mathbf{y}(x_{t-1})) = \frac{(r-t)y_{0,t}(x_{t-1})}{s(x_{t-1},\mathbf{y}(x_{t-1}))} + \frac{(r-t)y_{1,t-1}(x_{t-1})\lambda(x_{t-1},y(x_{t-1}))}{s(x_{t-1},\mathbf{y}(x_{t-1}))^2} > 0,$$

part (b') also holds in this case.

Recall that, to end the verification of part (c') of the assumption, we only have to consider the case $(j,k) = (1,t)$. By (6.2.15), we know that $y_{1,t}(x_{t-1}) = 0$. We shall show that the derivative of this function also vanishes at $x = x_{t-1}$, but the second derivative is positive at this point, which implies that item (ii) of part (c') holds in this case.

We start by calculating the first derivative of $y_{1,t}$. By equation (6.2.18),

$$\frac{dy_{1,t}}{dx} = \sum_{k=t}^{r-2} \alpha_k^{(t)} \phi_{1,t}^{(k)}. \tag{6.2.23}$$

Now, equation (6.2.19) implies that

$$\begin{aligned}\alpha_t^{(t)}(x_{t-1}, \mathbf{y}(x_{t-1})) &= \frac{1 - \sum_{l=t+1}^{r-2}(r-l-1)a_l(x_{t-1}, y(x_{t-1}))}{1 + \sum_{l=t+1}^{r-2}(l-t)a_l(x_{t-1}, y(x_{t-1}))} \\ &= \frac{(r-t-1)a_t(x_{t-1}, y(x_{t-1}))}{1 + \sum_{l=t+1}^{r-2}(l-t)a_l(x_{t-1}, y(x_{t-1}))},\end{aligned}$$

where the last step uses the fact that the quantity

$$1 - \sum_{l=t}^{r-2}(r-l-1)a_l(x_{t-1}, y(x_{t-1}))$$

is equal to the numerator of $\alpha_{t-1}^{(t-1)}(x_{t-1}, \mathbf{y}(x_{t-1}))$, which is equal to zero by the termination condition of Phase $t-1$. By the definitions of $\phi_{j,k}^{(t)}$ and of a_k, given in (6.2.1) and (6.2.6), respectively, we have

$$\phi_{1,t}^{(t)} = -1 + (r-t-1)a_t \text{ and } \phi_{1,t}^{(k)} = (r-k-1)a_t.$$

So, if we multiply the right-hand side of equation (6.2.23) by the positive quantity

$$1 + \sum_{l=t+1}^{r-2}(t-l)a_l,$$

we obtain

$$\begin{aligned}&(r-t-1)a_t\left(-1 + (r-t-1)a_t\right) + \sum_{k=t+1}^{r-2}(r-t-1)a_k(r-k-1)a_t \\ &= -(r-t-1)a_t\left(1 - \sum_{k=t}^{r-2}(r-k-1)a_k\right).\end{aligned}$$

But the term within brackets is precisely the numerator of $\alpha_{t-1}^{(t-1)}$, hence the first derivative of $y_{1,t}$ vanishes at x_{t-1}, as claimed.

We now calculate the second derivative of $y_{1,t}$, given by

$$\begin{aligned}\frac{d^2 y_{1,t}}{dx^2} &= \frac{d}{dx}\sum_{k=t}^{r-2}\alpha_k^{(t)}\phi_{1,t}^{(k)} \\ &= \sum_{k=t}^{r-2}\frac{d\alpha_k}{dx}\phi_{1,t}^{(k)} + \alpha_k^{(t)}\frac{d\phi_t^{(k)}}{dx}.\end{aligned} \tag{6.2.24}$$

Since our calculations will be somewhat more involved, we omit references to the point $(x_{t-1}, \mathbf{y}(x_{t-1}))$ from the equations. However, we shall consider the functions applied at this point. To have a more compact notation, we shall also use h' to denote the derivative $\frac{dh}{dx}$. Finally, each summand in the right-hand side of equation (6.2.24) will be referred as a term in the expansion of $\frac{d^2 y_{1,t}}{dx^2}$.

We now calculate the value of different types of terms in this expansion. Again, we use the definition of $\alpha^{(t)}$ in (6.2.19) and the identities

$$\phi_{1,t}^{(t)} = -1 + (r-t-1)a_t \text{ and } \phi_{1,k}^{(t)} = (r-k-1)a_t,$$

so that

$$\begin{aligned}\alpha_t^{(t)\prime}\phi_{1,t}^{(t)} &= \left(\frac{-(r-t-1)a_t}{D^2}\sum_{l=t+1}^{r-2}(l-t)a_l' - \frac{1}{D}\sum_{l=t+1}^{r-2}(r-l-1)a_l'\right) \\ &\quad \times \left(-1 + (r-t-1)a_t\right), \\ \alpha_k^{(t)\prime}\phi_{1,t}^{(k)} &= \left(\frac{-(r-t-1)a_k}{D^2}\sum_{l=t+1}^{r-2}(l-t)a_l' + \frac{(r-t-1)a_k'}{D}\right) \\ &\quad \times (r-k-1)a_t, \text{ if } t < k \leq r-2, \\ \alpha_k^{(t)}\phi_{1,k}^{(t)\prime} &= \frac{(r-t-1)a_k}{D}(r-k-1)a_t', \ t \leq k \leq r-2,\end{aligned} \tag{6.2.25}$$

where

$$D = 1 + \sum_{l=t+1}^{r-2}(l-t)a_l.$$

The last part holds in the case $k = t$ because the numerator of $\alpha_t^{(t)}$ is the sum of the numerator of $\alpha_{t-1}^{(t-1)}$, which vanishes at $x = x_{t-1}$, and the term $(r-t-1)a_t$.

If we sum the numerators of the terms above with quadratic denominator, we obtain

$$(r-t-1)a_t\sum_{l=t+1}^{r-2}(l-t)a_l'\left(1 - (r-t-1)a_t - \sum_{k=t+1}^{r-2}(r-k-1)a_k\right) = 0,$$

once again by the fact that $\alpha_{t-1}(x_{t-1}, \mathbf{y}(x_{t-1})) = 0$. As a consequence, the terms involving derivatives of $\alpha^{(t)}$ sum to

$$\sum_{l=t+1}^{r-2}(r-l-1)a_l'/D, \tag{6.2.26}$$

while the terms involving derivatives of $\phi^{(t)}$ in (6.2.24) sum to

$$(r-t-1)a_t'\sum_{k=t}^{r-2}(r-k-1)a_k/D = (r-t-1)a_t'/D. \tag{6.2.27}$$

Putting everything together, we obtain

$$\frac{d^2y_{1,t}}{dx^2}(x_{t-1}) = \sum_{l=t}^{r-2}(r-l-1)a_l'/D.$$

But note that

$$\frac{d\alpha_{t-1}^{(t-1)}}{dx}(x_{t-1}) = -\frac{\displaystyle\sum_{l=t}^{r-2}(r-l-1)a_l'}{1+\displaystyle\sum_{l=t}^{r-2}(l-t-1)a_l} < 0,$$

from which we deduce that the second derivative of $y_{1,t}$ is positive at x_{t-1}, as claimed.

Observe that, to obtain the results about the derivatives of $y_{1,t}$ at x_{t-1}, we only used the definitions of a_k and $\phi_{1,t}^{(k)}$, the fact that $\alpha_{t-1}(x_{t-1}, \mathbf{y}(x_{t-1})) = 0$ and the fact that $y_{1,k}(x_{t-1}) = 0$ if $t \le k \le r-2$.

Finally, to conclude the verification of Assumption 4.5.2, we note that part (d') is satisfied, since

$$y_{j,k}(x_{t-1}) < 1-\gamma \text{ for every } (j,k) \in \mathcal{I},$$

due to our choice of γ in (6.2.14). If Phase t is not the final phase, we may again redefine $\gamma > 0$ as a smaller positive number, if necessary, to ensure that

$$\text{(i)} \quad \gamma < y_{j,k}(x_t) < 1-\gamma, \text{ if } (j,k) \notin Y_t;$$

$$\text{(ii)} \quad \frac{d\alpha_t^{(t)}}{dx}(x_t) \le -\gamma.$$

Moreover, in the above calculations, we established that

$$\begin{aligned} y_{j,k}(x_t) &= 0, && \text{if } (j,k) = (1,k) \text{ and } k \ge t+1 \\ \gamma < y_{j,k}(x_0) &< 1-\gamma, && \text{otherwise.} \end{aligned}$$

validating the induction hypothesis in (6.2.17).

Again, γ is only redefined a finite number of times in this discussion, always being assigned a positive value, hence there exists $\gamma > 0$ for which the assumptions hold at the points $x_0 = 0$ and x_{t-1}, $t > 1$, as long as the only termination condition for Phase $t-1$ is $\alpha_{t-1}^{(t-1)}(x_{t-1}, y(x_{t-1})) = 0$, with the first derivative of $\alpha_{t-1}^{(t-1)}$ being negative at this point. Let b be the index of the final phase obtained by this method for such a value of $\gamma > 0$. As in the case of independent sets, we could potentially have $b(\gamma') < b(\gamma)$ if $\gamma' < \gamma$. However, the limit

$$b = \lim_{\gamma \to 0^+} b(\gamma), \tag{6.2.28}$$

is well defined, since $b(\gamma)$ is bounded above by $r-2$ and is non-decreasing as γ decreases. Moreover, since $b(\gamma)$ is integer-valued, we know that this limit is achieved for γ sufficiently small. For this value of b, we may also define

$$\xi(r) = \lim_{\gamma \to 0^+} x_b^0(\gamma) + y_{0,r}(x_b^0(\gamma)) + y_{1,r-1}(x_b^0(\gamma)), \tag{6.2.29}$$

where $y_{0,r}(x)$ and $y_{1,r-1}(x)$ are components of the solution to the system of differential equations obtained here. It is easy to see that this limit exists, since all the terms in this sum are non-decreasing as γ increases and they are bounded above by 1.

Second part As for independent sets, this argument provides a solution to the system of differential equations of interest, in this case (6.1.8), in the situation when $p_0 = 0$. To extend the result to $p_0 \in (0, p_0']$ for some $p_0' > 0$, we again use Lemma 2.5.1 and Lemma 5.2.1. Some of the details are omitted, since the argument used is very similar to the one in the case of independent sets.

Let $p_0 > 0$ and consider the system of differential equations

$$\begin{aligned} &\frac{dy_{j,k}}{dx} = \sum_{l=0}^{r-2} \alpha_l^{(0)}(x, \mathbf{y}(x)) \phi_{j,k}^{(l)}(x, \mathbf{y}(x)), \ (j,k) \in \mathcal{I}, \\ &y_{0,0}(0) = 1, \ y_{j,k}(0) = 0 \text{ if } (j,k) \neq (0,0), \end{aligned} \tag{6.2.30}$$

which is given in Definition 6.1.1 with this value of p_0 and $t = 0$. Also consider the region $\Omega_{\gamma,\mu,M}$, where γ, μ and M are defined as in the case $t = 0$.

As in the case of independent sets, the initial conditions $w_{0,k}(p_0)$ are continuous as functions of p_0, and it is clear that, if p_0 is sufficiently small, the point $(0, \mathbf{w}(0))$ lies in the interior of $\Omega_{\gamma,\mu,M}$, since $w_{0,k}(0)$ lies in the interior of this region. By Lemma 2.5.1, we may again uniquely extend the solution $\mathbf{y}^{p_0}$ to (6.2.30) arbitrarily close to the boundary of $\Omega_{\gamma,\mu,M}$.

To show that this solution can be extended so as to satisfy the properties in Property $(b, p_0', \mathbf{x})$, we now verify that, if $p_0 > 0$ is sufficiently small, Assumption 4.5.1, with the probability p_0 changed from 0 to a small positive constant, also holds for $\mathbf{y}^{p_0}$ at $x = 0$. The continuity of the initial conditions immediately implies that this is the case for part (a), for item (i) of parts (b) and (c), and for part (d). So, the only parts of Assumption 4.5.1 that require additional discussion involve the definition of the indices in Y_1 and A_1 or item (ii) in parts (b) and (c). In the analysis of the case $p_0 = 0$, these assumptions only appear in the verification of the following facts:

1. $$\frac{dy_{1,k}^0}{dx} \equiv 0, \text{ if } 1 \leq k \leq r-2,$$

2. $$y_{1,0}^0(0) = 0 \text{ and } \frac{dy_{1,0}^0}{dx}(0) > \gamma,$$

3.
$$\frac{d^\nu y^0_{0,k}}{dx^\nu}(0) = 0, \text{ if } \nu < 2k, \text{ and } \frac{d^{2k} y^0_{0,k}}{dx^{2k}}(0) > \gamma, \text{ if } k \geq 1,$$

4.
$$\frac{d^\nu \alpha^{(0)}_k}{dx^\nu}(0) = 0, \text{ if } \nu < 2k, \text{ and } \frac{d^{2k} \alpha^{(0)}_k}{dx^{2k}}(0) > \gamma, \text{ if } k \geq 1,$$

5.
$$\frac{d^\nu y^0_{1,r-1}}{dx^\nu}(0) = 0, \text{ if } \nu < 2r-1, \text{ and } \frac{d^{2r-1} y^0_{0,k}}{dx^{2r-1}}(0) > \gamma.$$

For $p_0 > 0$, since the definition of $\alpha^{(t)}_k$ is still such that

$$\frac{dy^{p_0}_{1,k}}{dx} \equiv 0 \text{ if } k \in \{2, \ldots, r-2\},$$

item 1 clearly holds in this case. Item 2 is also easy. Indeed, we now have $y^{p_0}_{1,0}(0) > 0$, and, by continuity of the initial conditions, we still have

$$\frac{dy^{p_0}_{1,0}}{dx}(0) > \gamma.$$

However, we need to be more careful in verifying that the remaining items are also satisfied. This is because, for $p_0 > 0$, the terms $\alpha^{(0)}_k(0, \mathbf{y}^{p_0}(0))$ are not equal to zero. For instance, Claim 6.2.1 does not hold for $\mathbf{y}^{p_0}$. However, by the same type of analysis in the proof of that result, we can show that, for $k \geq 1$ and $s < 2k$, the terms

$$\frac{d^s y^{p_0}_{0,k}}{dx^s}(0)$$

can be shown to be polynomials in the variable p_0 with minimum degree p_0^{2k-s} and such that the coefficient of the monomial with minimum degree is positive. In particular, for p_0 sufficiently small, the derivatives of $y^{p_0}_{0,k}$ of order smaller than $2k$ are all positive, but cannot be bounded below by a constant, since the lower bounds depend on p_0. However, by continuity of the initial conditions, we derive that the derivative of order $2k$ of $y^{p_0}_{0,k}$ is bounded below by a positive constant independent of p_0. This verifies item 3 when p_0 is a small positive integer. By definition of a_k in (6.2.6) and of $\alpha^{(0)}_k$ in (6.2.7), it is clear that item 4 also holds, whereas item 5 may be obtained by the same argument.

As with independent sets, we may now argue that, for $p'_0 > 0$ sufficiently small and $0 < p_0 \leq p'_0$, the solution to the system of differential equations (6.2.30) can be extended to a point $x^{p_0}_0$ that tends to x^0_0 as p_0 tends to 0. The additional claim that $y^{p_0}_{j,k}$ uniformly approximates $y^0_{j,k}$ is an easy consequence of Lemma 5.2.1.

The behaviour of $\mathbf{y}^{p_0}$ in the remaining phases can be analysed similarly. That is, we again derive it from the information we have about the solution in the case $p_0 = 0$, using Lemma 2.5.1 and Lemma 5.2.1. Note that several of the conditions are easier to verify, since we now

have $\mathbf{y}^0_{0,k}(x_{t-1}) > \gamma$ and $\alpha^{(t)}_k(x_{t-1}, \mathbf{y}(x_{t-1}) > \gamma$, so that an argument using Claim 6.2.1 is not needed. Finally, the case that was different for independent sets, namely the extension to $p_0 > 0$ of the formula

$$\frac{d^\nu y^0_{1,t}}{dx^\nu}(x_{t-1}) = 0, \text{ if } \nu \leq 1 \text{ and } \frac{d^2 y^0_{1,t}}{dx^2}(x_{t-1}) > \gamma,$$

can be easily treated here. This is because the calculations leading to this fact in the case $p_0 = 0$ did not use any properties of the solutions, but only the definitions of $\phi^{(t)}_{j,k}$, $\alpha^{(t)}_k$ and a_k, and the assumption that Phase $t-1$ was not the final phase. As a consequence, the same calculations are valid in the case $p_0 > 0$. All the other cases can be resolved easily, and therefore are omitted.

In particular, for p_0' sufficiently small and $p_0 \in [0, p_0']$, we find functions $\mathbf{y}^{p_0}(x) = (y^{p_0}_{j,k}(x))_{(j,k)\in\mathcal{I}}$ such that, for $t \in \{0, \ldots, b\}$,

$$\begin{aligned} &\frac{dy^{p_0}_{j,k}}{dx} = F^{(t)}_{j,k}(x, \mathbf{y}^{p_0}(x)) \text{ for } x \in [x^{p_0}_{t-1}, x^{p_0}_t),\ k = 0, \ldots, r, \\ &y^{p_0}_{j,k}(x^{p_0}_{t-1}) = \beta^{p_0}_{t-1,j,k}, \end{aligned}$$

where the initial conditions $\beta^{p_0}_{0,k}$ are defined as in Definition 6.1.1, and the constants $x^{p_0}_0 = 0 < \cdots < x^{p_0}_b$ are defined through the termination conditions.

Now, for $p_0 > 0$, all the components of $\mathbf{y}^{p_0}(0)$ are strictly positive. Moreover, for $p_0 \leq p_0'$, none of the termination conditions $y^{p_0}_{j,k}(x) = 0$, $(j,k) \in \mathcal{I}$ hold for x in the interval $[0, x^{p_0}_b]$, except possibly at $x = x^{p_0}_b$. Thus, for any $\xi > 0$,

$$\gamma^{p_0}_\xi = \inf\{y_{j,k}(x) \,:\, x \in [0, x^{p_0}_b - \xi],\ (j,k) \in \mathcal{I}\} > 0.$$

With these values of b, p_0' and $x^{p_0}_t$, we may define functions

$$x_0, \ldots, x_b : [0, p_0'] \to [0, 1]$$

by setting

$$x_t(p_0) = x^{p_0}_t, \text{ for every } t \in \{0, \ldots, b\} \text{ and every } p_0 \in [0, p_0'].$$

By the work done above, it is clear that, for these values of b and p_0', and these functions $x_0, \ldots, x_b$, the system of differential equations (6.1.6) satisfies Property $(b, p_0', \mathbf{x})$.

6.3 A lower bound on a largest induced forest

As in the case of independent sets, we combine the results of the last two sections to give new lower bounds on $\tau(G)$, the number of vertices in a largest induced forest of a graph G, when G is a regular graph with sufficiently large girth.

In the previous section, we proved that there exists a positive integer $b \in \{1, \ldots, r-1\}$, a constant $p_0' \in (0, 1]$ and functions

$$x_1, \ldots, x_b : [0, p_0'] \to [0, 1]$$

for which the system of differential equations in (6.1.6) satisfies Property $(b, p_0', \mathbf{x})$. To obtain these values, we first showed that exists $\gamma > 0$ for which the solution of (6.1.8) with initial conditions given with $p_0 = 0$ could be extended to a point $x_b(0)$, with the points of phase transition $x_0(0), \ldots, x_{b-1}(0)$ being determined by termination conditions. We then proved that, with this value of γ, we could find a constant $p_0' > 0$ and functions $x_1, \ldots, x_b : [0, p_0'] \to [0, 1]$ for which Property $(b, p_0', \mathbf{x})$ holds.

For each $r \geq 3$, consider the positive integer $b \in \{0, \ldots, r-2\}$ defined in (6.2.28) and the constant $\xi(r)$ defined in (6.2.29). We have the following theorem.

Theorem 6.3.1 *Let $\delta > 0$ and $r \geq 3$. Then, there exists $g > 0$ such that every r-regular graph G on n vertices with girth greater than or equal to g satisfies $\tau(G) \geq (\xi(r) - \delta)n$.*

Proof Let $r \geq 3$ and let $\delta' > 0$. Let $\gamma > 0$ be such that, in the definition of $\xi(r)$, we have

$$|x_b(0) + y_{0,r}(x_b(0)) + y_{1,r-1}(x_b(0)) - \xi(r)| \leq \frac{\delta'}{2}.$$

Consider $p_0' \in (0, 1]$ and the functions

$$x_0, \ldots, x_b : [0, p_0'] \to [0, 1],$$

as defined in the previous section, for which the system of differential equations in (6.1.6) satisfies Property $(b, p_0', \mathbf{x})$. By Theorem 6.1.2, there exists $g > 0$ such that every r-regular graph G on n vertices with girth greater than or equal to g satisfies

$$\tau(G) \geq \left(x_b(0) + y_{0,r}(x_b(0)) + y_{1,r-1}(x_b(0)) - \frac{\delta'}{2} \right) n.$$

By our choice of γ, this leads to

$$\tau(G) \geq (\xi(r) - \delta')n,$$

as required. ∎

We have shown in Chapter 2 that the property of inducing a forest in a graph is vertex monotone. In particular, by Lemma 2.1.2, the bounds for r-regular graphs in Theorem 6.3.1 can be extended to graphs with maximum degree r.

Corollary 6.3.2 *Let $\delta > 0$ and $r \geq 3$. Then there exists $g > 0$ such that every graph G on n vertices with maximum degree r and girth greater than or equal to g satisfies $\tau(G) \geq (\xi(r) - \delta)n$.*

Table 6.3.1: Lower and upper bounds on $\tau(G)/n$

r	$\xi(r)$	$\xi_1(r)$	$\Xi(r)$
3	0.7368	0.7268	0.7500
4	0.6351	0.6045	0.6667
5	0.5662	0.5269	0.6216
6	0.5149	0.4711	0.5776
7	0.4746	0.4283	0.5403
8	0.4415	0.3940	0.5086
9	0.4137	0.3658	0.4811
10	0.3898	0.3419	0.4570

Moreover, by Lemma 2.1.4, the numbers $\xi(r)$ are also asymptotic bounds for random regular graphs.

Corollary 6.3.3 *Let $\delta > 0$ and $r \geq 3$. Then a random r-regular graph G on n vertices a.a.s satisfies $\tau(G) \geq (\xi(r) - \delta)n$.*

Numerical approximations of the numbers $\xi(r)$, given in Table 6.3.1, have been obtained by numerically solving the system of differential equations (6.1.8) with initial conditions given with $p_0 = 0$. The method used in this computations is the same used for independent sets, so we refer to Section 5.3 for a discussion.

The numbers $\xi_1(r)$ and $\Xi(r)$ in the table are the lower bounds obtained in Chapter 2 and the upper bounds found in [7]. The numbers $\xi(r)$ improve the bounds previously known for all values of $r \geq 4$ for which they were calculated. (If $r = 3$, the best lower bound is $\tau(G) \geq 0.75n$ obtained in [24].) They are also new best asymptotic bounds for random regular graphs, improving, for all $r \geq 4$, the bounds given in [7].

Chapter 7

Conclusion

In this thesis, we have analysed the class of locally greedy algorithms, a class of probabilistic algorithms in regular graphs for which the random selection of a vertex v of the input graph is determined by the current state of the vertices within some fixed distance of v. This allows us to consider algorithms for which the different operations processed are ranked, so that some are given priority over others.

If the parameters d and ℓ in the class of algorithms are given by $d = 1$ and $\ell = 1$, the output of the algorithm is an independent set, while, with parameters $d = 1$ and $\ell = 2$, the algorithm produces a set inducing an acyclic subgraph of the input graph. This has been used to derive new lower bounds on the independence number $\alpha(G)$ of r-regular graphs with sufficiently large girth and on the the number $\tau(G)$ of vertices in a largest induced forest of G. These bounds are defined through the solution to the systems of ordinary differential equations defined in (5.1.7) and (6.1.6), respectively. For a few fixed values of r, numerical approximations of these lower were shown in Table 5.3.1 and in Table 6.3.1.

In the case of independent sets, the new bounds improve the previous best bounds for all values of r for which they were calculated. In Corollary 5.3.2, we extended these bounds to the more general case of graphs with large girth and maximum degree r. We also translated them into asymptotic bounds for random r-regular graphs. Even if, in this particular case, the numbers obtained here do not improve the bounds first given by Wormald in [56], our work provides a new proof of the validity of these bounds, the first proof of this result that does not rely on sharp concentration results.

For induced forests, the new bounds can again be extended to regular graphs with maximum degree r, and they improve the previous best bounds for all particular values of $r \geq 4$ for which they were determined. Furthermore, unlike in the case of independent sets, the associated bounds for random r-regular also improve the previous best bounds, which were given in [7], for every $r \geq 4$.

It is clear that the same method, or modifications thereof, could be applied to several

other graph problems. One such problem involves independent dominating sets, which we now define. In a graph $G = (V, E)$, a set $S \subseteq V$ is *k-dominating* if every vertex that does not lie in S is at distance at most k of a vertex in S. When $k = 1$, such a set is called a *dominating set*. An *independent dominating set* S in the graph is now a set that is both dominating and independent. Note that these are precisely the maximal independent sets in the graph, that is, the sets of vertices in the graph that are independent, but such that the addition of any new vertex would create a set that is not independent. The associated optimisation problem consists of determining the size of an independent dominating set of minimum cardinality in the graph.

Recall that a locally greedy algorithm with parameters $\ell = 1$ and $d = 1$ produces an independent set in the input graph. Moreover, when it is applied to find a large independent set in a graph, as with the instance introduced in Chapter 5, it actually finds a maximal independent set in a clever way, so as to include as many vertices as possible. Thus, if instead the probabilities were adjusted so as to generate a maximal independent set with as few vertices as possible, we may derive good upper bounds on the size of a smallest independent dominating set in the graph. Here, we should mention that the output of the algorithm is not exactly a maximal independent set, since there will still be a few white vertices at the end of the algorithm. This is due to the restriction on the number of steps performed by it. However, the solutions of the differential equations also give us the expected proportion of white vertices at the end of the algorithm. In particular, an upper bound on the proportion of vertices in a minimum independent dominating set is given by the sum of the proportion of purple vertices produced by the algorithm and the proportion of white vertices that remain at the end of the algorithm.

Another concept of vertex domination is as follows. A subset S of vertices is *k-fold-dominating* if each vertex not in S is joined to at least k vertices in S. This definition was proposed by Dai and Wu in [18], who referred to it as k-dominating, motivated by the study of wireless networks. In such a network, an important problem is to find a set of nodes to form a backbone that supports routing, which corresponds to a dominating set in the graph modelling this network. An efficient backbone is given by a small dominating set. Now, if dominating sets are replaced by k-fold-dominating sets, the backbone produced balances efficiency and fault tolerance. Again, the efficiency of the backbone is measured by its size, with smaller backbones being more efficient.

We now observe that the class of locally greedy algorithms may also be used to provide a small k-fold-dominating set in an r-regular graph. Indeed, if we fix $d = 1$ and $\ell = k$, the white vertices are vertices with less than k purple neighbours, that is, they are the vertices in the graph that are not k-dominated by the set of purple vertices. So, by assigning probabilities in such a way that there are "few" white vertices at the end of the algorithm, and, at the same time, keeping the set of purple vertices "small", we may find a good upper bound on

the cardinality of a smallest k-fold-dominating set in the graph.

Note that, to obtain the bounds suggested for the above problems, it suffices to determine the probabilities conveniently and to solve the differential equations associated with this choice of probability, since the remainder is covered by this thesis. To be precise, we would also have to obtain the recurrence equations when the parameters d and ℓ satisfy $d = 1$ and $\ell > 2$, but this can be easily done with arguments as in Section 3.2.

On the other hand, more work would be required to use locally greedy algorithms in applications for which the parameter d is larger than 1, but we believe that this can be done. Recall that the results in Chapter 2 are all proved for general d and, although this is not the case for Chapter 4, the results in that section can be easily generalised to $d > 1$. The real difficulty lies in obtaining the recurrence equations for $w_{i,j,k}$, since, as the choice of a vertex may now directly affect a vertex at distance d of it, in addition to the results involving independence of branches around white vertices, given by Corollaries 3.1.12, 3.1.13 and 3.1.14, we would also need information about the distribution of white vertices around yellow and purple neighbours to determine whether a white vertex v has been affected by the selection of a vertex not adjacent to v.

Generalisations of locally greedy algorithms may also be obtained by considering more classes of vertices. For instance, in the simple algorithm of Chapter 2, in addition to white, purple and yellow vertices, we considered an additional set of blue vertices. We have seen later that, by analysing this algorithm differently, we could group the blue and white vertices of the original algorithm in a single class. Nevertheless, this approach could be useful for other purposes. For instance, we could devise a class of algorithms for dominating sets as follows: purple vertices correspond to vertices in the dominating set, blue vertices correspond to the non-purple vertices dominated by purple (that is, with at least one purple neighbour) and white vertices correspond to the vertices not dominated by purple. At each step of the algorithm, white and blue vertices are randomly chosen with a probability determined by the number of neighbours of each colour. Now, the goal of the algorithm is to produce a "small" set of purple vertices in such a way that only a few white vertices remain, which would lead to a bound on the size of a minimum dominating set. The analysis of the expected performance of such an algorithm should strongly resemble the work done in this thesis. A lot of additional problem-oriented generalisations could be analysed by a similar method.

Moreover, Wormald and the author of this thesis are currently developing an alternative method to the study of properties of graphs with large girth, combining the approach of this thesis with other ideas. This work is motivated by [58], which analyses a quite general class of algorithms in random regular graphs. We believe that a general result concerning this class of algorithms can be obtained for graphs with sufficiently large girth. This is a very promising approach, since it should cover the dominating set problem and a number of other results in the literature.

Bibliography

[1] M. Ajtai, J. Komlós and E. Szemerédi, A dense infinite Sidon sequence. *European Journal of Combinatorics* 2:1–15, 1981.

[2] N. Alon, J. Kahn and P.D Seymour, Large induced degenerate subgraphs in graphs. *Graphs and Combinatorics* 3:203–211, 1987.

[3] N. Alon, D. Mubayi and R. Thomas, Large induced forests in sparse graphs. *Journal of Graph Theory* 38:113–123, 2001.

[4] N. Alon and J. Spencer, *The Probabilistic Method,* John Wiley and Sons, 2000.

[5] G. Avondo-Bodeno, *Economic applications of the theory of graphs,* Gordon and Breach Science, New York.

[6] S. Bau, L.W. Beineke, G.M. Du, Z.S. Liu, and R.C. Vandell, Decycling cubes and grids, *Utilitas Math.*, 59:129–137, 2001.

[7] S. Bau, N.C. Wormald, and S. Zhou, Decycling number of random regular graphs, *Random Structures and Algorithms*, 21:397–413, 2002.

[8] L.W. Beineke and R.C. Vandell, Decycling graphs, *J. Graph Theory*, 25:59–77, 1997.

[9] M. Beis, W. Duckworth and M. Zito, Large k-independent sets of random regular graphs, *Electronic Notes in Discrete Mathematics*, 19:321–327, 2005. Proceedings of the 2nd Brazilian Symposium on Graphs, Algorithms and Combinatorics. Angra dos Reis (Rio de Janeiro), Brazil, April 2005.

[10] I. Benjamini, C. Hoppen, E. Ofek, P. Prałat and N. Wormald, Geodesics and almost geodesic cycles in random regular graphs, preprint.

[11] C. Berge, *Graphes et hypergraphes,* Dunod, Paris, 1970.

[12] N. Biggs, Constructions for cubic graphs with large girth, *Electronic Journal of Combinatorics* 5(A1):25pp, 1998.

[13] B. Bollobás, A probabilistic proof of an asymptotic formula for the number of labelled regular graphs, *European Journal of Combinatorics*, 1:311–316, 1980.

[14] B. Bollobás, *Extremal Graph Theory,* Academic Press, 1978.

[15] S. Butenko, P. Pardalos, I. Sergienko, V. Shylo and P. Stetsyuk, Finding maximum independent sets in graphs arising from coding theory, *Proceedings of the 2002 ACM symposium on Applied computing,* Madrid, 542–546, 2002.

[16] S. Chastel, Querying for different similar images using maximum independent sets of a graph, *4th EURASIP Conference focused on Video/Image Processing and Multimedia Communications, 2003,* 1:149–154, 2003.

[17] D. M. Cvetković, M. Doob, and H. Sachs, *Spectra of Graphs,* Academic Press Inc., New York, 1980.

[18] F. Dai and J. Wu, On constructing k-connected k-dominating sets in wireless networks, accepted to *Journal of Parallel and Distributed Computing.*

[19] G. Davidoff, P. Sarnak, and A. Valette, *Elementary Number Theory, Group Theory and Ramanujan Graphs* Cambridge University Press, 2003.

[20] J. Díaz and M. Serna and N.C. Wormald, Computation of the bisection width for random d-regular graphs, *Lecture Notes in Computer Science, LATIN 2004: Theoretical informatics* 49–58, 2004.

[21] R. Diestel, *Graph Theory,* Springer Verlag, New York, 1980.

[22] W. Duckworth and N. C. Wormald, On the independent domination number of random regular graphs, *Combinatorics, Probability and Computing*, 15:513–522, 2006.

[23] W. Duckworth and B. Mans, Randomised greedy algorithms for finding small k-dominating sets of random regular graphs, *Random Structures and Algorithms*, 27(3):401–412, 2005.

[24] K. Edwards and G. Farr, Fragmentability of graphs, *Journal of Combinatorial Theory, Series B* 82:30–37, 2001.

[25] P. Erdős, Graph theory and probability, *Canadian Journal of Mathematics* 11:34–38, 1959.

[26] P. Erdős, M. Saks, and V. T. Sós, Maximum induced trees in graphs, *J. Combinatorial Theory, Series B*, 41:61–79, 1986.

[27] D. Gijswijt, On a packet scheduling problem for smart antennas and polyhedra defined by circular-ones matrices, *Electronic Notes in Discrete Mathematics*, 17:175–178, 2004.

[28] C. Godsil and G. Royle, *Algebraic Graph Theory,* Springer Verlag, New York, 2001.

[29] P. Haxell, O. Pikhurko and A. Thomason, Maximum acyclic and fragmented sets in regular graphs, *to appear in Journal of Graph Theory*, 7 pp.

[30] S. Hoory, N. Linial and A. Wigderson, Expander graphs and their applications, *Bulletin of the AMS*, 43:439-561, 2006.

[31] G. Hopkins and W. Staton, Girth and independence ratio. *Canadian Mathematical Bulletin* 25:179–186, 1982.

[32] C. Hoppen and N.C. Wormald, Induced forests in regular graphs with large girth, *Combinatorics, Probability and Computing* 17:1–22, 2007.

[33] W. Hurewicz, *Lectures on Ordinary Differential Equations,* MIT Press, 1958.

[34] E. L. Ince, *Ordinary Differential Equations,* Dover, New York, 1955.

[35] A. B. Kahng, S. Vaya, and A. Zelikovsky, New graph bipartizations for doubleexposure, bright field alternating phase-shift mask layout, *Proc. Asia and South Pacific Design Automation Conference,* 133–138, 2001.

[36] R.M. Karp, Reducibility among combinatorial problems, *Complexity of Computer Computation*, pages 85–103, 1972.

[37] G. Kirchhoff, Über die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Verteilung galvanischer Ströme geführt wird, *Ann. Phys. Chem.*, 72:497–508, 1847.

[38] D. Kühn and D. Osthus, Topological minors in graphs of large girth, *Journal of Combinatorial Theory, Series B* 86:364–380, 2002.

[39] J. Lauer and N.C. Wormald, Large independent sets in random graphs with large girth, *Journal of Combinatorial Theory, Series B* 97:999–1009, 2007.

[40] D.M. Li and Y.P. Liu, A polynomial algorithm for finding the minimum feedback vertex set of a 3-regular simple graph, *Acta Math. Sci.*, 19(4):375–381, 1999.

[41] Y.D. Liang, On the feedback vertex problem in permutation graphs, *Inform. Process. Lett.*, 52:123–129, 1994.

[42] Y.D. Liang and M.S. Chang, Minimum feedback vertex sets in cocomparability graphs and convex bipartite graphs, *Acta Inform.*, 34:337–346, 1997.

[43] R. de la Llave, *Computer Aided Proofs in Analysis,*(K. Meyer and D. Schmidt, eds), Springer Verlag, New York, 1991.

[44] B. D. McKay, Independent sets in regular graphs of high girth, *Proceedings of the Australia-Singapore Joint Conference on Information Processing and Combinatorial Mathematics, Singapore, 1986*, *Ars Combinatoria* 23A:179–185, 1987.

[45] M. Pop, D. S. Kosack, and S. L. Salzberg, Hierarchical scaffolding with Bambus, *Genome Research* 14(1):149–159, 2004.

[46] N. Robertson, D. Sanders, P. Seymour and R. Thomas, The four-colour theorem, *Journal of Combinatorial Theory, Series B* 70:2–44, 1997.

[47] R. W. Robinson and N. C. Wormald, Almost all regular graphs are hamiltonian, *Random Structures and Algorithms* 5:363–374, 1994.

[48] A. Schrijver, *Combinatorial Optimization: Polyhedra and Efficiency,* Springer Verlag, Berlin Heidelberg, 2003.

[49] J. B. Shearer, A note on the independence number of triangle-free graphs, *Discrete Mathematics* 46:83–87, 1983.

[50] J. B. Shearer, A note on the independence number of triangle-free graphs, II, *Journal of Combinatorial Theory, Series B* 53:300–307, 1991.

[51] L. Shi and N. C. Wormald, Colouring random 4-regular graphs, *Combinatorics, Probability and Computing* 16(2):309–344, 2007.

[52] C. Thomassen, Girth in graphs, *Journal of Combinatorial Theory, Series B* 35:129–141, 1983.

[53] C. Wang ,E. Lloyd and M. Soffa, Feedback vertex sets and cyclically reducible graphs, *Journal of the ACM* 32:296–313, 1985.

[54] N. C. Wormald, *Some Problems in the Enumeration of Labelled Graphs*, Doctoral thesis, Newcastle University, 1978.

[55] N. C. Wormald, The asymptotic distribution of short cycles in random regular graphs, *J. Combinatorial Theory, Series B* 31:168–182, 1981.

[56] N. C. Wormald, Differential equations for random processes and random graphs, *The Annals of Applied Probability* 5, No.4 :1217–1235, 1995.

[57] N. C. Wormald, The differential equation method for random graph processes and greedy algorithms, *Lectures on Approximation and Randomized Algorithms,* (M. Karoński and H.J. Prömel, eds) 73–155, 1999.

[58] N. C. Wormald, Analysis of greedy algorithms on graphs with bounded degrees, *Discrete Mathematics* 273:235–260, 2003.

Appendix A

Example 3.1.8 revisited

In this appendix, we add a detailed discussion of Example 3.1.8, in which the independence assumptions are proven and the probabilities are calculated in the probability space of sets of labels.

Example A.1.4 *We consider a version of Algorithm 3.1.1 with parameters $r = 3$, $d = 1$ and $\ell = 2$. Let $N = 1$ and fix probabilities $p_0 > 0$ and $p_{1,j,k}$, where $p_{1,0,k} = 0$, for every k, $p_{1,1,0} = p_{1,0} > 0$, $p_{1,1,1} = p_{1,1} > 0$ and $p_{1,1,2} = 1$. Let G be a 3-regular graph with girth larger than 9. We shall calculate the probability of the partial colouring of G in Figure A.1.1, at time 1. In the figure, the white vertices are represented by unfilled circles, while the purple vertices are filled circles.*

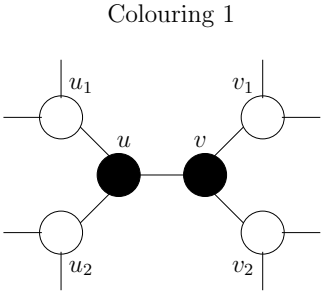

Figure A.1.1: A partial colouring of G.

We wish to calculate the probability that the adjacent vertices u and v are both purple and that their remaining neighbours are all white at time 1, that is, that they are white after step 1 in an application of the locally greedy algorithm. We shall consider cases according to the relevant labels of u and v. First note that the relevant labels cannot both be 1. If they were, u and v would not have a purple neighbour at time 0 and therefore would not turn purple at time 1, as $p_{1,0,k} = 0$ for every k. Moreover, the case for which u has relevant label 0 and v has relevant label 1 is symmetric to the case for which u has relevant label 1 and v

has relevant label 0, so their probabilities are the same by independence of vertex labelling. This leaves us with two cases to consider.

Case 1: In this case, u and v both have relevant label 0. Since u_1, u_2, v_1 and v_2 are white at time 1, they clearly cannot have relevant label 0, and neither can their neighbours other than u and v (otherwise one of u_1, u_2, v_1 and v_2 would be yellow). The probability that these events occur is $p_0^2(1-p_0)^{12}$.

Now, Corollary 3.1.6 implies that the colour of any vertex in the graph at time $i = 1$ depends only on the labels of vertices within distance $(d+1)i + d = 3$ of this vertex. In particular, only the labelling of the subgraph of G given in Figure A.1.2 is relevant for Colouring 1. Note that the colours in this figure represent the colouring of the vertices at

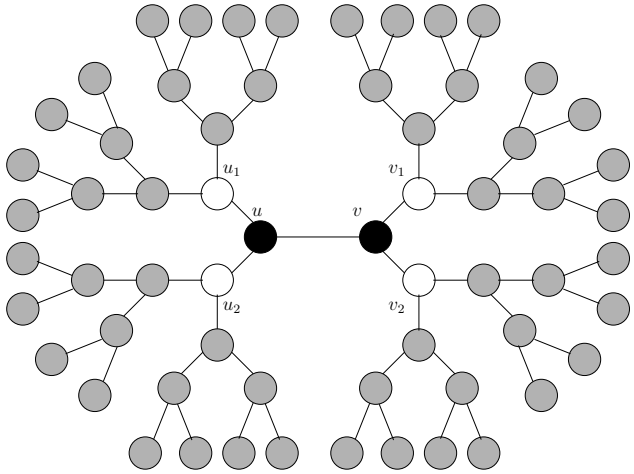

Figure A.1.2: The vertices whose labels may influence Colouring 1.

time 0 under the assumptions already made. A vertex whose colour is unknown is coloured grey in the picture.

The branches rooted at u_1, u_2, v_1 and v_2 are defined as $T_{u,1,4}$, $T_{u,2,4}$, $T_{v,1,4}$ and $T_{v,2,4}$, respectively. We prove that, conditional upon the events already assumed, the labelling of each branch cannot influence the colour, at time 1, of any vertex in the set $A = \{u, v, u_1, u_2, v_1, v_2\}$ other than its root vertex.

Claim A.1.5 *Conditional on the event that, among all vertices at distance at most 2 from u or v, only u and v have relevant label 0, the labelling of $T_{u,1,4}$ does not affect the colour, at time 1, of a vertex in A other than its root vertex u_1.*

Proof of the Claim Consider the branch $T_{u,1,4}$ rooted at u_1 and fix two labellings of the graph, $\mathcal{S}$ and $\hat{\mathcal{S}}$, differing only in $T_{u,1,4}$ but not inducing the same colouring on A at time 1.

Let $w_1, \ldots, w_k$ be the vertices whose labels with respect to $\mathcal{S}$ and $\hat{\mathcal{S}}$ are different, and let $\mathcal{S}_j$ be the labelling obtained by assigning the labels of $\hat{\mathcal{S}}$ to the vertices $w_1, \ldots, w_j$ and the

labels of $\mathcal{S}$ to all other vertices in the graph. In particular, $\mathcal{S}_0 = \mathcal{S}$ and $\mathcal{S}_k = \hat{\mathcal{S}}$, so there is a smallest $j \in \{1, \ldots, k\}$ such that the colourings induced to A by $\mathcal{S}_{j-1}$ and $\mathcal{S}_j$ differ. Let w be a vertex in A for which the colour changes. Clearly, w is not equal to u, since this vertex has relevant label 0, hence its colour cannot be modified. There are two reasons for the colour of w to change. On the one hand, it could have relevant label 1 with respect to exactly one of the labellings, and, on the other hand, one of its neighbours, call it $\hat{w}$, could have relevant label 0 or 1 with respect to only one of the labellings. In the former case, using Lemma 3.1.5, we conclude that the vertex w_j must be at distance at most 2 from w. This lemma also ensures that $\min\{L_{\mathcal{S}}(w_j), L_{\hat{\mathcal{S}}}(w_j)\} < \min\{L_{\mathcal{S}}(w), L_{\hat{\mathcal{S}}}(w)\} = 1$, so w_j must have relevant label 0 with respect to only one of the labellings. This is a contradiction, since the only vertex in $T_{u,1,4}$ within distance two of w is u_1, and we are conditioning upon u_1 not having relevant label 0. The same contradiction may be reached by applying the same argument to $\hat{w}$, since $\hat{w}$ cannot be equal to u. This establishes our claim. □

Therefore, the colour of u_1 at time 1 only depends on the labelling of the branch $T_{u,1,4}$ rooted at u_1, and the corresponding statement is true for the vertices u_2, v_1 and v_2. By symmetry, it suffices to look at the labellings of $T_{u,1,4}$ for which the vertex u_1 is white at time 1. The vertices in this branches are named as in Figure A.1.3.

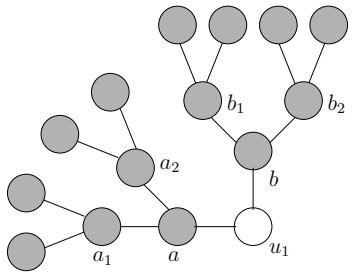

Figure A.1.3: The vertices in the branch rooted at u_1.

We already know that u_1, a and b do not have relevant label 0. Under these conditions, u_1 is not white at time 1 if and only if at least one of these three vertices has relevant label 1, since we already know that u is purple at time 1. We subdivide into cases according to the labels of a_1, a_2, b_1 and b_2 at time 0. Symmetric cases are treated together. Moreover, all the cases are considered conditional on the assumptions made so far, namely on the event that, among all vertices at distance at most two of u or v, only u and v have relevant label 0.

Case 1.0: None of the four vertices have relevant label 0. This happens with probability $(1 - p_0)^4$.

As a consequence, a and b are white at time 0, so u_1 does not have relevant label 1 with probability $(1 - p_{1,0})$. Moreover, none of the neighbours of a and b are purple at time 0, hence a and b cannot have relevant label 1 by our choice of probabilities.

Hence, the conditional probability associated with Case 1.0 conditional on the earlier discussion is equal to

$$(1-p_0)^4(1-p_{1,0}).$$

Case 1.1: Exactly one vertex in $\{a_1, a_2, b_1, b_2\}$ has relevant label 0. This happens with probability $4p_0(1-p_0)^3$.

Without loss of generality, assume that a_1 has relevant label 0. We again have that a and b are white at time 0. As a consequence, u_1 does not have relevant label 1 with probability $(1-p_{1,0})$. Since b does not have any purple neighbours at time 0, it cannot have relevant label 1. Now, a will be have relevant label 1 with probability $p_{1,1}$ if a_2 is yellow and with probability $p_{1,0}$ if a_2 is white. The former happens with probability p_0^2, since its two neighbours of other than a have to be purple. The latter occurs with probability $(1-p_0^2)$. We deduce that a does not have relevant label 1 with probability $1-(p_{1,1}p_0^2+p_{1,0}(1-p_0^2))$.

The probability associated with this case is therefore

$$4p_0(1-p_0)^3(1-p_{1,0})[1-(p_{1,1}p_0^2+p_{1,0}(1-p_0^2))].$$

Case 1.2: Exactly two vertices in $\{a_1, a_2, b_1, b_2\}$ have relevant label 0. Two distinct situations may arise in this case.

If the two purple vertices at time 0 are adjacent to the same vertex, call it a, then a is purple at time 0 and b is white at time 0. Now, u_1 does not have relevant label 1 with probability $(1-p_{1,1})$, while a and b cannot have relevant label 1.

If a and b are both adjacent to a purple vertex at time 0, the probability that none of u_1, a and b have relevant label 1 is

$$(1-p_{1,0})[1-(p_{1,1}p_0^2+p_{1,0}(1-p_0^2))]^2,$$

by proceeding as in Case 1.1.

The probability associated with this case is

$$2p_0^2(1-p_0)^2(1-p_{1,1})+4p_0^2(1-p_0)^2(1-p_{1,0})[1-(p_{1,1}p_0^2+p_{1,0}(1-p_0^2))]^2.$$

Case 1.3: Exactly three vertices in $\{a_1, a_2, b_1, b_2\}$ have relevant label 0. So one of a and b is yellow at time 0, while the other is white with a purple neighbour. Using the arguments of the previous cases, the probability associated with this case is

$$4p_0^3(1-p_0)(1-p_{1,1})[1-(p_{1,1}p_0^2+p_{1,0}(1-p_0^2))].$$

Case 1.4: The vertices a_1, a_2, b_1 and b_2 have relevant label 0. This situation cannot happen in any labelling implying our event, since u_1 would be white with a purple neighbour and two yellow neighbours at time 0, so it would have relevant label 1 with probability 1.

This concludes the analysis of Case 1. In particular, if $q = 1 - (p_{1,1}p_0^2 + p_{1,0}(1-p_0^2))$ and

$$r = (1-p_0)^4(1-p_{1,0}) + 4p_0(1-p_0)^3(1-p_{1,0})q + 2p_0^2(1-p_0)^2(1-p_{1,1}) \\ + 4p_0^2(1-p_0)^2(1-p_{1,0})q^2 + 4p_0^3(1-p_0)(1-p_{1,1})q,$$

the probability associated with Case 1, that is, the probability that u and v have relevant label 0 and that u_1, u_2, v_1 and v_2 are all white at time 1 is equal to.

$$p_0^2(1-p_0)^{12}r^4.$$

Case 2: By symmetry, we shall only consider the case for which u has relevant label 0 and v has relevant label 1, which will be multiplied by 2 to give the total probability associated with this case. Again, u_1, u_2, v_1 and v_2 cannot have relevant label 0, and neither can their neighbours other than u and v. Indeed, if one such vertex adjacent to u_1 had relevant label 0, u_1 would again be yellow at time 0, whereas, if a neighbour of v_1 had relevant label 0, v_1 would be yellow at time 1 . From here, we also deduce that v has relevant label 1 with probability $p_{1,0}$. So, these events all occur with probability $p_0(1-p_0)p_{1,0}(1-p_0)^{12}$. Conditional upon these events, we may proceed just as in Case 1. Here, we can also show that the labelling of the branches rooted at u_1, u_2, v_1 and v_2 cannot influence the colour, at time 1, of any vertex in the set $A = \{u, v, u_1, u_2, v_1, v_2\}$ other than its root vertex. The branches rooted at u_1 and u_2 can be treated exactly as before, while the branches rooted at v_1 and v_2 must take into account the fact that v does not have relevant label 0. Such calculations gives us the probability associated with Case 2, which is equal to

$$2p_0(1-p_0)p_{1,0}(1-p_0)^{12}r^2s^2,$$

where r comes from Case 1 and

$$s = (1-p_0)^4 + 4p_0(1-p_0)^3q + 2p_0^2(1-p_0)^2 + 4p_0^2(1-p_0)^2q^2 + 4p_0^3(1-p_0)q.$$

The difference between this term and r is that, because v does not have relevant label 0, its neighbour in the branch, say v_1, cannot have relevant label 1.

The probability that the labelling of G induces Colouring 1 at time 1 is the sum of the probabilities associated with Cases 1 and 2, that is,

$$p_0^2(1-p_0)^{12}r^4 + 2p_0(1-p_0)p_{1,0}(1-p_0)^{12}r^2s^2. \ \square$$

Appendix B

A determinant

In this appendix, we prove a formula for the determinant of a matrix that was used repeatedly in this thesis.

Proposition B.1.6 *Let r be a positive integer and let $a_1, \ldots, a_r$ be real numbers. Consider the square matrix*

$$A = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ ra_1 & (r-1)a_1 - 1 & (r-2)a_1 & \cdots & a_1 \\ ra_2 & (r-1)a_2 & (r-2)a_2 - 1 & \cdots & a_2 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ ra_{r-1} & (r-1)a_{r-1} & (r-2)a_{r-1} & \cdots & a_{r-1} - 1 \end{bmatrix}.$$

Then

$$\det A = (-1)^{r-1} \left(1 + \sum_{k=1}^{r-1} k\, a_k \right).$$

Proof If $B = (b_{i,j})_{i,j=1}^{n}$ is a matrix, basic linear algebra tells us that the determinant of B is given by

$$\det B = \sum_{\sigma} sgn(\sigma) \prod_{i=1}^{n} b_{i,\sigma(i)}, \tag{B.1.1}$$

where σ ranges over all permutations of the sequence $(1, 2, \ldots, n)$. The summand associated with a fixed permutation σ will be called the *term of the permutation* σ. The function $sgn(\sigma)$, called the *sign* of a permutation, assigns value 1 to *even* permutations, i.e., to permutations that can be obtained from $(1, 2, \ldots, n)$ by interchanging an even number of pairs, and assigns value -1 to the *odd* permutations, which are the permutations obtained from $(1, 2, \ldots, n)$ by interchanging an odd number of pairs.

Applying formula (B.1.1) to the matrix A, we can see that $\det A$ is a polynomial in the variables $a_1, \ldots, a_{r-1}$, which is linear with respect to each a_k. Here, we shall index the

rows and columns from 0 to $r-1$, since this will simplify the discussion. Analogously, the permutations are permutations of $(0, 1, \ldots, r-1)$.

As a polynomial the variables $a_1, \ldots, a_{r-1}$, the only terms with degree 0 in A are the elements in the first row and the summands -1 in each element of the diagonal. Thus, the only permutation whose term can generate a monomial of degree 0 is $\sigma = (0, \ldots, r-1)$. This permutation is even by definition, and the monomial of degree 0 that it generates is $(-1)^{r-1}$.

Now, a monomial of degree 1 on the variable a_k has to be generated by a permutation σ such that $\sigma(i) = i$ for every $i \in \{1, \ldots, r-1\} \setminus \{k\}$, since all the non-diagonal elements in the i-th row of the matrix are monomials of degree 1 in the variable a_i. The only permutations satisfying this are the identity and the permutation that interchanges the elements 0 and k. By definition, the former is even and generates the monomial $(-1)^{r-2}(r-k)a_k$, while the latter is odd, and therefore generates the monomial $(-1)^{r-1}ra_k$. In particular, the monomials of degree in $\det A$ sum to

$$\sum_{k=1}^{r-1}(-1)^{r-1}\left(ra_k - (r-k)a_k\right) = (-1)^{r-1}\sum_{k=1}^{r-1} k\, a_k.$$

Now, let $2 \leq l \leq r-1$. Fix $J = \{i_1, \ldots, i_l\} \subseteq \{1, \ldots, r-1\}$ and consider the coefficient of $a_{i_1}a_{i_2}\cdots a_{i_l}$ in $\det A$. As in the previous cases, we may again argue that any permutations whose term generates such a monomial must keep the elements in $\{1, \ldots, r-1\} \setminus J$ fixed. For each $k \in \{0\} \cup J$, we consider the set of permutations Σ_k for which $\{1, \ldots, r-1\} \setminus J$ is fixed and $\sigma(0) = k$. By the definition of the matrix a, it is easy to see that, in absolute value, the coefficient of $a_{i_1}a_{i_2}\cdots a_{i_l}$ generated by each $\sigma \in \Sigma_k$ is equal to

$$\frac{r!}{(r-k)\prod_{j \notin J}(r-j)}.$$

So, if we show that, for every k, exactly half the permutations in Σ_k are even, it will follow that the monomials generated by even and odd permutations will cancel out, so that the coefficient of $a_{i_1}a_{i_2}\cdots a_{i_l}$ in $\det A$ is equal to zero.

If $k = 0$, the permutations in Σ_0 are the ones that fix $\{0, \ldots, r-1\} \setminus J$. These can be viewed as permutations of the symbols in J, and, since J contains at least two elements, the same number of them are even or odd. Note that the sign of the permutation $\sigma \in \Sigma_0$ is equal to the sign of its restriction to J, since all other elements are fixed, so our result follows in this case.

If $k \neq 0$, each element in Σ_k is determined by a bijection between the sets J and $J' = (J \setminus \{k\}) \cup \{0\}$, since the image of all the other elements is known. By formally identifying the element k in J and the element 0 in J' as the same element, these can again be seen as permutations in a set of symbols containing at least two elements, so the same number of them are even or odd. Now, the sign of a permutation $\sigma \in \Sigma_k$ is clearly the opposite of

the sign of the permutation obtained by restricting this permutation to J and identifying the elements k and 0, so it is still true that exactly half of them are even.

This implies that the coefficient of any monomial of degree at least 2 in $\det A$ is 0, and therefore

$$\det A = (-1)^{r-1} \left(1 + \sum_{k=1}^{r-1} k\, a_k \right),$$

as required. ∎