Combinatorics, Probability and Computing (2007) 16, 459–494. © 2006 Cambridge University Press
doi:10.1017/S0963548306007954 First published online 14 August 2006 Printed in the United Kingdom

Colouring Random Regular Graphs

LINGSHENG SHI and NICHOLAS WORMALD[†]

Department of Combinatorics and Optimization, University of Waterloo, Waterloo ON, Canada N2L 3G1
(e-mail: lshi@uwaterloo.ca, nwormald@uwaterloo.ca)

Received 19 October 2004; revised 15 May 2006

In a previous paper we showed that a random 4-regular graph asymptotically almost surely (a.a.s.) has chromatic number 3. Here we extend the method to show that a random 6-regular graph asymptotically almost surely (a.a.s.) has chromatic number 4 and that the chromatic number of a random d-regular graph for other d between 5 and 10 inclusive is a.a.s. restricted to a range of two integer values: $\{3,4\}$ for $d=5$, $\{4,5\}$ for $d=7$, 8, 9, and $\{5,6\}$ for $d=10$. The proof uses efficient algorithms which a.a.s. colour these random graphs using the number of colours specified by the upper bound. These algorithms are analysed using the differential equation method, including an analysis of certain systems of differential equations with discontinuous right-hand sides.

1. Introduction

Random regular graphs are a commonly studied model of random graph. They have some features that are strikingly different from other models, for instance the connectivity is higher at lower densities. See Bollobás [4], or the more recent survey paper [20], for more information.

We showed [17] that the chromatic number of random 4-regular graphs is 3, with probability tending to 1 as the size of the graph tends to ∞. In this paper we extend this investigation to the uniform model $\mathscr{G}_{n,d}$ of random d-regular graphs on n labelled vertices. (We restrict nd to the even integers in all asymptotics.) We obtain significant new upper bounds on $\chi(G)$, holding asymptotically almost surely (a.a.s.) for $G \in \mathscr{G}_{n,d}$, for small values of d. These imply in particular that a random 6-regular graph a.a.s. has chromatic number 4. For other values of $d \leqslant 10$, the bounds restrict χ to one of two possible values. Along with the upper bounds, we obtain algorithms which a.a.s. colour these random graphs with the number of colours given by the upper bound. This compares with a recent non-algorithmic result of Achlioptas and Naor [2] that, for the common random

[†] Research supported by the Canada Research Chairs Program and NSERC.

https://doi.org/10.1017/S0963548306007954 Published online by Cambridge University Press

graph model $\mathcal{G}(n, d/n)$ (in which edges between the n vertices are chosen independently at random with probability d/n), the chromatic number is, when d is fixed, concentrated on at most two integers, and on one value for d in some intervals.

A number of results on the chromatic number of $\mathcal{G}_{n,d}$ have already appeared. Frieze and Łuczak [11] showed that χ is a.a.s. asymptotic to $\frac{d}{2\log d}$, in the sense that as $d \to \infty$ moderately slowly, there are functions of d asymptotic to $\frac{d}{2\log d}$ that are a.a.s. upper and lower bounds on $\chi(G)$ for $G \in \mathcal{G}_{n,d}$. This was extended by Cooper, Frieze, Reed and Riordan [8] to $d = o(n^{1-\eta})$ for all fixed $\eta > 0$ (provided $d \to \infty$), and, earlier than that, by Krivelevich, Sudakov, Vu and Wormald [14, Theorem 2.5] to $n^{6/7+\eta} \leqslant d \leqslant 0.9n$ with the formula adjusted to $(1 + o(1))n/(2\log_b d)$ where $b = n/(n-d)$. However, these results say nothing about fixed d. It is known (see Borodin and Kostochka [5], Catlin [7] and Lawrence [15]) that for all K_4-free graphs G with maximum degree d, $\chi(G) \leqslant \frac{3}{4}(d+2)$, which consequently gives a bound holding a.a.s. for random d-regular graphs (d fixed), since they a.a.s. contain no K_4. Also Kostochka [13] showed that a d-regular graph ($d \geqslant 5$ and fixed) with sufficiently large girth has chromatic number at most $2 + \lfloor d/2 \rfloor$. By the known properties of $\mathcal{G}_{n,d}$, this implies that for a random graph in $\mathcal{G}_{n,d}$, $\mathbf{P}(\chi \leqslant 2 + \lfloor d/2 \rfloor)$ is bounded below by a positive constant. Thus, for instance, when $d = 5$, $\mathbf{P}(\chi \leqslant 4)$ is bounded below by a positive constant.

On the other hand, Molloy and Reed [16] found lower bounds holding a.a.s. for d fixed. These arise simply from finding values k and d for which the expected number of k-colourings of a random d-regular graph tends to 0. Because it is not readily accessible, we provide a proof of these lower bounds, along the lines of the argument in [16], at the end of this section. They cannot be obtained by a simple direct evaluation of the expected number of k-colourings.

While the present paper was being prepared, Achlioptas and Moore [1] announced a proof of two things: firstly that the chromatic number of $\mathcal{G}_{n,d}$ is 2-point concentrated, so is a.a.s. k or $k+1$ for some $k = k(d)$, and secondly that if k is the smallest integer such that $d < 2k\log k$ then a.a.s. $\mathcal{G}_{n,d}$ has chromatic number k, $k+1$, or $k+2$. Additionally, for roughly half of all d, the lowest value k is eliminated by the Molloy–Reed lower bound.

Our upper bounds for small values of d, together with the known lower bounds, give the following results. These cut one value, the uppermost, off the range of possible values given in [1] for $d \leqslant 9$. A much less explicit theorem for arbitrary fixed values of d is given in Theorem 1.2 below. Note that $d = 5$ is the smallest unknown case (see [17]), and, after the following results, $d = 7$ is next.

Theorem 1.1. *Let G be a random d-regular graph and let χ be the chromatic number of G. Then a.a.s. $\chi = 4$ for $d = 6$. Moreover, a.a.s. $\chi = 3$ or 4 for $d = 5$, $\chi = 4$ or 5 for $7 \leqslant d \leqslant 9$, and $\chi = 5$ or 6 for $d = 10$.*

We also give 'simple' algorithms which a.a.s. colour a random d-regular graph with the number of colours given by the upper bound, based on a greedy colouring scheme. To minimize repetition of material in [17], our proof refers heavily to it, so the reader will need to consult that paper to make sense of the present paper from Section 4 onwards.

https://doi.org/10.1017/S0963548306007954 Published online by Cambridge University Press

Our method extends to $d > 10$ in theory, as shown by the next theorem, but computations were too difficult to carry out accurately enough to obtain useful results. To obtain the results in Theorem 1.1, a program was used with floating point arithmetic and precise bounding of the consequential errors in the calculations.

In [17] we defined the algorithm **Colour**(k) to colour the vertices of a graph, using colours $1,\dots,k$, in the following manner. First, define the *type* of a vertex v in a partially coloured graph to be the pair (i,j) where i is the number of neighbours of v which are already coloured, and j is the number of different colours appearing on such neighbours (*i.e.*, not available to legally colour v). The algorithm takes as input an ordering of the possible types, as a list of priorities, and then repeatedly randomly chooses an uncoloured vertex of the highest priority type and colours it with one of its available colours chosen at random.

To reduce the number of colours required, we perform a preprocessing step of 3-colouring the vertices in all short cycles in the graph. This results in the algorithm **Short**(g,k) defined as follows. The input is a d-regular graph G and a priority list of types. A *short* cycle is one of length at most g.

Algorithm Short(g,k)

(1) properly 3-colour the vertices in the short cycles of G;
(2) apply **Colour**(k).

We say this algorithm (and other colouring algorithms involving k colours) has *failed* at any time that a vertex has k differently coloured neighbours. It *succeeds* if, up until the time that all vertices of G have been coloured, it has not failed.

We claim that, for the values of d and k giving the upper bounds in Theorem 1.1, the algorithm **Short**(g,k) (with a suitable list of priorities) succeeds with probability at least $1-\epsilon_g+o(1)$ when applied to a random d-regular graph on n vertices, as $n\to\infty$. Here ϵ_g is, for each d and k, a function which tends to 0 as $g\to\infty$. It follows that, for these values of d and k, a random $G\in\mathscr{G}_{n,d}$ is a.a.s. k-colourable, where, as in all asymptotic statements throughout this paper, d is constant.

For reasons we will discuss, analysis of **Short**(g,k) is slightly more complicated than the algorithm we actually examine, which is a modification in which several types are grouped together. The modified algorithm selects between the types in this group with certain predetermined probabilities. By selecting these probabilities appropriately, this algorithm mimics an instance of **Short**(g,k), and in this case we refer to the modification loosely as a *partial deprioritization* of **Short**(g,k). This leads to a general theorem which requires some definitions before being stated.

We denote the priority of a type (i,j) by $\mathtt{prio}(i,j)$. For convenience, we assume that the set of priorities is $0,\dots,\ell_{\max}$, using ℓ to denote a particular priority $\mathtt{prio}(i,j)$, with an ordering specified by $\ell'>\ell$ if (i',j') has higher priority than (i,j). We introduce a vector function of a variable z, $\mathbf{y}=(y_0,\dots,y_{\ell_{\max}})$, where $y_\ell=y_\ell(z)$. Define $i(\ell)$ and $j(\ell)$ so that $\ell=\mathtt{prio}(i(\ell),j(\ell))$ for all priorities ℓ, and we refer to y_ℓ for convenience also as $y_{i,j}$ if $\mathtt{prio}(i,j)=\ell$ (*i.e.*, y_ℓ is also known as $y_{i(\ell),j(\ell)}$). Now define

$$R_1=R_1(\mathbf{y})=\max\{\ell : y_\ell>0\},\quad R_2=R_2(\mathbf{y})=\min\{\ell : F_{\ell'}(\mathbf{y})<1 \text{ for all } \ell'\geqslant\ell\},\qquad (1.1)$$

https://doi.org/10.1017/S0963548306007954 Published online by Cambridge University Press

where, with $i = i(\ell)$ and $j = j(\ell)$,

$$F_\ell(\mathbf{y}) = \sum_{\ell' > \ell} c(\ell') b_{\ell'}(\mathbf{y}), \qquad c(\ell) = (d-i)\binom{k}{j}, \tag{1.2}$$

$$b_\ell(\mathbf{y}) = \frac{(d-i+1)j}{(d-2z)k}(y_{i-1,j} + y_{i-1,j-1}) - \frac{d-i}{d-2z} y_{i,j}. \tag{1.3}$$

We will consider the differential equations

$$\frac{\mathrm{d}y_\ell}{\mathrm{d}z} = \begin{cases} b_\ell(\mathbf{y}) & \text{if } \ell < \ell_0, \\ 0 & \text{if } \ell > \ell_0, \\ b_\ell(\mathbf{y}) + (F_\ell(\mathbf{y}) - 1)/c(\ell) & \text{if } \ell = \ell_0, \end{cases} \tag{1.4}$$

where

$$\ell_0 = \mathtt{prio}(i_0, j_0) = \max\{R_1, R_2\} \tag{1.5}$$

is a function of $\mathbf{y}$. The right-hand sides of these equations are not always continuous, so we have to be careful (see Section 3).

A type (i, j) is *dangerous* if $j \geqslant k-1$ and $i < d$, and we define B' to be the set of dangerous types. In the following theorem, *simple* modification of **Short**(g, k) means that its second step, **Colour**(k), is partially deprioritized: when there are no dangerous vertices, the type of the next vertex is chosen not according to priority but according to a probability which can be specified in advance for all n and all steps of the process. These probabilities can be calculated using solutions of equations equivalent to (1.4), as will be evident from the proof.

Theorem 1.2. *Suppose that all dangerous types have higher priority than every non-dangerous type. Suppose furthermore that for every function* $\mathbf{y}$ *whose right-hand derivative satisfies the differential equations* (1.4) *with* ℓ_0 *defined in* (1.5) *and with initial conditions*

$$y_{0,0}(0) = 1, \ y_{i,j}(0) = 0 \text{ otherwise}, \tag{1.6}$$

there is a $z_1 > 0$ *such that:*

(i) *for* $0 < z < z_1$, $y_{0,0} > 0$ *and*

$$\sum_{\ell' \in B'} (d - i(\ell')) \binom{k}{j(\ell')} b_{\ell'}(\mathbf{y}(z)) < 1,$$

and

(ii)

$$\frac{\sum_{i,j}(d-i)(d-i-1)\binom{k}{j} y_{i,j}(z_1)}{\sum_{i,j}(d-i)\binom{k}{j} y_{i,j}(z_1)} < 1. \tag{1.7}$$

Then a random $G \in \mathscr{G}_{n,d}$ *is a.a.s. k-colourable. Moreover, a simple modification of* **Short**(g, k) *a.a.s. succeeds when applied to a random* $G \in \mathscr{G}_{n,d}$.

This theorem's proof is completed at the end of Section 4.

As with [17], this paper uses the differential equation method introduced in [18]. Unlike [17], some of the approach and terminology of [9] is used to deprioritize algorithms

https://doi.org/10.1017/S0963548306007954 Published online by Cambridge University Press

which perform many types of operations with a given list of priorities. This type of deprioritization was introduced in [21] to avoid phase changes occurring, such as when a high priority type suddenly disappears. These phase changes cause difficulties: not only does the associated system of differential equations change at these points, but also the argument showing that the equations model the graph process becomes quite awkward. In this paper, we extend the theory further, by working with systems of equations which specify right-hand derivatives only. The derivatives in the natural setting are sometimes discontinuous at phase changes. In the past ([9], [18], [19] and [21], for example), the associated solution functions were cut apart at the phase changes so that within each phase the derivative is smooth. Due to the difficulty of identifying the phase changes from the numerical solutions of differential equations, we introduce in this paper some extra theory which enables us to avoid the need to identify phase changes precisely.

The biggest open problem in this area is to settle the question of $d = 5$: do random 5-regular graphs have chromatic number 3 a.a.s.? (Or is it 4, or a mixture?) At the time of the final revisions of the present paper, the answer seems to be 3: Díaz, Grammatikopoulos, Kaporis, Kirousis, Perez and Sotiropoulos [10] showed that, assuming that the maximum of a certain function occurs at the place where it appears to, the probability the chromatic number is 3 is bounded away from 0 (they worked with multigraphs but a slight modification allows them to treat graphs). Combining their arguments with those of Kemkes and the second author [12] shows that, under the same assumption, the chromatic number of a random 5-regular graph is indeed a.a.s. 3. However, the assumption has not yet been proved. Another question is whether the ideas in this paper will improve the results for random graphs $\mathscr{G}(n, p)$.

For completeness, we include here what is essentially the proof of the Molloy–Reed lower bound in [16]. For this and later work in the paper, we use the *pairing* model of $\mathscr{G}_{n,d}$. A pairing is a perfect matching of dn points which are partitioned into n cells, each containing exactly d points. Such a pairing P corresponds to a multigraph $G(P)$ in which the cells are regarded as vertices and the $dn/2$ pairs are edges. A uniformly random pairing can be generated by repeatedly choosing (using any rule whatsoever) a point x to be paired, and then randomly choosing the point y to be paired with x from the remaining unpaired points. We refer to this choice of y as *exposing* the pair containing x. Conditional on $G(P)$ having no loop or multiple edge, it has the uniform distribution and is thus identical to $\mathscr{G}_{n,d}$. (See [4] or [20] for more details.)

Theorem 1.3 (Molloy–Reed). *If $k(1-1/k)^{d/2} < 1$, then a random $G \in \mathscr{G}_{n,d}$ is a.a.s. not k-colourable.*

Proof. The first moment method is used. To compute the expected number of proper k-colourings of the vertices of a random pairing, we need to compute the number of pairings with the k colour classes specified, and divide by the total number of pairings which is known to be $(dn-1)!!$.

It is straightforward to estimate this number within a polynomial factor, if one first shows that the number of pairings respecting a given colour partition is maximized when colour classes are almost equal. We omit some of the details justifying this.

https://doi.org/10.1017/S0963548306007954 Published online by Cambridge University Press

Suppose V_1 and V_2 are two of the colour classes in a k-colouring and $|V_1| > |V_2|$. A *good* pairing is one whose induced graph is properly coloured by the given k-colouring, otherwise it is *bad*. Now let $v \in V_1$ and move v to V_2. This forms a new k-colouring with $V_1' = V_1 \setminus \{v\}$ and $V_2' = V_2 \cup \{v\}$. Let B be the set of formerly good pairings which are bad with respect to the new colouring, and G the set of formerly bad pairings which are now good.

Let $M_1 \in B$, and let S be the set of edges in M_1 from V_1' to V_2. Choose a vertex w not incident with S, let w' be the matched neighbour of w. Then $w' \notin V_1$ or V_2. Let $v' \in V_2$ be the vertex matched to v. Create an element of G by switching the edges vv' and ww' of M_1 to vw and $v'w'$. This gives $|V_1'| - |S|$ good pairings. Conversely, each good pairing comes from $|V_2| - |S|$ bad pairings. Hence $|B| \leqslant |G|$. So the maximum number of good pairings occurs when the colour classes are roughly equal in size.

Now consider the number of pairings with k colour classes that are equal in size, and for simplicity assume the size is dn/k. Let $a_{i,j}$ be the number of pairs with colours i and j. Then for each i,

$$\sum_{1 \leqslant j \leqslant k} a_{i,j} = dn/k. \tag{1.8}$$

Let $\mathscr{A}$ be the set of all possible sequences $(a_{i,j})$ satisfying (1.8). Then the number of pairings is at most

$$\sum_{\mathscr{A}} \prod_{i<j} \binom{dn/k}{a_{i,j}}^2 a_{i,j}! \leqslant n^{O(1)} \binom{dn/k}{\frac{dn}{k(k-1)}}^{2\binom{k}{2}} \left(\frac{dn}{k(k-1)}!\right)^{\binom{k}{2}}.$$

The inequality follows on taking the logorithm of the product in the left-hand side (using Stirling's formula), and observing that each term then has negative second derivative. The maximum, subject to (1.8), thus occurs when all $a_{i,j} = dn/(k(k-1))$ (again assuming divisibility for simplicity). Hence for a fixed k-colouring, the probability that a random pairing is properly coloured is at most the value above divided by $(dn-1)!!$, giving $n^{O(1)}(k-1)^{dn}\{dn/(k(k-1))\}^{dn/2}/(dn)^{dn/2} = n^{O(1)}(1-1/k)^{dn/2}$. Now let X be the number of proper k-colourings of a random pairing. Since the total number of k-colourings is at most k^n, we have

$$\mathbf{P}(X > 0) \leqslant \mathbf{E}(X) \leqslant n^{O(1)} k^n (1-1/k)^{dn/2}.$$

Therefore if $k(1-1/k)^{d/2} < 1$ we have $\mathbf{P}(X > 0) \to 0$ as $n \to \infty$, that is, there is a.a.s. no proper k-colouring of the multigraph $G(P)$ of a random pairing. The well-known relationship between the pairing model and $\mathscr{G}_{n,d}$ (see [4] or [20]) now implies the same for $G \in \mathscr{G}_{n,d}$. □

2. Summary of the argument

In this section we give a summary of the argument for $d = 4$ in [17], and describe how this will be modified for general d in the present paper. The main part of this, Section 4, will refer heavily to the argument in [17], to which the reader should also refer if extra detail is sought regarding the following description.

https://doi.org/10.1017/S0963548306007954 Published online by Cambridge University Press

In [17], the algorithm **Short**(g,k) was analysed on $G \in \mathscr{G}_{n,d}$ using the pairing model described in Section 1. Almost all of that analysis was carried out for arbitrary fixed $d \geqslant 4$. A process **PColour**(k) was defined, corresponding to **Colour**(k), but colouring the vertices of the graph $G(P)$ of the random pairing P simultaneously with generating P at random. For a specified set $\mathscr{S}$ of pairs of a potential pairing, which induce disjoint cycles in $G(P)$, the process **PComplete**$(k,\mathscr{S})$ was defined. This begins with the pairs in $\mathscr{S}$, properly 3-colours the vertices in the induced cycles, and from there it proceeds as **PColour**(k) would. This process was used to analyse **Short**(g,k) applied to a random graph G_0 such that $G_0 \cup \mathscr{S}$ is a d-regular graph on n vertices, with no short cycles other than the cycles in $\mathscr{S}$ (with uniform probability distribution subject to this constraint). The process was restricted so that it avoided *unwanted* cycles, which are cycles created during the execution of this process that are short but not contained in $\mathscr{S}$. (In practice, loops, or cycles of length 1, are always unwanted, since $g \geqslant 3$.)

It was shown that, conditional upon there being no unwanted cycles up to a certain time $t_1 = t_1(n)$, defined below from the solution of a system of differential equations, with probability at least $1 - \epsilon_g + o(1)$ the process **PComplete**$(k,\mathscr{S})$ does not fail before t_1, where $\epsilon_g \to 0$ as $g \to \infty$. Here 'time' refers to the number of vertices of basic type which have been coloured, where 'basic' is defined below.

The differential equations define functions whose values (at the appropriate points) provide good approximations to the numbers of vertices of the various types (scaled, in this case, by dividing by n). Before defining the equations, a set of types is specified as 'basic'. In the context of [17], only the type $(1,1)$ is basic, and B' is the set of types of higher priority than basic. (This is related to the B' defined in Theorem 1.2, for which the interpretation is that all non-dangerous types are basic.) These equations were first derived only by a heuristic argument resulting in [17, equation (4.20)], which is easily seen to be equivalent to (1.4) with $\ell_0 = \mathtt{prio}(1,1)$. The rigorous derivation of equations came later, resulting in ones of a slightly different form (see [17, equations (6.16)–(6.20) and Theorem 10.1]; here $\mathbf{y} = (y_{0,0},\dots)$ as before and differentiation is with respect to a new variable x):

$$y'_{i,j} = \alpha(\mathbf{y})b_{i,j}(\mathbf{y}) - \nu_{i,j}(\mathbf{y}), \tag{2.1}$$

$$z' = \alpha(\mathbf{y}) \tag{2.2}$$

with initial conditions $y_{0,0}(0) = 1$ and $y_{i,j}(0) = 0$ otherwise, and $z(0) = 0$. Here

$$s(\mathbf{y}) = \sum_{0\leqslant i<d;\ 0\leqslant j<k} (d-i)\binom{k}{j} y_{i,j},$$

$$b_{i,j}(\mathbf{y}) = (d-i+1)j(y_{i-1,j} + y_{i-1,j-1})/\big(ks(\mathbf{y})\big) - (d-i)y_{i,j}/s(\mathbf{y}), \tag{2.3}$$

$$\alpha(\mathbf{y}) = (d-i_0)\big(1-F(\mathbf{y})\big)^{-1}, \tag{2.4}$$

$$\nu_{i,j}(\mathbf{y}) = \begin{cases} 1/\binom{k}{j} & \text{if } (i,j) = (i_0,j_0), \\ \alpha(\mathbf{y})b_{i,j}(\mathbf{y}) & \text{if } (i,j) \in B', \end{cases} \tag{2.5}$$

and

$$F(\mathbf{y}) = \sum_{(i,j)\in B'} (d-i)\binom{k}{j} b_{i,j}(\mathbf{y}). \tag{2.6}$$

https://doi.org/10.1017/S0963548306007954 Published online by Cambridge University Press

(The latter is the same as the quantity F_ℓ in (1.2) provided that ℓ is the highest priority of a basic type. It may help to observe that F is a function of B'.) Then t_1 is defined to be $\lfloor nx_0^- \rfloor$ for some $x_0^- < x_0$, where (see [17, equation (6.21)])

$$x_0 = \inf\{x > 0 \,:\, s(\mathbf{y}) = \epsilon \text{ or } F(\mathbf{y}) \geqslant 1 - \epsilon \text{ or } y_{i,j}(x) = 0 \text{ for all basic } (i,j)\}. \tag{2.7}$$

Equations (2.1) and (2.2) were shown to be equivalent to (1.4) at the end of [17, Section 6].

We return to the main thread of the argument. An edge, or the corresponding pair, is called *free* if it is incident with at least one uncoloured vertex. If the condition (1.7) holds we say that *sparseness* holds. It was also shown that if the sparseness holds at $z = z(x_0^-)$, and no unwanted cycle has occurred by time t_1, then the process **PComplete**$(k, \mathscr{S})$ succeeds with probability at least $1 - \epsilon_g + o(1)$, provided the dangerous types of vertices have priority over all others. This is because the graph induced by the free edges is sufficiently sparse at time t_1. It was also shown that the act of conditioning on unwanted cycles does not cause a problem for the analysis.

Thus, provided the sparseness holds at $z(x_0^-)$, the algorithm gives a proper k-colouring, with fixed $\mathscr{S}$, with probability at least $1 - \epsilon_g + o(1)$. Letting $g \to \infty$ and summing over all suitable $\mathscr{S}$ then gave the result in [17]. The value of x_0 arising when $d = 4$ produces the required sparseness property (for a suitable set of priorities).

For $d \geqslant 5$, it was clear that no set of priorities would do the job for the values of k of interest (such as those upper bounds in Theorem 1.1). Regardless of the priorities, vertices of type $(1, 1)$ will grow in number. The condition $F(\mathbf{y}) \geqslant 1 - \epsilon$ in (2.7) occurs long before the required sparseness condition holds, due simply to the growth in size of the variable $y_{1,1}$. (For some details when $d = 5$, see the start of Section 4.) When $F(\mathbf{y}) = 1$ we say that a change of phase has occurred, and what happens afterwards was not addressed in [17]. Typically, some other type of vertex of higher priority than $(1, 1)$ will start to proliferate.

Changes of phase were addressed in earlier applications of the method, such as [18]. At any typical part of a prioritized greedy algorithm (*i.e.*, one with several operations to choose from according to a priority list) such as **Colour**(k), vertices of some particular type (in this case, (i, j) say) will regularly be chosen as v in the algorithm and yet be (or become) plentiful in number. Vertices of types with higher priority may also be regularly chosen, but will be rare and regularly be used up entirely (at which point another vertex of type (i, j) again will be coloured. In this situation, we say informally that (i, j) is the *basic* type (which will correspond to the definition of basic type for the above equations), and a phase change is a point at which the basic type changes. A phase finishes when either the vertices of higher priorities than the current basic type become plentiful, or those with the current basic type become very scarce. We sometimes define the boundaries of the phases precisely in terms of the solution of systems of differential equations.

After a phase change, the same method of describing the algorithm in terms of the solutions of the differential equations is still valid. The equations change when the phase changes, since the equations, as given above, depend on the basic type. If the sparseness condition mentioned above is satisfied before any dangerous type becomes basic, the conclusion will be that the greedy colouring algorithm is a.a.s. successful. This is the idea that eventually leads to Theorem 1.2.

https://doi.org/10.1017/S0963548306007954 Published online by Cambridge University Press

When analysing a prioritized algorithm, we often deprioritize some of the choices, meaning that we choose between those operations with specific probabilities (and these probabilities in general will change through the course of the algorithm). The net result achieved by the algorithm (on random input) is (we believe) essentially the same as with prioritization. The advantage of deprioritizing is that the analysis becomes simpler during changes of phase, which eventually makes it possible to state a theorem without lots of conditions ensuring smooth phase transitions.

Unfortunately, in general the derivatives of the functions which model **PColour**(k) are not Lipschitz, and not even continuous, since at points where the basic priority changes, the derivatives of the old and new basic variables can jump. In related problems (see [18] and [21]), arguments have been used to show that the system in question makes a reasonably smooth transition at points where the derivative is discontinuous. However, such arguments require dividing the domain up into pieces on which the derivatives are continuous, together with verification of the value of derivatives or second derivatives at the points of transition. This verification is normally done numerically. For the problem at hand, the numerical work will be carried out to the limits of available accuracy. So this is an appropriate and useful time, for this line of research, to examine the solution of the systems arising in these problems without subdividing into separate systems with continuous derivatives.

In this way, systems of right-hand derivatives arise naturally, and a class of equations of the type we need to consider are treated in Section 3. Phases are described in more detail in Section 4, and then the resulting analysis is performed. We require considerable effort in Sections 3 and 4 because we are introducing a quite new theoretical method of examining the systems of differential equations that arise in problems such as this, especially for random regular graphs. These systems exhibit changes of phase, by which we mean points where the derivative is discontinuous. The analysis in Section 4 is aimed at a smoothed version of the solutions of such systems that arise by modifying the colouring process via partial deprioritization. In Section 5 the final numerical calculations are described.

3. A system with discontinuous derivatives

In this section we consider a slightly abstracted version of the system of differential equations (1.4) for use in Section 4. The variables are w_ℓ, $\ell \in \mathscr{L} = \{1, \dots, \ell_{\max}\}$, and as usual we write $\mathbf{w} = (w_1, \dots, w_{\ell_{\max}})$. The equations are

$$\frac{\mathrm{d}w_\ell}{\mathrm{d}z} = f_\ell^{\ell_0}(\mathbf{w}) = \begin{cases} g_\ell(\mathbf{w}) & \text{if } \ell < \ell_0, \\ 0 & \text{if } \ell > \ell_0, \\ F_{\ell_0-1}(\mathbf{w}) - 1 & \text{if } \ell = \ell_0, \end{cases} \tag{3.1}$$

where

$$F_\ell(\mathbf{w}) = \sum_{\ell' > \ell} g_{\ell'}(\mathbf{w}), \tag{3.2}$$

and

$$\ell_0 = \max\{R_1, R_2\}, \tag{3.3}$$

https://doi.org/10.1017/S0963548306007954 Published online by Cambridge University Press

where

$$R_1 = R_1(\mathbf{w}) = \max\{\ell : w_\ell > 0\}, \quad R_2 = R_2(\mathbf{w}) = \min\{\ell : F_{\ell'}(\mathbf{w}) < 1 \text{ for all } \ell' \geqslant \ell\}. \quad (3.4)$$

We will impose some conditions on the functions $g_\ell(\mathbf{w})$.

The correct interpretation of *solution* needs specifying. We take it to mean a continuous function whose right-hand derivative at any point is given by the right-hand side of (3.1). General theorems treating discontinuous derivatives are known but do not seem to apply to this situation; see Bressan [6], for example, for some conditions on existence of solutions almost everywhere. Nevertheless, we will prove an existence theorem, which is not totally necessary for our later arguments, but reduces complications later, besides adding to the general understanding of the situation.

There are various systems involved in (3.1), one system resulting when each value of ℓ_0 is held fixed. We call this system $\mathscr{E}_{\ell_0}$. Equations (3.3) and (3.4) specify which system is *active* (*i.e.*, defining the derivative of $\mathbf{w}$) for any given value of $\mathbf{w}$. When these definitions are in force, we refer to it as the *composite system* to distinguish it from the individual systems $\mathscr{E}_{\ell_0}$.

For the purposes of the present argument (and as is common when treating differential equations), a function f of $\mathbf{y}$ is *Lipschitz* on a given domain D if for some constant L, we have $|f(\mathbf{y}_1) - f(\mathbf{y}_2)| < L|\mathbf{y}_1 - \mathbf{y}_2|$. If each derivative y'_ℓ is given by a Lipschitz function f_ℓ then the system will have a unique solution.

We first note two simple results. These may comfort a reader who is not used to interpreting systems of differential equations as right-hand derivatives. The first result will be useful later. The second result, though not used explicitly, shows that wherever it is not obviously impossible due to discontinuity of the derivative, the solution of the system, as defined above, is differentiable.

Lemma 3.1. *If a continuous function f has a right-hand derivative which is bounded above by M on an interval $[a,b]$ then $f(b) - f(a) \leqslant M(b-a)$.*

Proof. If $f(b) - f(a) > (M+c)(b-a)$ with $c > 0$, let $S = \{x \in [a,b] : f(x) - f(a) \leqslant (M + c/2)(x-a)\}$. Since a is in the set but not b, the set S has a supremum, say x_0. Since f is continuous, $x_0 < b$. Since the right derivative of f at x_0 is at most M, we find that there are points between x_0 and b which are also in S, a contradiction. □

Lemma 3.2. *If f and g are continuous functions on an interval $[a,b]$ and g is the right-hand derivative of f on $[a,b]$, then f is differentiable on (a,b) and g is its derivative.*

Proof. For $\delta > 0$, let $M^+ = \sup\{g(x) : x \in [c-\delta, c]\}$. By Lemma 3.1, $(f(c) - f(c-\delta))/\delta \leqslant M^+$. Similarly $(f(c) - f(c-\delta))/\delta \geqslant M^- := \inf\{g(x) : x \in [c-\delta, c]\}$. Since g is continuous, both M^+ and M^- tend to $g(c)$ as $\delta \to 0$. □

It is interesting to compare these facts with the example given by Billingsley [3, Example 31.3], which is a continuous probability distribution function strictly increasing on the interval $[0,1]$ and yet with derivative equal to 0 almost everywhere. By Lemma 3.1,

https://doi.org/10.1017/S0963548306007954 Published online by Cambridge University Press

this function cannot have a continuous right derivative. (In fact it is not right differentiable everywhere.)

We may isolate the equation (3.1) with given fixed value of ℓ_0 as determining a system of differential equations in its own right, and we refer to this as $\mathscr{E}_{\ell_0}$. Of course (3.3) defines ℓ_0 as a function $\ell_0(\mathbf{w})$ for the system at hand. At any point, we call the variable w_{ℓ_0} *basic* and those w_ℓ with $\ell > \ell_0$ *super-basic* (these have derivative specified to be 0). It seems to be quite crucial, for the case we consider to be amenable to analysis, that, at points where ℓ_0 can increase, the function specifying the derivative of the new basic variable is 0.

For simplicity we take $z = 0$ as the initial point. The following theorem is ample for our needs. We do not need to prove uniqueness of the solution; in general, systems with non-Lipschitz right-hand sides do not have unique solutions.

Theorem 3.3. *Let $\mathscr{L}$ be finite. Suppose a differential equation system (referred to as the 'composite' system) is defined by the system $\mathscr{E}_{\ell_0}$ in variables w_ℓ, $\ell \in \mathscr{L}$, given in* (3.1) *and* (3.2) *where ℓ_0 (determining the 'active' system) is defined by* (3.3) *and* (3.4). *Assume furthermore that each $g_\ell \geqslant -O(w_\ell)$ and is Lipschitz on an open domain $\mathscr{D}_0$ in $(z, \mathbf{w})$-space, and that $\sum_{\ell \in \mathscr{L}} w_\ell > 0$ on $\mathscr{D}_0$. Then the composite system has a solution $\mathbf{w}^{(0)}$ given any nonnegative initial value (i.e., with $w_i \geqslant 0$ for all i) $\mathbf{w}^{(0)}(0) = \mathbf{w}_0$ in $\mathscr{D}_0$, and continuing to a point on the boundary of $\mathscr{D}_0$.*

Proof. We proceed by a method of successive approximations. The discontinuous nature of the derivatives requires special treatment at points where w_{ℓ_0} approaches 0. For this we judiciously alter the length of the intervals being used in the approximations.

Fix some $\delta > 0$, which we can imagine to be small. We will define $\mathbf{w}^{(\delta)}(t\delta)$, an approximation to $\mathbf{w}(t\delta)$, iteratively for $t = 1, 2, \ldots$. The definition is given by the usual one-step Euler approximation

$$E(\mathbf{w}, \ell, z, h) = w_\ell(z) + h f_\ell^{\ell_0(\mathbf{w}(z))}(\mathbf{w}(z)). \tag{3.5}$$

However, changes of phase require special alterations of the value of h, so we have defined the Euler approximation for a general step size h to accommodate the full definition.

Suppose that $\mathbf{w}^{(\delta)}$ is given at some point z. It turns out that the derivative can only be discontinuous near points where the derivative of $w_{\ell_0}^{(\delta)}$ is negative. So, to define $w_\ell^{(\delta)}(z + \delta)$, first put $z_0 = z$ and then for $i \geqslant 0$

$$z_{i+1} = \begin{cases} z + \delta & \text{if } Q \geqslant 0 \text{ or } -w_{\ell_0}^{(\delta)}(z_i)/Q \geqslant \delta, \\ z - w_{\ell_0}^{(\delta)}(z_i)/Q & \text{otherwise,} \end{cases} \tag{3.6}$$

where $Q = f_{\ell_0}^{\ell_0(\mathbf{w})}(\mathbf{w})$ for $\mathbf{w} = \mathbf{w}^{(\delta)}(z_i)$. Then define

$$w_\ell^{(\delta)}(z_{i+1}) = E(\mathbf{w}^{(\delta)}, \ell, z_i, z_{i+1} - z_i) \tag{3.7}$$

for each ℓ. Note first that this is nonnegative for $\ell = \ell_0$, since it is equal to 0 if $z_{i+1} \neq z + \delta$. Secondly, since $g_\ell > -O(w_\ell)$, we have for sufficiently small δ that $w_\ell^{(\delta)}(z_{i+1})$ is nonnegative for all $\ell < \ell_0$. The derivative is of course 0 for $\ell > \ell_0$. So we may assume nonnegativity of $w_\ell^{(\delta)}$ henceforth. We also observe that the sequence $z_0, z_1, \ldots$ terminates after a finite

https://doi.org/10.1017/S0963548306007954 Published online by Cambridge University Press

number of steps with $z_i = z + \delta$ for some i. This is because $z_{i+1} \neq z + \delta$ implies that $Q < 0$ and $w^{(\delta)}_{\ell_0}(z_{i+1}) = 0$, and thus $w^{(\delta)}_{\ell_0}(z_i) > 0$. For $z_{i+2} \neq z + \delta$, these conditions can only repeat themselves, with i replaced by $i+1$, if ℓ_0 is smaller at z_{i+1} than at z_i. It follows that the largest value, $i_{\max}$, of i, is bounded by the number $|\mathscr{L}|$ of variables. Thus iterations of (3.7) define $w^{(\delta)}_\ell(z+\delta) = w^{(\delta)}_\ell(z_{i_{\max}})$ given $w^{(\delta)}_\ell(z) = w^{(\delta)}_\ell(z_0)$. The definition of the vector function $\mathbf{w}^{(\delta)}$ is completed at all points between the z_i in this interval of length δ by linear interpolation.

The above defines $\mathbf{w}^{(\delta)}$ on the interval $[z, z+\delta]$ given its value at z. By iteration using $z = t\delta$, $t = 0, 1, 2, \ldots$, we obtain the definition of $\mathbf{w}^{(\delta)}$ extending to the boundary of $\mathscr{D}_0$.

We next fix δ and examine the difference between $\mathbf{w}^{(\delta)}$ and $\mathbf{w}^{(\delta/2)}$. Define a *full node* of $\mathbf{w}^{(\delta)}$ to be any point z_i in any of the intervals $[z, z+\delta]$ for $z = t\delta$ in the definition of $\mathbf{w}^{(\delta)}$ (including the node $z = t\delta$). Then between any two consecutive full nodes, the function is linear and has derivative determined, at its previous full node, via (3.1). Also define a *seminode* of $\mathbf{w}^{(\delta)}$ to be any full node of $\mathbf{w}^{(\delta/2)}$, and a *node* of $\mathbf{w}^{(\delta)}$ to be any of its full nodes or seminodes.

Assume that z is a node of $\mathbf{w}^{(\delta)}$ and

$$\sum_\ell |w^{(\delta)}_\ell(z) - w^{(\delta/2)}_\ell(z)| \leqslant \xi. \tag{3.8}$$

Let z_1 denote the greatest full node of $\mathbf{w}^{(\delta)}$ not greater than z, and similarly z_2 the greatest full node of $\mathbf{w}^{(\delta/2)}$ (or seminode of $\mathbf{w}^{(\delta)}$) not greater than z. Let ℓ_1 denote the value of $\ell_0(\mathbf{w}^{(\delta)})$ at z_1, and ℓ_2 similarly for $\ell_0(\mathbf{w}^{(\delta/2)})$ at z_2. Also define z' to be the node immediately following z. Then on the interval (z, z'), the derivative of $\mathbf{w}^{(\delta)}$ is given by the system $\mathscr{E}_{\ell_1}$ at z_1, and similarly for $\mathbf{w}^{(\delta/2)}$, the system $\mathscr{E}_{\ell_2}$, and z_2. By the Lipschitz property of each system, it follows that these derivatives are only $O(\delta)$ different from the derivatives given by $\mathscr{E}_{\ell_1}$ and $\mathscr{E}_{\ell_2}$ at z. Similarly, by the boundedness of (right) derivatives, the functions $\mathbf{w}^{(\delta)}$ and $\mathbf{w}^{(\delta/2)}$ themselves change by $O(\delta)$ between z_1 or z_2 and z. We will use these two observations repeatedly in the following, so call them the δ-*property*.

We wish to bound

$$\xi' = \sum_\ell |w^{(\delta)}_\ell(z') - w^{(\delta/2)}_\ell(z')|$$

in terms of

$$h = z' - z$$

and ξ. We consider cases depending on the relationship between ℓ_1 and ℓ_2.

Case 1: $\ell_1 = \ell_2$. Since each system has Lipschitz right-hand sides (and recalling the δ-property), the difference between derivatives of w_{ℓ_0} at $\mathbf{w}^{(\delta)}(z)$ and at $\mathbf{w}^{(\delta/2)}(z)$ is $O(\xi + \delta)$, thus the difference between the functions at z' is

$$|w^{(\delta)}_{\ell_0}(z') - w^{(\delta/2)}_{\ell_0}(z')| = |w^{(\delta)}_{\ell_0}(z) - w^{(\delta/2)}_{\ell_0}(z)| + O(h\xi + h\delta).$$

The details of this bound are quite routine. The corresponding result applies for the same reason to the functions w_ℓ for $\ell < \ell_0$ since the basic variable does not change. For $\ell > \ell_0$ of course the derivative is 0. Thus $\xi' \leqslant \xi + O(h(\xi + \delta))$.

https://doi.org/10.1017/S0963548306007954 Published online by Cambridge University Press

Case 2: $\ell_1 < \ell_2$. Write $D_z(f)$ for the (right) derivative of a function f at z. By the definition of ℓ_0 as the maximum of R_1 and R_2 in (3.4), and recalling that the derivative of every super-basic variable is 0, we have for each $\ell_1 < \ell \leqslant \ell_2$

$$w_\ell^{(\delta)}(z) = D_z\big(w_\ell^{(\delta)}(z)\big) = 0. \tag{3.9}$$

It then follows from the theorem hypotheses that for such ℓ,

$$g_\ell\big(\mathbf{w}^{(\delta)}(z)\big) \geqslant 0, \tag{3.10}$$

and we also have

$$w_\ell^{(\delta/2)}(z) = O(\xi). \tag{3.11}$$

Using (3.9), the δ-property and the nonnegativity of $w_\ell^{(\delta/2)}$, and then later the definition (3.2) of F_ℓ,

$$\begin{aligned} D_z\left(\sum_{\ell=\ell_1}^{\ell_2} |w_\ell^{(\delta/2)} - w_\ell^{(\delta)}|\right) &= D_z\left(|w_{\ell_1}^{(\delta/2)} - w_{\ell_1}^{(\delta)}| + \sum_{\ell=\ell_1+1}^{\ell_2} w_\ell^{(\delta/2)}\right) \\ &= |F_{\ell_1-1}\big(\mathbf{w}^{(\delta)}(z)\big) - 1 - g_{\ell_1}\big(\mathbf{w}^{(\delta/2)}(z)\big)| \\ &\quad + \sum_{\ell=\ell_1+1}^{\ell_2-1} g_\ell\big(\mathbf{w}^{(\delta/2)}(z)\big) + F_{\ell_2-1}\big(\mathbf{w}^{(\delta/2)}(z)\big) - 1 + O(\delta) \\ &\leqslant |F_{\ell_1}\big(\mathbf{w}^{(\delta)}(z)\big) - 1| + |g_{\ell_1}\big(\mathbf{w}^{(\delta)}(z)\big) - g_{\ell_1}(\mathbf{w}^{(\delta/2)}(z))| \\ &\quad + F_{\ell_1}\big(\mathbf{w}^{(\delta/2)}(z)\big) - 1 + O(\delta). \end{aligned} \tag{3.12}$$

The term $F_{\ell_1}\big(\mathbf{w}^{(\delta)}(z)\big) - 1$ is negative by the definition of R_2 and ℓ_1, and so finally the bound (3.12) is $O(\delta + \xi)$. Hence this group of variables w_ℓ behaves collectively like a single variable in Case 1, and the other variables individually behave likewise, and the same conclusion is reached.

Case 3: $\ell_2 < \ell_1$. This is covered by exactly the same argument as in Case 2, switching the roles of $\mathbf{w}^{(\delta)}$ and $\mathbf{w}^{(\delta/2)}$.

We have in all cases $\xi' \leqslant \xi + O\big(h(\xi + \delta)\big) = \xi \exp\big(O(h)\big) + O(h\delta)$ since $h \leqslant \delta$. Thus by induction on the nodes z beginning with $z = 0$, the error ξ given in (3.8) is $O(\delta e^z)$. The approximants $\mathbf{w}^{(\delta)}$ hence converge (uniformly) on the bounded domain $\mathscr{D}_0$, and so there is a uniform limiting function which we may denote by $\mathbf{w}^{(0)}$.

It is immediate that $\mathbf{w}^{(0)}$ is continuous, by the uniform bound (irrespective of δ) on the derivative of $\mathbf{w}^{(\delta)}$ (at its points of differentiability). All that remains is to verify the claim that $\mathbf{w}^{(0)}$ has right-hand derivative given by (3.1) at every point z_0 interior to $\mathscr{D}_0$.

First, let ℓ_0 denote $\ell_0\big(\mathbf{w}^{(0)}(z_0)\big)$. Then $w_\ell^{(0)}(z_0) = 0$ for $\ell > \ell_0$. By nonnegativity of $w_\ell^{(\delta)}$ and the condition on g_ℓ, it follows that $F_\ell\big(\mathbf{w}^{(0)}(z_0)\big)$ is monotonic non-increasing, as a function of ℓ, for $\ell \geqslant \ell_0$, and we also have $F_{\ell_0}\big(\mathbf{w}^{(0)}(z_0)\big) < 1$. For $\ell > \ell_0$, consider the value of $w_\ell^{(0)}(z_0 + h)$ when $h > 0$ is arbitrarily small. Fixing h for the next part of the discussion, first note that $F_{\ell-1}\big(\mathbf{w}^{(\delta)}(z_0)\big) - 1$ is strictly less than $\big(F_{\ell_0}(\mathbf{w}^{(0)}(z_0)) - 1\big)/2 < 0$ for δ sufficiently small (h fixed) by the bound on variation of $\mathbf{w}^{(\delta)}$ and by the approximability of $\mathbf{w}^{(0)}$

https://doi.org/10.1017/S0963548306007954 Published online by Cambridge University Press

by $\mathbf{w}^{(\delta)}$. It follows that for arbitrarily small $\epsilon > 0$ (with ϵ fixed in the following statement):

$$\text{for } \delta \text{ sufficiently small, } w_\ell^{(\delta)}(z_0 + \epsilon h) = 0 \ (\ell > \ell_0) \text{ and } \ell_0\big(\mathbf{w}^{(\delta)}(z_0 + \epsilon h)\big) \leqslant \ell_0. \quad (3.13)$$

Further use of the bounded variation properties shows that $w_\ell^{(\delta)}$ remains 0 in the interval $[z_0 + \epsilon h, z_0 + h]$. By the approximability properties, $w_\ell^{(0)}(z_0 + h) = 0$, and so $w_\ell^{(0)}$ has right-hand derivative 0 at z_0, as required to satisfy (3.1).

It only remains to examine the right derivative of $w_\ell^{(0)}$ for $\ell \leqslant \ell_0$, which we split into two cases.

Case i: $w_{\ell_0}^{(0)}(z_0) > 0$. Here, using arguments as above, we see from (3.13) that for δ, h and $\epsilon > 0$ sufficiently small, $w_{\ell_0}^{(\delta)}(z_0 + \epsilon h) > 0$, $\ell_0\big(\mathbf{w}^{(\delta)}(z_0 + \epsilon h)\big) = \ell_0$ and only one system $\mathscr{E}_{\ell_0}$ is active for $\mathbf{w}^{(\delta)}(z)$ for $z \in [z_0 + \epsilon h, z_0 + h]$. The claim then follows immediately from the good approximation of $\mathbf{w}^{(0)}$ by $\mathbf{w}^{(\delta)}$, since such $\mathbf{w}^{(\delta)}$ have well-behaved gradients at their differentiable points inside such a neighbourhood.

Case ii: $w_{\ell_0}^{(0)}(z_0) = 0$. Then, by the definition of ℓ_0, we have $F_{\ell_0 - 1}\big(\mathbf{w}^{(0)}(z_0)\big) \geqslant 1$. First suppose that $F_{\ell_0 - 1}\big(\mathbf{w}^{(0)}(z_0)\big) > 1$ or $\mathbf{w}_{\ell_0 - 1}^{(0)}(z_0) > 0$. Recall by definition that $F_{\ell_0}\big(\mathbf{w}^{(0)}(z_0)\big) < 1$. Then, arguing as for (3.13) and Case i, we have as in Case i, $\ell_0\big(\mathbf{w}^{(\delta)}(z_0 + \epsilon h)\big) = \ell_0$ for δ, ϵ and h sufficiently small, and the argument is virtually the same as Case i, so we omit the further details. If on the other hand $F_{\ell_0 - 1}\big(\mathbf{w}^{(0)}(z_0)\big) = 1$ and $\mathbf{w}_{\ell_0 - 1}^{(0)}(z_0) = 0$, then define ℓ_1 to be the maximum $\ell < \ell_0$ for which $F_{\ell - 1}\big(\mathbf{w}^{(0)}(z_0)\big) > 1$ or $\mathbf{w}_{\ell - 1}^{(0)}(z_0) > 0$. By the theorem's hypothesis that $\sum_{\ell \in \mathscr{L}} \mathbf{w}_\ell > 0$ on $\mathscr{D}_0$, such an ℓ exists.

Again by the bounded variation properties (and the arguments above which exclude $\ell > \ell_0$ becoming active), in the appropriately defined intervals $[z_0 + \epsilon h, z_0 + h]$ (as above), the only possible active systems are just $\mathscr{E}_\ell$ for $\ell_1 \leqslant \ell \leqslant \ell_0$. (Moreover, it seems possible that the active system does indeed change on the trajectory of $\mathbf{w}^{(\delta)}$, in arbitrarily small neighbourhoods of z_0, even for arbitrarily small δ, due to F_ℓ dropping below 1 for one of these values of ℓ.) Note that by definition of F, it must be that $g_\ell = 0$ for $\ell_1 < \ell < \ell_0$ (referring to $\mathbf{w}^{(0)}(z_0)$). It is now easy to check that the gradients specified by (3.1) are (on the same appropriate intervals as above) arbitrarily close to 0 for $\ell > \ell_1$ and $g_{\ell_0}(\mathbf{w}^{(0)}(z_0))$ for $\ell = \ell_1$, which shows, using the ideas above, that $w_\ell^{(0)}$ has the desired right-hand derivative at z_0. The case $\ell < \ell_1$ follows as before since for such ℓ the derivative specified in (3.1) does not depend on which of the systems is active. □

4. Phases and deprioritization

For $d > 4$, we are not as lucky as with $d = 4$. With $d = 5$ and $k = 4$ we may start with $(1, 1)$ as the basic type, and it is sensible to allocate lower priority to (2,1), (3,1) and (4,1), and higher priority to (2,2), (3,2) and (4,2). The resulting differential equations (2.1) and (2.2) can be solved numerically, and the point x_0 in (2.7) is reached quite quickly (using $\epsilon \approx 0$), well before the sparseness condition arising from (1.7) can hold. At x_0, the expression on the left side of (1.7) is

$$\frac{\sum_{i,j}(d-i)(d-i-1)\binom{k}{j}y_{i,j}(x_0)}{\sum_{i,j}(d-i)\binom{k}{j}y_{i,j}(x_0)} \approx 3.916319 > 1.$$

https://doi.org/10.1017/S0963548306007954 Published online by Cambridge University Press

At the point x_0, $F(\mathbf{y})$ basically reaches 1 (see (2.7)), which we loosely call the *explosion* condition, as the time between processing vertices of basic type explodes. Now, another phase will occur: the basic type is no longer processed and the lowest priority type that is commonly processed should be renamed as basic. In general, which type becomes basic depends on the priority list. Under the list to be defined in Section 5, in the second phase, the type (2,2) becomes basic for $(d,k) = (5,4)$, whereas (3,2) becomes basic for $(d,k) = (6,4)$ and $(d-2,2)$ becomes basic for $d > 6$ and $k > 4$. We can then use essentially the same argument as before, except that there is a gap near the phase change-over where $F(\mathbf{y})$ exceeds 1 and so the conditions necessary for analysis of the first phase do not occur. This causes some difficulty. In this section, we will deal with these gaps and later phases all in one go by altering the algorithm, using the approach from [21] (with a slight modification).

We henceforth assume that $k \geqslant 4$. Even $k \geqslant 3$ ensures that the type $(1,1)$ is not dangerous. Where appropriate, the results will hypothesize that all dangerous types have higher priority than all non-dangerous types.

From this point onwards, we assume complete familiarity with [17], and refer to the equations and other things from that paper concisely as, for example, (4.18):[17] and Section 4:[17].

As described in Section 2, we are interested in the system (1.4), which arose first in the guise of (4.20):[17]. This equation, if we ignore (1.5), specifies a different system for each ℓ_0; in the setting of (4.20):[17] this corresponds to different types (i_0, j_0) being regarded as basic; as in Section 3, we call this system $\mathscr{E}_{\ell_0}$ or $\mathscr{E}_{i_0,j_0}$. We give a differential equation for the vector-valued function $\mathbf{y}$ by specifying which system is *active* (*i.e.*, defining the derivative of $\mathbf{y}$) for any given value of $\mathbf{y}$. Initially, $y_{0,0}(0) = 1$, and $y_{i,j}(0) = 0$ for all other i and j. For $z > 0$, system $\mathscr{E}_{i_0,j_0}$ will be active if and only if (i_0, j_0) is the basic type, defined immediately below. This gives a composite system as considered in Section 3.

For the purpose of defining the composite system, we require a precise theoretical definition of the 'basic type'. Recall from Section 1 that for convenience, we use y_ℓ to denote that variable $y_{i,j}$ for which $\mathtt{prio}(i,j) = \ell$, and use this to determine i and j as functions $i(\ell)$ and $j(\ell)$ of ℓ. In other contexts, ℓ and (i,j) or i,j are used interchangeably to stand for types, for purposes of convenience, where it should cause no confusion. Recalling (1.1):

$$R_1 = R_1(\mathbf{y}) = \max\{\ell : y_\ell > 0\}, \quad R_2 = R_2(\mathbf{y}) = \min\{\ell : F_{\ell'}(\mathbf{y}) < 1 \text{ for all } \ell' \geqslant \ell\}, \quad (4.1)$$

where

$$F_\ell(\mathbf{y}) = \sum_{\ell' > \ell} \big(d - i(\ell')\big) \binom{k}{j(\ell')} b_{\ell'}(\mathbf{y}).$$

(Note this is the same function F as defined in (2.6) provided that ℓ is the basic priority.) Here R_1 can be used to ensure that the basic type of vertex is available to be chosen as the next vertex in the process, and R_2 can be used to detect the explosion condition. So the *basic type* (i_0, j_0) (or ℓ_0) is defined to be the one such that

$$\ell_0 = \mathtt{prio}(i_0, j_0) = \max\{R_1, R_2\} \quad (4.2)$$

in accordance with (3.3) and (1.5). Thus, the basic type is simply a function $\ell_0(\mathbf{y})$. It is easy to see, though not required for our argument, that for any priority list with type

https://doi.org/10.1017/S0963548306007954 Published online by Cambridge University Press

$(1,1)$ ranked above $(0,0)$, the basic type defined by (4.2) will initially be $(1,1)$. For, letting $\ell = \texttt{prio}(1,1)$, if $\ell > \texttt{prio}(0,0)$, we have $F_\ell(\mathbf{y}(0)) = 0 < 1$ and $F_{\ell-1}(\mathbf{y}(0)) = d - 1 > 1$.

We will need to consider the relationship between the systems $\mathscr{E}_\ell$. Note that all these systems are well behaved on any bounded open domain $\mathscr{D}_0$ in $(z, \mathbf{y})$-space with s bounded away from 0 (even including negative values of the y_ℓ). Of course, the 'time' variable, z, can be ignored in defining the domain since the systems are all autonomous. We assume for the following discussion of the solution of the composite system that a given point $\mathbf{y}$ is inside such a domain $\mathscr{D}_0$. Other domain-type restrictions will be imposed on the process later.

Note that the equation in each system $\mathscr{E}_\ell$ on $\mathscr{D}_0$ has a Lipschitz right-hand side because the specified derivative is given by a polynomial in the y_ℓ, with bounded coefficients. (The significant denominator is $d - 2z$, which is bounded away from 0 on $\mathscr{D}_0$.)

The differential equation method as in [19] requires Lipschitz derivatives, so the arguments in Section 6:[17] do not apply directly to (1.4). Thus, at this point the applicability of equations (1.4) with the auxiliary definition of ℓ_0 in (1.5) have no direct rigorous interpretation for the process. To overcome this difficulty we use partial deprioritization. To obtain a general deprioritized process, those types (i, j) which are not dangerous but have basic, or higher, priority are treated in a different way: instead of selecting such a type in **PColour**(k) according to priority, it is chosen with a specified probability $\hat{\phi}_{i,j}(t)$. Defining these probabilities in any continuous manner gives what we call a general deprioritized algorithm or process, not necessarily tuned to the original one; the resulting differential equation determines what result is obtained. By using the proportions $\phi_{i,j}(t)$ calculated in Section 4:[17] to define $\hat{\phi}_{i,j}(t)$ as below, we will obtain a sequence of such general deprioritized processes whose behaviour approaches arbitrarily closely to the behaviour indicated by the equations we derived (heuristically) relating to the original process. We will call this sequence of processes the *partially deprioritized process* **PColour***(k). Note that it depends on the specified functions $\hat{\phi}_{i,j}(t)$.

One difficulty with this general idea is that the selection of a vertex of the desired type (at every step) may not be possible unless there is a good supply of such vertices available. This might seem to contradict the idea that the ones above basic priority always run out at the end of a clutch (in the original process). However, these vertices can be maintained in numbers large enough to avoid the problem, but nevertheless small enough so as not to affect the final results. Additionally, this deprioritization should not apply to the dangerous types, or they would build up in numbers and cause problems. So dangerous types will continue to be prioritized. The whole plan requires an initial phase called a *preparatory* phase, to build up the numbers of vertices of non-dangerous types. In the preparatory phase, the first step is to process a random vertex of type $(0,0)$. After that, the non-dangerous types (i, j) have the specified probability $\hat{\phi}_{i,j} = 0$ except for $(i, j) = (1,1)$, and the dangerous types as usual are processed with (their specified) high priority. Thus, after the initial step (and provided the vertices of type $(1,1)$ do not become exhausted) this is equivalent to rearranging (for the duration of this phase only) the priorities of the process **PColour**(k) so that $(1,1)$ has the highest priority of the non-dangerous types, and all dangerous types have priority higher than that. We specify that, for some suitable small constant ϵ_1, the preparatory phase runs for $t_1 = \lfloor \epsilon_1 n \rfloor$ steps. Thus ϵ_1 becomes a parameter which is input for **PColour***(k). Section 7:[17] shows that the number of vertices of type

https://doi.org/10.1017/S0963548306007954 Published online by Cambridge University Press

$(1,1)$ a.a.s. does not reach 0 before time t_1, and thus it is irrelevant what is prescribed to cover this eventuality: for instance we may stop the process.

The state of the process at the end of the preparatory phase is given in the following.

Lemma 4.1. *Suppose that the preparatory phase of* **PColour***(k) *is begun with an initial pairing either empty or containing just the pairs in* $\mathscr{S}$. *For* ϵ_1 *a sufficiently small constant, a.a.s.* $Y_{i,C}(t_1) = q_\ell n + o(n)$ *for each* (i, C), *where each* q_ℓ *is a positive constant depending on* ϵ_1, *and* ℓ *denotes* $(i, |C|)$. *Similarly, a.a.s.* $Z(t_1) = \hat{\epsilon}_1 n + o(n)$ *where* $\hat{\epsilon}_1$ *is a positive constant.*

Proof. By Lemma 7.1:[17], the type $(1,1)$ is common. We can thus apply Corollary 6.4:[17], referring to the definition of the functions $y_{i,j}$ satisfying the differential equation (6.16):[17]. We wish to find a lower bound for x_0. At $x = 0$ we have $y_{0,0} = 1$ and the other y are 0, so $s(\mathbf{y}) = d$ and in view of (4.14):[17] the only positive value of $b_{i,j}$ is $b_{1,1} = 1/k$. Thus with $\ell = (1,1)$ we have $F_{\ell-1} = (d-1)kb_\ell = d - 1 > 1$, and hence the derivative of $y_{1,1}$ is strictly positive. From this it follows by induction on type that $y_{\ell'}$ has positive derivative for all other non-dangerous types ℓ'. The lemma thus follows from Corollary 6.4:[17]. □

After the preparatory phase, all priorities now revert to their original definitions. Define (with reference to Section 6:[17]) B^- to be the set of non-dangerous types. Thus, all non-dangerous types are lumped together into one super-type called *expanded basic*. Each non-dangerous type is now called basic. Regarding priorities, **PColour***(k) operates on the expanded basic type in the same way that **PColour**(k) operates on the basic type, and we will apply this with all dangerous types, which are treated exactly as before, having priority higher than the expanded basic. During the execution of **PColour***(k), when there is no vertex of type higher than basic, a vertex of expanded basic type is processed. The type of this vertex is first chosen randomly, type (i, j) being chosen in proportion to some function $\hat{\phi}_{i,j}$ (which may be a function of $\mathbf{Y}$), and this type is then processed in the usual manner.

Before defining $\hat{\phi}_{i,j}$, to ensure that it is Lipschitz, we introduce an extra smoothing parameter η, which will also become part of the input to **PColour***(k). Our main difficulty is to show how to define $\hat{\phi}_{i,j}$ as a function of η so that the solution of the resulting differential equations converges to a suitable function, as $\eta \to 0$. The function $\hat{\phi}_{i,j}$ is defined similarly to $\phi_{i,j}$ given in (4.5):[17] and (4.6):[17] (note that this gave rise to the function $\theta_{i,j}(\mathbf{y})$ given in (4.17):[17], that implicitly appears in (4.20):[17]), apart from the smoothing, which blurs the distinction between phases. It will use a function λ to weight the types, defined recursively, from largest $\ell \in B^-$ to smallest, by

$$\lambda(\ell) = \max\{0, \min\{\overline{Y}_\ell/(\eta_\ell n) - 1, 1\}\}\left(1 - \sum_{\ell' > \ell, \ell' \in B^-} \lambda(\ell')\right)$$

where $\eta_\ell = \eta d^{-3i(\ell)}$. Note that λ depends upon both $\mathbf{Y}$ and η.

We will impose some extra restrictions, (D1) and (D2), on the values of the vector $\mathbf{Y}$; if the process reaches a point where this is violated, its further steps are unspecified (as

https://doi.org/10.1017/S0963548306007954 Published online by Cambridge University Press

we will show, this will happen rarely enough to be ignored). The first restriction is

$$\text{(D1)} \qquad Y_{0,\emptyset} > 2\eta n, \qquad \overline{Y}_\ell > \frac{1}{2}\eta_\ell n \quad \forall \ell \in B^-. \tag{4.3}$$

From the first condition here and by scrutiny of the definition of λ,

$$\sum_{\ell \in B^-} \lambda(\ell) = 1, \quad \lambda_\ell \geqslant 0 \text{ for all } \ell \in B^-, \tag{4.4}$$

(the first condition holding since the type $\ell = (0,0)$ always takes up the slack if the sum of $\lambda(\ell)$ of the higher priority types is less than 1) and

$$\text{for } \ell \in B^-,\ \lambda(\ell) = 0 \text{ if } \ell < \ell' \text{ for some } \ell' \in B^- \text{ with } \overline{Y}_{\ell'} \geqslant 2\eta_\ell n. \tag{4.5}$$

Before proceeding, we need the analogue of $F_\ell(\mathbf{y})$ defined as a function of $\mathbf{Y}$. Recalling that $r_{i,j}$ from (4.7):[17] is written in terms of $\overline{\mathbf{Y}}$ and hence is a function of $\mathbf{Y}$, define

$$\overline{F}_\ell(\mathbf{Y}) = \sum_{\ell' > \ell} c(\ell') r_{\ell'}(\mathbf{Y}), \tag{4.6}$$

with $c(\ell')$ defined as in (1.2).

The function $\phi_{i,j}(\mathbf{Y})$ as given in (4.5):[17] and (4.6):[17] actually depends on (i_0, j_0), so we denote it now by $\phi_{i,j}^{(\ell_1)}(\mathbf{Y})$ with $\ell_1 = (i_0, j_0)$. We do not use ℓ_0 here because that is a special function determined by the current value of $\mathbf{Y}$, whilst the definition of $\phi_\ell^{(\ell_1)}(\mathbf{Y})$ we are making now is applicable for all types $\ell_1 \leqslant \ell$ for every $\mathbf{Y}$. At this point we extend that definition by declaring that $\phi_\ell^{(\ell_1)}(\mathbf{Y}) = 0$ if $\ell < \ell_1$. We will make no use of $\phi_\ell^{(\ell_1)}$ for those ℓ_1 such that $\overline{F}_{\ell_1} \geqslant 1$ for the purpose of defining $\hat{\phi}_{i,j}$, so the condition $\overline{F}_{\ell_1} < 1$ will appear often in the following.

We need to check that $\phi_{i,j}^{(\ell_1)}(\mathbf{Y})$ is clearly nonnegative. Its definition uses the function T defined in (4.10):[17]. For the context of this, along with (4.5):[17] and (4.6):[17], the value of (i_0, j_0) determines B'; for *any* value ℓ_1 of (i_0, j_0) a function $T^{(\ell_1)}$ results.

We require $T^{(\ell_1)}$ to be nonzero, and we can note at this point that (for ϵ and η sufficiently small)

$$\text{for all } \ell_1, \quad T^{(\ell_1)} > \epsilon \quad \text{provided that} \quad \overline{F}_{\ell_1} < 1,\ S > \epsilon n \text{ and } \lambda(\ell_1) > 0 \tag{4.7}$$

where S is defined in (4.1):[17]. This follows because, from (4.10):[17],

$$T^{(\ell_1)} = 1 - \overline{F}_{\ell_1} + \sum_{(i,j) > \ell_1} (d - i_0) \binom{k}{j} r_{i,j}. \tag{4.8}$$

By (4.5), if $\lambda(\ell_1) > 0$, then for $\ell_1 < \ell \in B^-$, $\overline{Y}_\ell < 2\eta_\ell n$ and thus by (4.7):[17] and (4.3), ℓ cannot equal $(0,0)$ and furthermore $r_\ell > 0$. Considering (4.6) and the fact that types (i,j) with $i = d$ have negative priority and can be ignored, $\overline{F}_{\ell_1}$ is hence bounded above by d times the value of the summation in (4.8). From this, (4.7) follows.

Given the conditions in (4.7), we deduce T in (4.10):[17] is positive and so by Lemma 4.1:[17],

$$\phi_{\ell_1}^{(\ell_1)} = (1 - \overline{F}_{\ell_1}) \left(T \binom{k}{j_0} \right)^{-1} > 0. \tag{4.9}$$

https://doi.org/10.1017/S0963548306007954 Published online by Cambridge University Press

It remains to consider $\phi_\ell^{(\ell_1)}$ for $\ell > \ell_1$, $\ell \in B^-$. Recall from just above that we can assume $r_\ell > 0$. Therefore (recalling that $\phi_\ell^{(\ell_1)} = 0$ for $\ell < \ell_1$)

$\phi_\ell^{(\ell_1)}$ is well defined and nonnegative provided that $\overline{F}_{\ell_1} < 1$, $S > \epsilon n$ and $\lambda(\ell_1) > 0$. (4.10)

Define $\overline{R}_2$ as R_2 in (3.4) but with F replaced by $\overline{F}$. We define $\hat{\phi}_{i,j}$ to be the following convex combination of the $\phi_{i,j}^{(\ell_1)}$ functions:

$$\hat{\phi}_{i,j}(\mathbf{Y}) = \sum_{\ell_1 \in B^-} \lambda(\ell_1) \phi_{i,j}^{(m(\ell_1))}(\mathbf{Y}), \tag{4.11}$$

where

$$m(\ell) = \max\{\ell, \overline{R}_2\}.$$

This is well defined by (4.10), since $\overline{F}_{m(\ell)} < 1$ always by definition. We will need the following result.

Lemma 4.2. *Provided $S > \epsilon n$, $\hat{\phi}_{i,j}$ is Lipschitz.*

Proof. In Section 6:[17] the functions μ were obtained by solving a linear system which has determinant 0 when F is 1 (see (6.18):[17]), and this was transformed to the equations for $\phi_{i,j}$ obtained in Section 4:[17] which are well defined. (To avoid the singularity, F is bounded away from 1 in (6.21):[17].) So it is not immediate that $\phi_{i,j}^{(\ell_1)}$ is Lipschitz, or even continuous. However, from the observations leading to (4.10) we deduce that $\phi_\ell^{(\ell_1)}$ is Lipschitz provided $\lambda(\ell_1) > 0$ and $\overline{F}_{\ell_1} < 1$. The condition $\lambda(\ell_1) > 0$ is irrelevant for $\hat{\phi}_{i,j}$ because $\phi_{i,j}^{(m(\ell_1))}$ is bounded and multiplied by $\lambda(\ell_1)$. Moreover, as already observed, $\overline{F}_{m(\ell)} < 1$ always. Thus, to complete the proof, we only need continuity of the function $\phi_{i,j}^{(m(\ell_1))}$ (where (i,j) and ℓ_1 are fixed), for which it suffices to show that

$$\overline{F}_{\ell_1} = 1 \quad \text{implies} \quad \phi_{i,j}^{(\ell_1)} = \phi_{i,j}^{(\ell_1+1)}. \tag{4.12}$$

If $\overline{F}_{\ell_1} = 1$, we have from (4.10):[17] and (4.6), with $\ell_1 = (i_0, j_0)$, $\ell_1 + 1 = (i_0^+, j_0^+)$ and $\ell' = (i', j')$,

$$\begin{aligned} T^{(\ell_1)} &= \sum_{\ell' > \ell_1} c(\ell') r_{\ell'} + (i' - i_0) r_{\ell'} \binom{k}{j'} \\ &= \sum_{\ell' > \ell_1} (d - i_0) \binom{k}{j'} r_{\ell'}, \\ T^{(\ell_1+1)} &= c(\ell_1 + 1) r_{\ell_1+1} + \sum_{\ell' > \ell_1+1} c(\ell') r_{\ell'} + (i' - i_0^+) r_{\ell'} \binom{k}{j'} \\ &= \sum_{\ell' > \ell_1} (d - i_0^+) \binom{k}{j'} r_{\ell'}, \end{aligned}$$

and thus for $(i,j) > \ell_1 + 1$,

$$\phi_{i,j}^{(\ell_1)} = \frac{d - i_0}{T^{(\ell_1)}} r_{i,j} = \frac{d - i_0^+}{T^{(\ell_1+1)}} r_{i,j} = \phi_{i,j}^{(\ell_1+1)}.$$

From this we obtain (4.12) in all cases, given that $\phi_{\ell_1}^{(\ell_1)}$ is always determined from the other $\phi_{i,j}^{(\ell_1)}$ via (4.6):[17]. □

https://doi.org/10.1017/S0963548306007954 Published online by Cambridge University Press

To complete the definition of **PColour***(k), for a step at which the type of every vertex in the graph is expanded basic, the type $(i,j) \in B^-$ is chosen in proportion to $\binom{k}{j}\hat{\phi}_{i,j}(\mathbf{Y})$, *i.e.*, with probability

$$p_{i,j} = \frac{\binom{k}{j}\hat{\phi}_{i,j}(\mathbf{Y})}{\sum_{(i',j')\in B^-}\binom{k}{j'}\hat{\phi}_{i',j'}(\mathbf{Y})}. \tag{4.13}$$

To ensure that this is a well-defined probability, we first make the second domain restriction that

$$\text{(D2)} \qquad \overline{F}_{\ell_2}(\mathbf{Y}) < 1-\epsilon \tag{4.14}$$

where

$$\ell_2 = \max B^-. \tag{4.15}$$

Then, provided $S > \epsilon n$, (4.9) and (4.10) ensure that the numerator of (4.13) is nonnegative (noting that only ℓ_1 with $\lambda(\ell_1) > 0$ is relevant in (4.11), and for such ℓ_1, $\overline{F}_{m(\ell_1)} < 1$). We also note that for sufficiently small $\epsilon_2 > 0$ we have

$$\sum_{(i',j')\in B^-}\binom{k}{j'}\hat{\phi}_{i',j'}(\mathbf{Y}) > \epsilon_2. \tag{4.16}$$

This is for similar reasons, with the additional observation that $\phi_{\ell_2}^{(\ell_2)}$ is bounded below by a positive function of ϵ using (4.14) and the fact that T is bounded above in (4.9).

Here is a rough informal description of what happens in **PColour***(k). By (4.3), all $\overline{Y}_\ell$ will remain positive. Those which drop below $2\eta_\ell n$ for $\ell \geqslant R_2$ will have a small proportion of participation as an initial clutch step via the formula (4.13) and will continue to drop close to $\eta_\ell n$, provided $\overline{F}_{\ell_1} < 1$ for some $\ell_1 < \ell$. For all other initial clutch steps, the type chosen is the one of highest priority amongst those above $2\eta_\ell n$.

To analyse **PColour***(k) precisely, we re-examine Section 6:[17], recalling the new definition of B^- as the set of all non-dangerous types. The expanded basic type replaces the basic type, and so a few modifications are required for definitions. Now B' is the set of dangerous types. For the approximation result we need to state the (symmetrical) system of differential equations corresponding to (6.16):[17] and (6.17):[17] which will be referred to in the following lemma. These have more or less the same definitions of $\alpha(\mathbf{y})$ and $b_{i,j}(\mathbf{y})$, but require different $v_{i,j}(\mathbf{y})$ for $(i,j) \in B^-$ since the set of the basic types B^- is now expanded to all non-dangerous types:

$$y'_{i,j} = \alpha(\mathbf{y})b_{i,j}(\mathbf{y}) - v_{i,j}(\mathbf{y}), \tag{4.17}$$

$$z' = \alpha(\mathbf{y}), \tag{4.18}$$

where

$$\alpha(\mathbf{y}) = \sum_{(i,j)\in B}(d-i)\binom{k}{j}v_{i,j}(\mathbf{y}), \tag{4.19}$$

$$v_{i,j}(\mathbf{y}) = \alpha(\mathbf{y})b_{i,j}(\mathbf{y}) \text{ if } (i,j) \in B', \tag{4.20}$$

https://doi.org/10.1017/S0963548306007954 Published online by Cambridge University Press

$$v_{i,j}(\mathbf{y}) = \sigma(\mathbf{y})\hat{\theta}_{i,j}(\mathbf{y}) \text{ if } (i,j) \in B^-, \tag{4.21}$$

$$\sigma(\mathbf{y}) = \left(\sum_{(i,j)\in B^-} \binom{k}{j} \hat{\theta}_{i,j}(\mathbf{y}) \right)^{-1} \text{ and } \hat{\theta}_{i,j}(\mathbf{y}) = \sum_{\ell \in B^-} \hat{\lambda}(\ell)\theta_{i,j}^{(m(\ell))}(\mathbf{y}), \tag{4.22}$$

$$\hat{\lambda}(\ell) = \max\{0, \min\{y_\ell/\eta_\ell - 1, 1\}\} \left(1 - \sum_{\ell' > \ell, \ell' \in B^-} \hat{\lambda}(\ell') \right),$$

$m(\ell)$ is defined (only a slight re-definition) as $\max\{\ell, R_2\}$, and the function $\theta_{i,j}(\mathbf{y})$ as given in (4.17):[17] appears now with an extra index $m(\ell)$ since its definition depends, just as ϕ does, on the type $m(\ell)$.

In analogy to Corollary 6.4:[17] we have the following, defining **PComplete***$(k, \mathscr{S})$ from **PColour***(k) as **PComplete**$(k, \mathscr{S})$ was defined from **PColour**(k). Recall F_ℓ from (1.2).

Lemma 4.3. *Assume the hypotheses of Lemma 4.1, and let ϵ_1 and ϵ be arbitrarily small positive constants. Define $\tilde{\mathbf{y}}$ and $\tilde{z}$ to be the solution to* (4.17) *and* (4.18) *above with initial conditions $y_\ell(\epsilon_1) = q_\ell$ for each ℓ, and $z(\epsilon_1) = \hat{\epsilon}_1$. Define $\tilde{y}_{i,C} = \tilde{y}_{i,j}$ for all C with $|C| = j$. Given $\eta > 0$, define*

$$x_0 = \inf\{x > 0 \,:\, \tilde{\mathbf{y}}_{0,0} \leqslant \epsilon \text{ or } F_{\ell_2}(\tilde{\mathbf{y}}) \geqslant 1 - \epsilon \text{ or } \tilde{y}_\ell(x) = \eta_\ell/2 \text{ for some } \ell \in B^-\} \tag{4.23}$$

where ℓ_2 is as in (4.15). *If η is sufficiently small and $x_0^- < x_0$, then in* **PComplete***$(k, \mathscr{S})$ *with $|\mathscr{S}|$ bounded, a.a.s.*

$$Y_{i,C}(t) = n\tilde{y}_{i,C}(t/n) + o(n) \text{ uniformly for all } i, C, \text{and all } \epsilon_1 n \leqslant t < x_0^- n, \tag{4.24}$$

and

$$Z(t) = n\tilde{z}(t/n) + o(n) \text{ uniformly for all } \epsilon_1 n \leqslant t < x_0^- n. \tag{4.25}$$

Proof. We modify the argument in Section 6:[17], noting that Lemma 4.1 implies that each of the basic types is common, and hence permits us to argue about the behaviour of process **PComplete***$(k, \mathscr{S})$ after the preparatory phase in the same way that Section 6:[17] considers **PColour**(k) with the assumption that the basic type is common. Lemma 6.1:[17] and its proof are still valid, since it relies on a clutch containing just one vertex of basic type, in this case any of the expanded basic types. The discussion leading up to Theorem 6.2:[17] still holds in the new context except that now a number of different variables v refer to basic types. Thus, equations (6.12):[17] and (6.13):[17], which are still valid, determine $v_{i,C}$ only for the nonbasic types (i, C). Clearly we have from (6.7):[17] and the definition of the rules of choosing between basic types, $\sum_{|C|=j} \mu_{i,C} = p_{i,j}$ for every basic type (i, j), and thus

$$\mu_{i,C} = \frac{p_{i,|C|} Y_{i,C}}{\sum_{|C|=j} Y_{i,C}}. \tag{4.26}$$

We may observe that translating these equations and (4.13) to the v variables, in the

https://doi.org/10.1017/S0963548306007954 Published online by Cambridge University Press

symmetrical case of course (the possible asymmetry in the coordinates of $\mathbf{Y}$ will be dealt with shortly) gives (4.21).

The definition of the domain $\mathscr{D}_{\epsilon,\epsilon_0}$ is hereby altered (though, for convenience, we do not change its name) to incorporate the condition on the determinant of the matrix M for the new system of equations which arise for the variables $v_{i,C}$ by translating equations (4.26) to the v variables, together with (6.12):[17] and (6.13):[17]. We further constrain $\mathscr{D}_{\epsilon,\epsilon_0}$ for the current setting by adding the condition that (D1) and (D2), given by (4.3) and (4.14), must hold for $(x, \mathbf{Y}/n) \in \mathscr{D}_{\epsilon,\epsilon_0}$. The old condition concerning F is deleted, since it only entered the argument to ensure the Lipschitz conditions. By Lemma 4.2 and (4.16), $p_{i,j}$ is Lipschitz when viewed as a function of $\mathbf{Y}/n$, for $\mathbf{Y}/n \in \mathscr{D}_{\epsilon,\epsilon_0}$, and hence by (D1) and (4.26), so is $\mu_{i,C}$.

Hence the conclusion of Theorem 6.2:[17] also holds in the present setting, and the proof is a direct translation to the present setting. There is one additional twist that needs to be mentioned: the initial conditions for this application of [19, Theorem 6.1] are only known asymptotically since they result from the conclusion of Lemma 4.1. However, as with the observation in the proof of Corollary 6.4:[17], if the initial conditions are asymptotically equal then the solutions will be asymptotically equal. It follows that we may use the differential equations with initial conditions being exactly the constants q_ℓ and $\hat{\epsilon}_1$ referred to in Lemma 4.1.

We do not write the resulting differential equations explicitly, but move on to deduce that, in **PColour***(k), a.a.s. the approximation expressed in (4.24) and (4.25) holds at least until the solution first leaves the domain given by the intersection of $\mathscr{D}_{\epsilon,\epsilon_0}$ with the conditions (4.3) and (4.14). Again it is straightforward to verify that the solution of the differential equations is symmetrical since the initial conditions are, and so we write $y_{i,j}$ and $v_{i,j}$. The resulting symmetrical equations (corresponding to (6.16):[17]) are (4.17) and (4.18), where the supporting equation (4.21) follows from (4.26) and (4.13).

The only nontrivial alterations required to adapt the argument in Section 6:[17] to the present setting revolve around the relationship between the values of $v_{i,j}$ for (expanded) basic types given in (4.21), and the values of $v_{i,j}$ in Section 6:[17], where the only basic type is $(1,1)$. This relationship needs to be understood in order to establish the validity of the conclusion of Theorem 6.3:[17] (which allows some conditions on the domain to be ignored) in the present context.

In order to understand $\alpha(\mathbf{y})$ and later to simplify the above differential equations, we will use the following fact. With the system $\mathscr{E}_\ell$, let B_1 be the set of non-dangerous types with higher priority than the type $\ell = (i_0, j_0)$. Then the set of the expanded basic types B^- contains B_1 as a subset. The denominator in the following formula is bounded away from 0 for $\mathbf{y} \in \mathscr{D}_{\epsilon,\epsilon_0}$ by (4.14). Note also that, by its definition, $\theta_{i,j}^{(\ell)}$ is zero if $(i,j) < \ell$.

Proposition 4.4. *For any $\ell \in B^-$, define $\zeta^{(\ell)}$ by* (4.16):[17] *and* (4.15):[17] *in the context of system $\mathscr{E}_\ell$. Then*

$$\zeta^{(\ell)}(\mathbf{y}) = \sum_{(i,j)\in B^-} (d-i)\binom{k}{j}\theta_{i,j}^{(\ell)}(\mathbf{y})/(1 - F_{\ell_2}(\mathbf{y})).$$

https://doi.org/10.1017/S0963548306007954 Published online by Cambridge University Press

Proof. The expression on the right here is, using (4.17):[17] for the first step and (4.16):[17] for the second-last, equal to

$$\begin{aligned}
&\frac{(d-i_0)\binom{k}{j_0}\theta^{(\ell)}_{i_0,j_0}(\mathbf{y})+\sum_{(i,j)\in B_1}(d-i)\binom{k}{j}\theta^{(\ell)}_{i,j}(\mathbf{y})}{1-F_{\ell_2}(\mathbf{y})}\\
&\quad=\frac{(d-i_0)\big(1-\sum_{(i,j)\in B_1\cup B'}\binom{k}{j}\theta^{(\ell)}_{i,j}(\mathbf{y})\big)+\zeta^{(\ell)}(\mathbf{y})\sum_{(i,j)\in B_1}(d-i)\binom{k}{j}b_{i,j}(\mathbf{y})}{1-F_{\ell_2}(\mathbf{y})}\\
&\quad=\frac{(d-i_0)-\zeta^{(\ell)}(\mathbf{y})\big((d-i_0)\sum_{(i,j)\in B'}\binom{k}{j}b_{i,j}(\mathbf{y})-\sum_{(i,j)\in B_1}(d-i)\binom{k}{j}b_{i,j}(\mathbf{y})\big)}{1-F_{\ell_2}(\mathbf{y})}\\
&\quad=\frac{\zeta^{(\ell)}(\mathbf{y})a(\mathbf{y})-\zeta^{(\ell)}(\mathbf{y})\big((i-i_0)\sum_{(i,j)\in B_1}\binom{k}{j}b_{i,j}(\mathbf{y})+(d-i_0)\sum_{(i,j)\in B'}\binom{k}{j}b_{i,j}(\mathbf{y})\big)}{1-F_{\ell_2}(\mathbf{y})}\\
&\quad=\zeta^{(\ell)}(\mathbf{y})\text{ by (4.15):[17].}
\end{aligned}$$

□

Now we are ready to re-express $\alpha(\mathbf{y})$. First note that by (4.19)–(4.21),

$$\begin{aligned}
\alpha(\mathbf{y})&=\sum_{(i,j)\in B^-}(d-i)\binom{k}{j}\sigma(\mathbf{y})\hat{\theta}_{i,j}(\mathbf{y})+\alpha(\mathbf{y})\sum_{(i,j)\in B'}(d-i)\binom{k}{j}b_{i,j}(\mathbf{y})\\
&=\sum_{(i,j)\in B^-}(d-i)\binom{k}{j}\sigma(\mathbf{y})\hat{\theta}_{i,j}(\mathbf{y})+\alpha(\mathbf{y})F_{\ell_2}(\mathbf{y}).
\end{aligned}$$

This gives, using Proposition 4.4 for the last step,

$$\begin{aligned}
\alpha(\mathbf{y})&=\sum_{(i,j)\in B^-}(d-i)\binom{k}{j}\sigma(\mathbf{y})\hat{\theta}_{i,j}(\mathbf{y})/\big(1-F_{\ell_2}(\mathbf{y})\big)\\
&=\sum_{(i,j)\in B^-}(d-i)\binom{k}{j}\sigma(\mathbf{y})\sum_{\ell\in B^-}\hat{\lambda}(\ell)\theta^{(m(\ell))}_{i,j}(\mathbf{y})/\big(1-F_{\ell_2}(\mathbf{y})\big)\\
&=\sigma(\mathbf{y})\sum_{\ell\in B^-}\hat{\lambda}(\ell)\sum_{(i,j)\in B^-}(d-i)\binom{k}{j}\theta^{(m(\ell))}_{i,j}(\mathbf{y})/\big(1-F_{\ell_2}(\mathbf{y})\big)\\
&=\sigma(\mathbf{y})\sum_{\ell\in B^-}\hat{\lambda}(\ell)\zeta^{(m(\ell))}(\mathbf{y}).
\end{aligned}\tag{4.27}$$

In view of this, equations (4.20)–(4.22) give a unique solution for v, showing that $|\det M|$ is also nonzero in the current setting, and as in Section 6:[17], this implies that it is bounded away from 0 in a neighbourhood of the symmetrical solution. Thus the conclusion of Lemma 6.3:[17] also holds in the present setting, provided only that we alter the definition of x_0 to (4.23). (The condition $s(\mathbf{y})>\epsilon$ and the first condition in (4.3) are both superseded by a stronger lower bound on $y_{0,0}$.) Recalling the first observations in this proof, this actually provides us with the analogue of Corollary 6.4:[17], thereby completing the proof of Lemma 4.3. □

For manipulating the differential equations (4.17), we can clearly assume that the values of $\mathbf{y}$ under consideration lie inside $\mathscr{D}_{\epsilon,\epsilon_0}$ as (re)defined in the proof of Lemma 4.3. Hence, the equations derived there can be applied. By (4.21) and (4.27), (4.17) becomes for $\ell\in B^-$

$$y'_\ell=\sigma(\mathbf{y})\sum_{\ell'\in B^-}\hat{\lambda}(\ell')\Big(\zeta^{(m(\ell'))}(\mathbf{y})b_\ell(\mathbf{y})-\theta^{(m(\ell'))}_\ell(\mathbf{y})\Big)\tag{4.28}$$

https://doi.org/10.1017/S0963548306007954 Published online by Cambridge University Press

and of course for $\ell \notin B^-$, $y'_\ell = 0$. By (4.18), the definition of $\hat{\theta}_{i,j}$ in (4.22), Lemma 4.1:[17] and (4.16):[17], it then follows (with $\hat{\lambda}(\ell) = 0$ if $\ell \notin B^-$) that

$$y'_\ell = \sigma(\mathbf{y}) b_\ell(\mathbf{y}) \sum_{m(\ell') \geqslant \ell} \hat{\lambda}(\ell') \zeta^{(m(\ell'))}(\mathbf{y}) + \sigma(\mathbf{y}) \frac{F_\ell(\mathbf{y}) - 1}{c(\ell)} \sum_{m(\ell') = \ell} \hat{\lambda}(\ell') \zeta^{(m(\ell'))}(\mathbf{y}), \tag{4.29}$$

and so by (4.27) and (4.18)

$$\frac{\mathrm{d}y_\ell}{\mathrm{d}z} = b_\ell(\mathbf{y}) \sum_{m(\ell') \geqslant \ell} \lambda^*(\ell') + \frac{F_\ell(\mathbf{y}) - 1}{c(\ell)} \sum_{m(\ell') = \ell} \lambda^*(\ell'), \tag{4.30}$$

where

$$\lambda^*(\ell') = \frac{\hat{\lambda}(\ell') \zeta^{(m(\ell'))}(\mathbf{y})}{\sum_{\ell' \in B^-} \hat{\lambda}(\ell') \zeta^{(m(\ell'))}(\mathbf{y})}.$$

Note that $\zeta^{(\ell)}$ is bounded above and below by positive constants for ℓ satisfying the conditions in (4.7), by its formula (4.9):[17]. Hence,

$$\zeta^{(m(\ell'))}(\mathbf{y}) \text{ is bounded above and below by positive constants.} \tag{4.31}$$

So the λ^* are all nonnegative and their sum is 1.

For application of Lemma 4.3, and also to examine the behaviour of the solutions, it is important to know that the lower bound on $\tilde{\mathbf{y}}$ in the definition (4.23) of x_0 is never achieved before reaching x_0.

Lemma 4.5. *Define the functions $\tilde{\mathbf{y}}$ and $\tilde{z}$ of x as in Lemma 4.3. Then for η sufficiently small, $\tilde{y}_\ell(x) \geqslant \eta_\ell$ for all $\ell \in B^-$, $\epsilon_1 \leqslant x \leqslant x_0$.*

Proof. Assume the contrary. For η sufficiently small, the strict versions of the required inequalities are satisfied at $x = \epsilon_1$, so let x_1 be the infimum of those x for which one of them fails. Then $\tilde{y}_\ell(x_2) < \eta_\ell$ for some ℓ and some x_2 arbitrarily close to x_1, and we may assume $\epsilon_1 < x_1 < x_0$. By the Mean Value Theorem, the derivative of $\tilde{y}_\ell$ is negative at some x_3 arbitrarily close to x_1 for which $\tilde{y}_\ell(x_3) < \eta_\ell$.

Note that the derivative α of z is always positive, by (4.27) and (4.31) and the positivity of σ and the fact that the $\hat{\lambda}$ are nonnegative and sum to 1. Thus $\frac{\mathrm{d}y_\ell}{\mathrm{d}z}$ is negative at $z_3 = z(x_3)$. (For simplicity, let $\mathbf{y}$ denote $\tilde{\mathbf{y}}$ as a function of z.)

However, we may rewrite (4.30) using (1.2) as

$$\frac{\mathrm{d}y_\ell}{\mathrm{d}z} = b_\ell(\mathbf{y}) \sum_{m(\ell') > \ell} \lambda^*(\ell') + \frac{F_{\ell-1}(\mathbf{y}) - 1}{c(\ell)} \sum_{m(\ell') = \ell} \lambda^*(\ell'),$$

where the two summations are nonnegative. We will show that the other factors are both nonnegative at z_3 when the cofactor summations are positive. It follows that the derivative is nonnegative at z_3, a contradiction which proves the lemma.

Consider the second term first. There are two cases. If $R_2 = \ell$ then either $F_{\ell-1}(\mathbf{y}) - 1 \geqslant 0$ as required, or ℓ must be the minimum type in B^-. However, we claim that its cofactor summation is then zero. For otherwise, the definition of λ^* (and λ) implies that $y_{\ell'} \leqslant 2\eta$ for all $\ell' > \ell$, yet we also have $y_\ell < 2\eta$ at z_2, contradicting (for sufficiently small η) the

https://doi.org/10.1017/S0963548306007954 Published online by Cambridge University Press

fact that $y_{0,0} \geqslant \epsilon$, which follows from the definition of x_0. In the second case, $R_2 < \ell$. Then the coefficient of $F_{\ell-1}(\mathbf{y}) - 1$ is just $\lambda^*(\ell)$, a multiple of $\hat{\lambda}(\ell)$. This is 0 by definition of λ since $y_\ell(z_3) < \eta_\ell$.

We now turn to b_ℓ, as defined just after (4.19):[17], and recall that ℓ denotes (i, j). There are two cases. The first is that $i \geqslant 1$. Then one of the types $(i-1, j)$ and $(i-1, j-1)$ is a member of B^- (If one of these is not, its corresponding variable y is treated as 0 by definition.) So by the choice of x_1, $y_{i-1,j} + y_{i-1,j-1} \geqslant \eta_{\ell-1} = d^3\eta_\ell \geqslant d^3 y_{i,j}$ at x_1. For x_3 sufficiently close to x_1, continuity therefore ensures that $b_\ell > 0$ at z_3. The second case for b_ℓ is that $i = 0$, so $\ell = (0, 0)$. But the definition of x_0 ensures that $\tilde{\mathbf{y}}_{0,0} \geqslant \epsilon$, so this is not a relevant case for η sufficiently small. □

Before proceeding, we transform the problem to one which has a slightly more convenient form. With $\ell = \texttt{prio}(i, j)$, define $c(i, j) = c(\ell)$,

$$w_{i,j}(z) = w_\ell = c(\ell) y_\ell(z)$$

and

$$b_\ell(\mathbf{w}) = \frac{(d-i+1)j}{(d-2z)k}\left(\frac{w_{i-1,j}}{c(i-1,j)} + \frac{w_{i-1,j-1}}{c(i-1,j-1)}\right) - \frac{d-i}{d-2z}\left(\frac{w_{i,j}}{c(i,j)}\right) \tag{4.32}$$

so that the value of b_ℓ at a given point in $\mathbf{w}$-space is actually $b_\ell(\mathbf{y}(\mathbf{w}))$ in terms of the previous function b_ℓ. (This abuse of notation should cause no problem because the two functions can easily be distinguished by the argument: $\mathbf{w}$ or $\mathbf{y}$.) The (invertible) linear transformation $w_\ell(z) = c(\ell) y_\ell(z)$ also results in the differential equations for $\mathbf{w}$-space corresponding to (4.30). Since the notation $\hat{\lambda}(\ell)$ hides the fact that this depends on $\mathbf{y}$, we need new functions

$$\tilde{\lambda}(\ell) = \max\{0, \min\{w_\ell/(\eta_\ell c(\ell)) - 1, 1\}\}\left(1 - \sum_{\ell' > \ell, \ell' \in B^-} \tilde{\lambda}(\ell')\right), \tag{4.33}$$

$$\tilde{\lambda}^*(\ell) = \frac{\tilde{\lambda}(\ell)\zeta^{(m(\ell))}(\mathbf{y}(\mathbf{w}))}{\sum_{\ell \in B^-} \tilde{\lambda}(\ell)\zeta^{(m(\ell))}(\mathbf{y}(\mathbf{w}))}, \tag{4.34}$$

which compute the same things as before, but as functions of $\mathbf{w}$. The differential equation system (4.30) thus becomes

$$\frac{\mathrm{d}w_\ell}{\mathrm{d}z} = c(\ell) b_\ell(\mathbf{w}) \sum_{m(\ell') \geqslant \ell} \tilde{\lambda}^*(\ell') + \big(F_\ell(\mathbf{w}) - 1\big) \sum_{m(\ell') = \ell} \tilde{\lambda}^*(\ell') \tag{4.35}$$

where, as in (1.2) (*cf.* (3.2)),

$$F_\ell(\mathbf{w}) = \sum_{\ell' > \ell} c(\ell') b_{\ell'}(\mathbf{w})$$

is the relevant symmetric version of $\overline{F}_\ell$. This derivative is a convex combination of functions (renaming ℓ' as ℓ_1):

$$\frac{\mathrm{d}w_\ell}{\mathrm{d}z} = \sum_{\ell_1 \in B^-} \tilde{\lambda}^*(\ell_1) f_\ell^{\ell_1} \tag{4.36}$$

https://doi.org/10.1017/S0963548306007954 Published online by Cambridge University Press

where, since $c(\ell)b_\ell(\mathbf{y}) + F_\ell(\mathbf{y}) = F_{\ell-1}(\mathbf{y})$,

$$f_\ell^{\ell_1}(\mathbf{w}) = \begin{cases} c(\ell)b_\ell(\mathbf{w}) & \text{if } \ell < m(\ell_1), \\ 0 & \text{if } \ell > m(\ell_1), \\ F_{\ell_1-1}(\mathbf{w}) - 1 & \text{if } \ell = m(\ell_1). \end{cases} \tag{4.37}$$

We refer to the solution of these differential equations as $\mathbf{w}^{\langle\eta\rangle}$.

Our next task is to show that $\mathbf{w}^{\langle\eta\rangle}$ converges to some limiting function (in the appropriate domain) as $\eta \to 0$. The linear transformation $w_\ell(z) = c(\ell)y_\ell(z)$ results in differential equations for $\mathbf{w}$-space corresponding to the equations (1.4) for $\mathbf{y}$ which arise from the original prioritized process. These are

$$\frac{\mathrm{d}w_\ell}{\mathrm{d}z} = f_\ell^{\ell_0}(\mathbf{w}), \tag{4.38}$$

where (4.2) defines ℓ_0 as a function $\ell_0(\mathbf{w})$ (*cf.* ℓ_1 occurring in (4.36)), and (recalling from Section 3)

$$R_1 = R_1(\mathbf{w}) = \max\{\ell : w_\ell > 0\}, \quad R_2 = R_2(\mathbf{w}) = \min\{\ell : F_{\ell'}(\mathbf{w}) < 1 \text{ for all } \ell' \geqslant \ell\}. \tag{4.39}$$

Define $\mathbf{w}^{(0)}$ to be the solution of (4.38) obtained in the proof of Theorem 3.3 with initial conditions defined at $\hat{\epsilon}_1 = z(\epsilon_1)$ corresponding to the initial conditions for $\tilde{y}$ in Lemma 4.3 relating to the end of the prephase: $w_\ell(\hat{\epsilon}_1) = c(\ell)q_\ell$ for each ℓ. Also define z for $x > \epsilon_1$ as the solution of (4.18) with the initial condition $z(\epsilon_1) = \hat{\epsilon}_1$.

We may take the point of view that there are different systems each playing a partial role: (4.37) gives the derivative in each individual system. This observation is unfortunately not easily used because the L_1-norm of a vector is not differentiable when one of the coordinates is 0. Points where this occurs play a special role in the proof of the following.

Note that x_0 is a function of η, so we now denote it by $x_0^{(\eta)}$.

Lemma 4.6. *Define $z_0^{(\eta)}$ to be $z(x_0^{(\eta)})$ Then, as $\eta \to 0$,*

$$\sup_{\hat{\epsilon}_1 \leqslant z \leqslant z_0^{(\eta)}} |\mathbf{w}^{\langle\eta\rangle}(z) - \mathbf{w}^{(0)}(z)| \to 0.$$

Proof. We use precisely the full nodes of $\mathbf{w}^{(\delta)}$ in an iterative scheme as in the proof of Theorem 3.3, and show that the difference between $\mathbf{w}^{(\delta)}$ and $\mathbf{w}^{\langle\eta\rangle}$ is a function of δ and η which tends to zero as both δ and η go to zero in a certain manner. In fact we specify now that $\eta = O(\delta^2)$. Before proceeding, note that the δ-property, as in the proof of Theorem 3.3, is still valid, in that both $\mathbf{w}^{(\delta)}$ and its derivative are Lipschitz. We extend the notion of δ-property also to include the Lipschitz property of $\mathbf{w}^{\langle\eta\rangle}$, which holds because it has a bounded derivative, though its derivative does not have a Lipschitz property independent of η.

Assume that z_1 is a full node of $\mathbf{w}^{(\delta)}$ and that for some ξ (which, if necessary, could be specified to be $O(\delta)$ in the induction)

$$\sum_{\ell \in B^-} J_\ell(z_1) \leqslant \xi \tag{4.40}$$

https://doi.org/10.1017/S0963548306007954 Published online by Cambridge University Press

where

$$J_\ell(z) = |w_\ell^{(\delta)}(z) - w_\ell^{\langle\eta\rangle}(z)|.$$

Let z_2 be the next full node of $\mathbf{w}^{(\delta)}$ after z_1. We are interested in the increase in the error of approximation of $\mathbf{w}^{(\delta)}$ by $\mathbf{w}^{\langle\eta\rangle}$ in the interval $I = [z_1, z_2]$, that is, in

$$J = \sum_{\ell \in B^-} J_\ell(z_2) - J_\ell(z_1). \tag{4.41}$$

In this proof, it is the behaviour of $\mathbf{w}^{\langle\eta\rangle}$ on I which determines the nature of the inductive step. For simplicity, let ℓ_0 denote $\ell_0(\mathbf{w}^{(\delta)}(z_1))$. The analogous function we require for $\mathbf{w}^{\langle\eta\rangle}$ is

$$\hat{\ell}(z) = \max\left(\hat{R}_1(\mathbf{w}^{\langle\eta\rangle}(z)), R_2(\mathbf{w}^{\langle\eta\rangle}(z))\right)$$

where $\hat{R}_1(\mathbf{w}^{\langle\eta\rangle})$ is the greatest ℓ such that $w_\ell^{\langle\eta\rangle} \geqslant 2c(\ell)\eta_\ell$, and R_2 as before takes the least value ℓ such that $F_{\ell'} < 1$ for all $\ell' \geqslant \ell$.

First we observe that $m(\ell') \geqslant \hat{\ell}(z)$ for all ℓ' with $\tilde{\lambda}^*(\ell') > 0$. Thus, if $\ell < \hat{\ell}(z)$ for all $z \in I$, then it follows from (4.35) that on the interval I,

$$\frac{\mathrm{d}w_\ell^{\langle\eta\rangle}}{\mathrm{d}z} = c(\ell) b_\ell(\mathbf{w}^{\langle\eta\rangle}).$$

If, in addition, $\ell < \ell_0$, then by the δ-property, the Lipschitz property of b_ℓ, and (4.40), this is equal to $c(\ell)b_\ell(\mathbf{w}^{(\delta)}) + O(\delta + \xi) = \frac{\mathrm{d}w_\ell^{(\delta)}}{\mathrm{d}z} + O(\delta + \xi)$. It then follows, by integrating the two derivatives on the interval I, that

$$J_\ell(z_2) - J_\ell(z_1) = O(\delta\xi + \delta^2) \qquad (\text{if } \ell < \min\{\hat{\ell}(z), \ell_0\} \text{ for all } z \in I). \tag{4.42}$$

We next proceed to consider the terms in (4.41) with $\ell \geqslant \min\{\ell_0, \min_{z\in I} \hat{\ell}(z)\}$. First note that, from (4.35), for any $\ell_1 \in B^-$,

$$\begin{aligned}\sum_{\ell \geqslant \ell_1} \frac{\mathrm{d}w_\ell^{\langle\eta\rangle}}{\mathrm{d}z} &= \sum_{\ell' : m(\ell') \geqslant \ell_1} \tilde{\lambda}^*(\ell') \left(\left(F_{m(\ell')}(\mathbf{w}^{\langle\eta\rangle}) - 1\right) + \sum_{\ell_1 \leqslant \ell \leqslant m(\ell')} c(\ell) b_\ell(\mathbf{w}^{\langle\eta\rangle}) \right) \\ &= \left(F_{\ell_1 - 1}(\mathbf{w}^{\langle\eta\rangle}) - 1\right) \sum_{\ell' : m(\ell') \geqslant \ell_1} \tilde{\lambda}^*(\ell'). \end{aligned} \tag{4.43}$$

Case 1: $\hat{\ell}(z) \geqslant \ell_0$ for all $z \in I$. Then $\tilde{\lambda}^*(\ell) = 0$ for $m(\ell) < \ell_0$ and so using (4.4)

$$\sum_{m(\ell) \geqslant \ell_0} \tilde{\lambda}^*(\ell) = 1. \tag{4.44}$$

By (4.42) we only need to consider terms with $\ell \geqslant \ell_0$ in (4.41). By the δ-property and the approximation (3.8), we have

$$F_{\ell_0 - 1}\left(\mathbf{w}^{\langle\eta\rangle}(z)\right) - 1 = F_{\ell_0 - 1}\left(\mathbf{w}^{(\delta)}(z_1)\right) - 1 + O(\xi + \delta)$$

for $z \in I$. Hence, from (4.43) and (4.44),

$$\begin{aligned}\sum_{\ell \geqslant \ell_0} w_\ell^{\langle\eta\rangle}(z_2) - w_\ell^{\langle\eta\rangle}(z_1) &= (z_2 - z_1)\left(F_{\ell_0 - 1}(\mathbf{w}^{(\delta)}(z_1)) - 1\right) + O(\delta\xi + \delta^2) \\ &= w_{\ell_0}^{(\delta)}(z_2) - w_{\ell_0}^{(\delta)}(z_1) + O(\delta\xi + \delta^2). \end{aligned} \tag{4.45}$$

https://doi.org/10.1017/S0963548306007954 Published online by Cambridge University Press

We may now apply the same observations to $\sum_{\ell>\ell_0} w_\ell^{\langle\eta\rangle}(z_2) - w_\ell^{\langle\eta\rangle}(z_1)$. Note that, by the definition of ℓ_0, $F_{\ell_0}(\mathbf{w}^{(\delta)}(z_1)) - 1 < 0$. Hence this summation is, if not negative, at most $O(\delta\xi + \delta^2)$. We may thus proceed first from (4.45) to

$$w_{\ell_0}^{(\delta)}(z_2) - w_{\ell_0}^{\langle\eta\rangle}(z_2) = w_{\ell_0}^{(\delta)}(z_1) - w_{\ell_0}^{\langle\eta\rangle}(z_1) + \sum_{\ell>\ell_0} w_\ell^{\langle\eta\rangle}(z_2) - w_\ell^{\langle\eta\rangle}(z_1) + O(\delta\xi + \delta^2)$$

and then to

$$|w_{\ell_0}^{(\delta)}(z_2) - w_{\ell_0}^{\langle\eta\rangle}(z_2)| \leqslant |w_{\ell_0}^{(\delta)}(z_1) - w_{\ell_0}^{\langle\eta\rangle}(z_1)| - \sum_{\ell>\ell_0} w_\ell^{\langle\eta\rangle}(z_2) - w_\ell^{\langle\eta\rangle}(z_1) + O(\delta\xi + \delta^2). \quad (4.46)$$

The definition of ℓ_0 implies that $w_\ell^{(\delta)} = 0$ for $\ell > \ell_0$. Hence, by the nonnegativity of the functions $\mathbf{w}^{(\delta)}$ (by definition) and $\mathbf{w}^{\langle\eta\rangle}$ (by definition of $z_0^{(\eta)}$ through x_0),

$$\begin{aligned}
&\sum_{\ell_0\leqslant\ell\in B^-} J_\ell(z_2) - J_\ell(z_1) \\
&\quad\leqslant |w_{\ell_0}^{(\delta)}(z_2) - w_{\ell_0}^{(\eta)}(z_2)| - |w_{\ell_0}^{(\delta)}(z_1) - w_{\ell_0}^{(\eta)}(z_1)| + \sum_{\ell>\ell_0} w_\ell^{\langle\eta\rangle}(z_2) - w_\ell^{\langle\eta\rangle}(z_1) \\
&\quad= O(\delta\xi + \delta^2)
\end{aligned}$$

by (4.46). So, by (4.42), $J = O(\delta\xi + \delta^2)$.

Case 2: $\hat{\ell}(z) < \ell_0$ for some $z \in I$. Let z_3 be a point in I which minimizes the value of $\hat{\ell}$ and write $\hat{\ell} = \hat{\ell}(z_3)$ for short. By definition,

$$w_\ell^{\langle\eta\rangle}(z_3) = O(\eta) \quad (4.47)$$

for all $\ell > \hat{\ell}$. Since $F_{\hat{\ell}}(\mathbf{w}^{\langle\eta\rangle}(z_3)) < 1$, it follows (by the δ-property) that $F_{\hat{\ell}}\big(\mathbf{w}^{\langle\eta\rangle}(z)\big) < 1 + O(\delta + \xi)$ for $z \in I$. By (4.47), on the same interval

$$F_{\ell-1}\big(\mathbf{w}^{\langle\eta\rangle}(z)\big) < 1 + O(\delta + \xi + \eta) \quad (4.48)$$

for all $\ell > \hat{\ell}$, including $\ell = \ell_0$. Hence, focusing on the fact from (4.47) that each such function $w_\ell^{\langle\eta\rangle}$ is small somewhere on I,

$$w_\ell^{\langle\eta\rangle}(z_2) = O\big(\delta(\delta + \xi + \eta)\big) = O(\delta^2 + \delta\xi),\ \ell_0 \leqslant \ell \in B^-. \quad (4.49)$$

Using this together with (4.43),

$$\begin{aligned}
\sum_{\ell\geqslant\ell_0} w_\ell^{\langle\eta\rangle}(z_1) &= \sum_{\ell\geqslant\ell_0} \big(w_\ell^{\langle\eta\rangle}(z_1) - w_\ell^{\langle\eta\rangle}(z_2)\big) + O(\delta^2 + \delta\xi) \\
&= -(z_2 - z_1)\big(F_{\ell_0-1}(\mathbf{w}^{\langle\eta\rangle}(z_1)) - 1\big) \sum_{m(\ell)\geqslant\ell_0} \tilde{\lambda}^*(\ell) + O(\delta^2 + \delta\xi) \\
&\leqslant -(z_2 - z_1)\big(F_{\ell_0-1}(\mathbf{w}^{\langle\eta\rangle}(z_1)) - 1\big) + O(\delta^2 + \delta\xi) \quad (4.50)
\end{aligned}$$

using the fact that the $\tilde{\lambda}^*$ are nonnegative and sum to 1, and also (4.48) for $\ell = \ell_0$. Hence using (4.40) and the δ-property, and then the derivative of $w_{\ell_0}^{(\delta)}$,

$$\begin{aligned}
\sum_{\ell\geqslant\ell_0} w_\ell^{\langle\eta\rangle}(z_1) &\leqslant -(z_2 - z_1)\big(F_{\ell_0-1}(\mathbf{w}^{(\delta)}(z_1)) - 1\big) + O(\delta^2 + \delta\xi) \\
&\leqslant w_{\ell_0}^{(\delta)}(z_1) - w_{\ell_0}^{(\delta)}(z_2) + O(\delta^2 + \delta\xi). \quad (4.51)
\end{aligned}$$

https://doi.org/10.1017/S0963548306007954 Published online by Cambridge University Press

Now note that

$$\begin{aligned}\sum_{\ell_0\leqslant\ell\in B^-} & J_\ell(z_2)-J_\ell(z_1)\\ &=\sum_{\ell_0\leqslant\ell\in B^-}\big(|w_\ell^{(\delta)}(z_2)-w_\ell^{\langle\eta\rangle}(z_2)|-|w_\ell^{(\delta)}(z_1)-w_\ell^{\langle\eta\rangle}(z_1)|\big)\\ &=w_{\ell_0}^{(\delta)}(z_2)-|w_{\ell_0}^{(\delta)}(z_1)-w_{\ell_0}^{\langle\eta\rangle}(z_1)|-\sum_{\ell_0<\ell\in B^-}w_\ell^{\langle\eta\rangle}(z_1)+O(\delta^2+\delta\xi)\end{aligned}$$

by (4.49) and the fact that $w_\ell^{(\delta)}(z_1)=0$ for $\ell>\ell_0$. By (4.51) and the nonnegativity of $\mathbf{w}^{(\delta)}$ and $\mathbf{w}^{\langle\eta\rangle}$ as noted above, we have

$$w_{\ell_0}^{\langle\eta\rangle}(z_1)\leqslant w_{\ell_0}^{(\delta)}(z_1)+O(\delta^2+\delta\xi),$$

and so the term $-|w_{\ell_0}^{(\delta)}(z_1)-w_{\ell_0}^{\langle\eta\rangle}(z_1)|$ is at most $-\big(w_{\ell_0}^{(\delta)}(z_1)-w_{\ell_0}^{\langle\eta\rangle}(z_1)\big)+O(\delta^2+\delta\xi)$. Hence, again using (4.49) and reversing the negative sign in the summation,

$$\sum_{\ell_0\leqslant\ell\in B^-}J_\ell(z_2)-J_\ell(z_1)\leqslant w_{\ell_0}^{(\delta)}(z_2)-w_{\ell_0}^{(\delta)}(z_1)+\sum_{\ell\geqslant\ell_0}\big(w_\ell^{\langle\eta\rangle}(z_1)-w_\ell^{\langle\eta\rangle}(z_2)\big)+O(\delta^2+\delta\xi).$$

Noting that $\tilde\lambda^*$ is continuous and so integrable, this gives

$$\begin{aligned}\sum_{\ell_0\leqslant\ell\in B^-} & J_\ell(z_2)-J_\ell(z_1)\\ &\leqslant(z_2-z_1)\big(F_{\ell_0-1}(\mathbf{w}^{(\delta)})-1\big)-\big(F_{\ell_0-1}(\mathbf{w}^{\langle\eta\rangle})-1\big)\int_{z_1}^{z_2}\sum_{m(\ell)\geqslant\ell_0}\tilde\lambda^*(\ell)\,\mathrm{d}z+O(\delta^2+\delta\xi)\\ &=\big(F_{\ell_0-1}(\mathbf{w}^{\langle\eta\rangle})-1\big)\int_{z_1}^{z_2}\sum_{m(\ell)<\ell_0}\tilde\lambda^*(\ell)\,\mathrm{d}z+O(\delta^2+\delta\xi).\end{aligned}\tag{4.52}$$

It only remains to consider terms in (4.41) with $\hat\ell\leqslant\ell<\ell_0$, for which it will be important to recall that $\hat\ell=\hat\ell(z_3)$. What is lost (in terms of approximation) from such ℓ will be compensated for by the negative value of (4.52) (apart from the error term). From (4.35), for $\ell\in B^-$ the derivative of $w_\ell^{\langle\eta\rangle}$ is $c(\ell)b_\ell(\mathbf{w}^{\langle\eta\rangle})-G(\ell)$, where (by (4.4))

$$G(\ell)=c(\ell)b_\ell(\mathbf{w}^{\langle\eta\rangle})\sum_{m(\ell')<\ell}\tilde\lambda^*(\ell')+\big(1-F_\ell(\mathbf{w}^{\langle\eta\rangle})\big)\sum_{m(\ell')=\ell}\tilde\lambda^*(\ell'),\tag{4.53}$$

whilst for $\ell<\ell_0$ the derivative of $w_\ell^{(\delta)}$ on the interval I is (by the δ-property)

$$c(\ell)b_\ell(\mathbf{w}^{(\delta)})=c(\ell)b_\ell(\mathbf{w}^{\langle\eta\rangle})+O(\delta+\xi).$$

Thus, for $\ell<\ell_0$,

$$|J_\ell(z_2)-J_\ell(z_1)|\leqslant\int_{z_1}^{z_2}|G(\ell)|\,\mathrm{d}z+O(\delta^2+\delta\xi).\tag{4.54}$$

We examine the sign of terms in the definition (4.53) of $G(\ell)$. The two summations are of course nonnegative. For those ℓ such that the first summation is nonzero, $\tilde\lambda^*(\ell')\neq0$ for some $\ell'<\ell$ and so the minimum in the definition (4.33) of $\tilde\lambda(\ell)$ must be less than 1. Thus $w_\ell^{\langle\eta\rangle}=O(\eta)$ and so, by the definition (4.32) of b_ℓ (and the nonnegativity of $\mathbf{w}^{\langle\eta\rangle}$), the first

https://doi.org/10.1017/S0963548306007954 Published online by Cambridge University Press

term of $G(\ell)$ is at least $-O(\eta)$. The second term is also either positive or $O(\delta)$, since for sufficiently small δ, using $\ell \geqslant \hat{\ell}$ and the δ-property,

$$F_\ell(\mathbf{w}^{\langle\eta\rangle}) - 1 \leqslant F_{\hat{\ell}}(\mathbf{w}^{\langle\eta\rangle}) - 1 + O(\delta) < O(\delta).$$

Thus $G(\ell) \geqslant -O(\delta+\eta) = -O(\delta)$ and consequently we may replace $|G(\ell)|$ in (4.54) by $G(\ell)$. Hence

$$\sum_{\hat{\ell}\leqslant\ell<\ell_0} |J_\ell(z_2) - J_\ell(z_1)| = \int_{z_1}^{z_2} \sum_{\hat{\ell}\leqslant\ell<\ell_0} G(\ell)\mathrm{d}z + O(\delta^2 + \delta\xi).$$

We find by the definition of F, and using the fact that $m(\ell') \geqslant R_2 \geqslant \hat{\ell}$ for all ℓ', that

$$\sum_{\hat{\ell}\leqslant\ell<\ell_0} G(\ell) = \big(1 - F_{\ell_0-1}(\mathbf{w}^{\langle\eta\rangle})\big) \sum_{l\,:\,m(\ell)<\ell_0} \tilde{\lambda}^*(\ell).$$

Thus, using (4.52),

$$\sum_{\hat{\ell}\leqslant\ell\leqslant\ell\in B^-} J_\ell(z_2) - J_\ell(z_1) \leqslant \sum_{\hat{\ell}\leqslant\ell<\ell_0} |J_\ell(z_2) - J_\ell(z_1)| + \sum_{\ell_0\leqslant\ell\in B^-} J_\ell(z_2) - J_\ell(z_1) = O(\delta^2 + \delta\xi).$$

We conclude that in Case 2, as in Case 1, $J = O(\delta^2 + \delta\xi)$. The usual inductive argument on errors now shows that $\mathbf{w}^{\langle\eta\rangle}$ and $\mathbf{w}^{(\delta)}$ differ by at most $O(\delta)$ on the interval in question. The proof of Theorem 3.3 shows that $\mathbf{w}^{(\delta)}$ approximates $\mathbf{w}^{(0)}$ to the same accuracy, which gives the lemma. □

We also modify the discussion in Section 10:[17] here, in order to complete the proof of Theorem 1.2.

Proof of Theorem 1.2. Assume that the hypotheses of the theorem hold. Then with **PColour***(k) specified as in Lemma 4.1, the conclusions of Lemma 4.3 hold for **PComplete***$(k, \mathscr{S})$. This is for all ϵ and $\eta > 0$. Taking η sufficiently small, Lemma 4.5 ensures that the definition (4.23) of x_0 is equivalent to

$$x_0 = \inf\{x > 0 \,:\, \tilde{\mathbf{y}}_{0,0} \leqslant \epsilon \text{ or } F_{\ell_2}(\tilde{\mathbf{y}}) \geqslant 1 - \epsilon\}. \tag{4.55}$$

As we saw, the equation (4.17) determining $\tilde{\mathbf{y}}$ is equivalent to (4.30) in $(z, \mathbf{y})$-space, which is then equivalent to the equation (4.36) for $\mathbf{w}^{\langle\eta\rangle}$ by a simple linear transformation. Lemma 4.6 shows that this $\mathbf{w}^{\langle\eta\rangle}$ is approximated arbitrarily closely by $\mathbf{w}^{(0)}$, by taking η sufficiently small, on the interval $[\hat{\epsilon}_1, z(x_0^{(\eta)})]$. The initial conditions of $\mathbf{w}^{(0)}$ are a function of $\hat{\epsilon}_1$ as determined by the analysis of the end of the preparatory phase, and as $\hat{\epsilon}_1 \to 0$ they must tend to

$$w_{0,0}(0) = d,\ w_{i,j}(0) = 0 \text{ otherwise}, \tag{4.56}$$

which correspond to (1.6). The successive approximation method of proof of Theorem 3.3 shows that, as with ordinary differential equations, a small perturbation of the initial conditions produces a small perturbation in the right-hand derivative solutions obtained. Hence, as $\hat{\epsilon}_1 \to 0$, the function $\mathbf{w}^{(0)}$ tends towards its value with initial conditions (4.56). Back in $(z, \mathbf{y})$-space, this implies the solutions of (4.30) converge to the function $\mathbf{y}$ defined

https://doi.org/10.1017/S0963548306007954 Published online by Cambridge University Press

in the statement of the present theorem. We now conclude in view of (4.55) that $z(x_0) \geqslant z_1$ for ϵ sufficiently small (in addition to the smallness already imposed on $\hat{\epsilon}_1$ and η). At this point we may finally fix such ϵ, and we conclude that for such ϵ, there are functions $\hat{\epsilon}_1(n)$ and $\eta(n)$ tending to 0 as $n \to \infty$, such that for the corresponding process **PComplete***$(k, \mathscr{S})$

$$Y_{i,C}(t) = n y_{i,|C|}(Z(t)/n) + o(n) \text{ uniformly for all } i, C, \text{and all } Z(t) \leqslant z_1 n. \tag{4.57}$$

Note that this is also uniform over all sets $\mathscr{S}$ of bounded size.

To use the argument in Section 8:[17], we note that the analogous equation holds back in $(x, \mathbf{y})$-space because of the equivalence of the differential equations, *i.e.*, (4.24) and (4.25) are now valid with the range of t altered to $0 \leqslant t \leqslant x(z_1)n$ and with $\tilde{y}_{i,C}(t/n)$ redefined as $y_{i,j}(z(t/n))$.

As noted in the proof of Lemma 4.3, we have that Lemma 6.1:[17] is valid for the expanded basic type, as well as the analogue of Theorem 6.2:[17]. Hence the whole argument of Section 8:[17] is valid in the present context. Lemmas 9.1:[17] and 9.2:[17] are proved without relying on the priorities other than the fact that dangerous types have higher priorities than others, so these are also valid in the present context. The argument proving Theorem 10.1:[17] also applies without change (making the appropriate translation: the analogue of (6.22):[17] follows from (4.57)). We conclude that **Short***(g, k) a.a.s. succeeds when applied to a random $G \in \mathscr{G}_{n,d}$, where **Short***(g, k) is defined by modifying **Colour**(k) so that, when a vertex of expanded basic priority is called for, its type is chosen according to the probability $\hat{\phi}_{i,j}(t)$ just as in the definition of process **PColour***(k). The theorem follows. □

5. Numerical and rigorous results for small $d \geqslant 5$

In this section, we finish the proof of Theorem 1.1 for $5 \leqslant d \leqslant 10$. We will use the following priority list for $d > 4$, defined inductively:

$$\texttt{prio}(0,0) = 1,\ \texttt{prio}(d-2, k-2) = 0,$$

$$\texttt{prio}(i,j) = \begin{cases} \texttt{prio}(i+1,j)+1 & \text{if } j \neq k-2,\ j \leqslant i < j+d-k \\ & \text{or } j = k-2,\ j \leqslant i < d-3, \\ \texttt{prio}(j-1,j-1)+1 & \text{if } j \neq k-2,\ i = j+d-k \\ & \text{or } (i,j) = (d-3, k-2), \\ -1 & \text{otherwise.} \end{cases}$$

This list is generated by taking the approach that a vertex of type (i, j) should have higher priority than one of type $(i, j-1)$ because it has more different colours on neighbours, and higher priority than one of type $(i+1, j)$ because it has more neighbours yet to be coloured. If $i > j + d - k$, then a vertex v of type (i, j) has $k - j$ colours available for $d - i$ (which is less than $k - j$) uncoloured neighbours and thus we can properly colour the vertex v no matter how its uncoloured neighbours are coloured. So this type is given the lowest priority, which is -1. The type $(d-2, k-2)$ is given priority 0 since vertices of this type do not increase the value of the left-hand side of (1.7) as $y_{d-2,k-2}$ increases

https://doi.org/10.1017/S0963548306007954 Published online by Cambridge University Press

and this tactic will help the stability of the differential equations, regarding phases, when computing numerically.

We 'solved' the system (1.4) as called for in Theorem 1.2 numerically, using a Runge–Kutta method of order 2. This was without using careful error bounds but just apparent good convergence of the numerical solution as the stepsize was made smaller. This gave, to a high degree of confidence, the upper bounds in Theorem 1.1 for d up to 10, as well as the upper bound 6 for $d = 11$, 12 and 13, and the upper bound 7 for $d = 14$. These last four are all just one above the Molloy–Reed lower bound. The rest of this section describes an algorithm used to obtain the rigorous bounds reported in Theorem 1.1.

It is desirable to obtain analytic results for all d, as for $d = 4$ in Section 11:[17]. The end of phase condition, when the basic variable changes, can be detected analytically and this provides the starting point for the next phase.

Unfortunately, however, for large d (even $d = 5$) the analytic solutions become messy from the fact that the initial conditions for later phases are rational functions of exponentials. This is especially daunting when the number of phases to be considered becomes large. (For example, for $d = 10$ it appears that more than 20 phases are necessary.)

Theorem 1.2 refers to the solution of (1.4), which, as noted early in Section 4:[17], is equivalent to (4.20):[17], or, as noted in Section 4:[17], we may use (4.22):[17]. So for $d > 4$ we solve the system of differential equations (4.22):[17] numerically. These equations are:

$$\frac{\mathrm{d}u_{i,j}}{\mathrm{d}v} = \begin{cases} b_{i,j}(\mathbf{u}) + 2u_{i,j} & \text{if } (i,j) \notin B, \\ 0 & \text{if } (i,j) \in B', \\ b_{i,j}(\mathbf{u}) + 2u_{i,j} + \frac{-k+\sum_{(i',j')\in B'}(d-i')\binom{k}{j'}b_{i',j'}(\mathbf{u})}{k(d-i)\binom{k}{j}} & \text{if } (i,j) = (i_0,j_0), \end{cases} \tag{5.1}$$

where

$$b_{i,j}(\mathbf{u}) = (d-i+1)j(u_{i-1,j} + u_{i-1,j-1})/k - (d-i)u_{i,j}.$$

We need to show why this is equivalent to solving (1.4), given (1.5). For $\mathbf{u}$-space, we can also define the basic type $\ell_0(\mathbf{u})$ by taking the maximum of $R_1(\mathbf{u})$ and $R_2(\mathbf{u})$ where the definition of R_i is similar to (4.1) upon substitution of the variable $\mathbf{y}$ by $\mathbf{u}$. Then in (4.22):[17], the active system depends on which type ℓ_0 is basic. Also note that $u_\ell > 0$ if and only if $y_\ell > 0$, and $b_\ell(\mathbf{u}) = b_\ell(\mathbf{y})$ implies $R_2(\mathbf{u}) = R_2(\mathbf{y})$. This implies that $\ell_0(\mathbf{u}) = \ell_0(\mathbf{y})$ and thus the system active at a point in $\mathbf{u}$-space corresponds to the system active at the corresponding point in $\mathbf{y}$-space, and a phase change in $\mathbf{u}$-space corresponds to a phase change in $\mathbf{y}$-space. In each system, the right derivatives of variables are given by the right-hand side of (5.1) and changes of variable for right derivatives work in the same way as for ordinary differential equations. Thus (1.4) and (5.1) are equivalent.

To obtain rigorous results from the numerical solutions, we need to bound the errors. In order to do more work with integers (rather than floating point calculations), we use

$$\beta_{i,j} = (d-i+1)j(u_{i-1,j} + u_{i-1,j-1}) - k(d-i)u_{i,j}$$

and

$$\frac{\mathrm{d}u_{i,j}}{\mathrm{d}v} = f_{i,j}(\mathbf{u}), \tag{5.2}$$

https://doi.org/10.1017/S0963548306007954 Published online by Cambridge University Press

where

$$f_{i,j}(\mathbf{u}) = \begin{cases} (\beta_{i,j} + 2u_{i,j})/k & \text{if } (i,j) \notin B, \\ 0 & \text{if } (i,j) \in B', \\ \frac{(d-i)\binom{k}{j}(\beta_{i,j}+2ku_{i,j})-k+\sum_{(i',j')\in B'}(d-i')\binom{k}{j'}\beta_{i',j'}}{k(d-i)\binom{k}{j}} & \text{if } (i,j) = (i_0,j_0), \end{cases}$$

and write

$$\mathbf{f}(\mathbf{u}) = (f_{i,j}(\mathbf{u})).$$

Here the coefficients of the $u_{i,j}$ in the case $(i,j) = (i_0,j_0)$ are integers which can be precomputed.

In a C program, we store upper and lower bounds $\mathbf{u}^+$ and $\mathbf{u}^-$ on the variables, beginning with the initial conditions at $v = 0$, $u_{0,0}(0) = 1/d$ and $u_{i,j}(0) = 0$ otherwise, and repeatedly step through increasing values of v, updating the bounds for each new value of v, based on upper and lower bounds $\mathbf{f}^+$ and $\mathbf{f}^-$ on the right-hand derivative of $\mathbf{u}$ in the interval between the old and new values of v. In each floating point calculation, the bounds are widened to allow for the maximum possible error due to floating point arithmetic. The basic method of calculation used is as follows. Note that for these calculations, what we call phase changes, *i.e.*, places where a different system is active, do not need to be found exactly: just closely enough that the required upper and lower bounds remain valid until the terminating condition is reached.

At the start of any interval $[v, v+h]$, we first compute rough upper and lower bounds on all variables in this interval. We can assume $j \leqslant k-2$ since we are considering non-dangerous types. We also have $j \geqslant 1$ unless the type is $(0,0)$. Hence, for all variables u_ℓ under consideration, we have $c(\ell) \geqslant k$ (the worst case being type $(d-1,1)$). We have the bound $u_\ell \leqslant 1/c(\ell)$ by definition of s just before (4.14):[17], and hence

$$u_\ell \leqslant 1/k. \tag{5.3}$$

Now, suppose that $\frac{\mathrm{d}y_{i,j}}{\mathrm{d}s} > \binom{k}{j}^{-1}$ at some point inside the domain where the approximations in the proof of Theorem 1.2 in Section 4:[17] are valid. Then in the process, the rate of change of $Y_{i,C}$ with respect to S is a.a.s. greater than $\binom{k}{j}^{-1}$ for some period of time, for each of the $\binom{k}{j}$ variables $Y_{i,C}$. But each edge added decreases S by 2 and decreases $\sum_{|C|=j} Y_{i,C}$ by at most 2. So this is impossible. Hence $\frac{\mathrm{d}y_{i,j}}{\mathrm{d}s} \leqslant \binom{k}{j}^{-1}$, and a similar argument shows that $\frac{\mathrm{d}y_{i,j}}{\mathrm{d}s} \geqslant -\frac{1}{2}\binom{k}{j}^{-1}$, the extra factor 2 coming from the fact that the vertex whose points are all matched in one step does not create a new vertex of relevant type (i,j). We now have (noting $s = \frac{\mathrm{d}z}{\mathrm{d}v}$ and $\frac{\mathrm{d}s}{\mathrm{d}z} = -2$)

$$\begin{aligned} \frac{\mathrm{d}u_{i,j}}{\mathrm{d}v} &= \frac{\mathrm{d}u_{i,j}}{\mathrm{d}s}\frac{\mathrm{d}s}{\mathrm{d}z}\frac{\mathrm{d}z}{\mathrm{d}v} \\ &= \left(\frac{1}{s}\frac{\mathrm{d}y_{i,j}}{\mathrm{d}s} - \frac{y_{i,j}}{s^2}\right)(-2s) \\ &= -2\left(\frac{\mathrm{d}y_{i,j}}{\mathrm{d}s} - u_{i,j}\right). \end{aligned}$$

https://doi.org/10.1017/S0963548306007954 Published online by Cambridge University Press

Finally, note that $\binom{k}{j} \geqslant k$ for $j \geqslant 1$ (since we only consider $j < k$), and for $j = 0$ we only consider type $(0,0)$ for which $\frac{\mathrm{d}y_{0,0}}{\mathrm{d}s}$ is never negative (as this variable can only decrease as x increases, *i.e.*, s decreases), we have the lower bound on $\frac{\mathrm{d}u}{\mathrm{d}v}$ of -2 and the upper bound $3/k$ (for convenience, we take this as 1). Thus $u_\ell^+(v) + h$ is an upper bound on the value of u_ℓ on the interval, and $u_\ell^-(v) - 2h$ is a lower bound. Moreover, since these bounds are valid up to the point, if any, that the solution leaves the domain, it is easy to verify in practice that the solution must remain in the domain inside the interval, by checking that no points satisfying these bounds on the interval are at the boundary of the domain.

Then we compute a range of ℓ which contains, for certain, the basic type defined by (3.3) throughout the entire interval $[v, v+h]$, by using the rough bounds on $\mathbf{u}$ for the whole interval. This can be done as follows. Let

$$\ell_0^- = \max\{R_1^-, R_2^-\}$$

where

$$R_1^- = \max\{\ell : u_\ell^- > 0\}, \qquad R_2^- = \max\{\ell : F_{\ell-1}(\mathbf{u}^-) \geqslant 1\}. \tag{5.4}$$

Similarly, set

$$\ell_0^+ = \max\{R_1^+, R_2^+\}$$

where

$$R_1^+ = \min\{\ell : u_{\ell+1}^+ = 0\}, \qquad R_2^+ = \min\{\ell : F_{\ell'}(\mathbf{u}^+) < 1 \ \forall \ell' \geqslant \ell\}. \tag{5.5}$$

Then it follows that

$$\ell_0(w) = \mathtt{prio}(i_0, j_0) \in [\ell_0^-, \ell_0^+] \quad \forall w \in [v, v+h].$$

Having the basic type ℓ_0 bounded roughly by this, and the rough bounds $\mathbf{u}^+$ and $\mathbf{u}^-$ on $\mathbf{u}$ for the interval $[v, v+h]$, we can refine the bounds for the right derivatives $\mathbf{f}$ given by the formula (5.2) for the whole interval $[v, v+h]$. For a given possible active system, let $\mathbf{f}^+(\mathbf{u}(w))$ denote the maximum possible right derivative in the interval $[v, v+h]$ thus obtained, and similarly $\mathbf{f}^-(\mathbf{u}(w))$ for the lower bound. Then, again using Lemma 3.1, we may use the Euler formula,

$$\mathbf{u}^\pm(v+h) = \mathbf{u}^\pm(v) + h\mathbf{f}^\pm(\mathbf{u}^\pm(\kappa)),$$

where $\kappa \in (v, v+h)$, and update the upper bound $\mathbf{u}^+$ for the interval $[v, v+h]$ to be the maximum of $\mathbf{u}^+(v)$ and $\mathbf{u}^+(v+h)$, maximized over all possible active systems (and similarly for the lower bound). This maximum value will be an upper bound on $\mathbf{u}$ for the whole interval since $\mathbf{u}^+(v+h)$ is an upper bound for the whole interval if $\mathbf{f}^+ > 0$, and $\mathbf{u}^+(v)$ otherwise. The symmetric statements hold for lower bounds.

Note that for all variables $\ell < \ell_0^-$, each system (5.2) which is possibly active on the interval $[v, v+h]$ specifies the derivative of u_ℓ to be β_ℓ/k, and thus u_ℓ is differentiable on this interval. Furthermore, β_ℓ ($\ell = \mathtt{prio}(i,j)$) is differentiable provided $u_{i-1,j}$, $u_{i-1,j-1}$ and $u_{i,j}$ are differentiable. Hence, if we define

$$L = \{\ell = \mathtt{prio}(i,j) \mid \ell < \ell_0^- \text{ and } \{\mathtt{prio}(i-1,j-1), \mathtt{prio}(i-1,j)\} \cap [\ell_0^-, \ell_0^+] = \emptyset\},$$

https://doi.org/10.1017/S0963548306007954 Published online by Cambridge University Press

Table 1.

d,k	Largest priority	Last basic type	v_0	Density
5,4	$\texttt{prio}(3,3)=6$	(1,1)	1.533897	0.999958
6,4	$\texttt{prio}(3,3)=9$	(3,2)	1.841634	0.999928
8,5	$\texttt{prio}(4,4)=16$	[(4,2),(3,2)]	1.955822	0.999997
9,5	$\texttt{prio}(4,4)=20$	(6,3)	2.189075	0.999999
10,6	$\texttt{prio}(5,5)=25$	(7,3)	2.128767	0.999992

then for $\ell \in L$, u_ℓ has a second derivative on $(v, v+h)$. So we may use Taylor's theorem for $\ell \in L$:

$$u_\ell(v+h) = u_\ell(v) + hu'_\ell(v) + \frac{1}{2}h^2 u''_\ell(\kappa) \tag{5.6}$$

where $\kappa \in (v, v+h)$. Based on the bounds on u'_ℓ, we compute the bounds on their second derivatives and hence update the bounds for $\mathbf{u}$ using (5.6).

Note that any system (i,j) with $\texttt{prio}(i,j) \in [\ell_0^-, \ell_0^+]$ might be active. In this case, we update all variables $\mathbf{u}$ for $\texttt{prio}(i,j) \leqslant \ell_0^+$ by letting $\mathbf{u}^+$ be the maximal upper bound over all active systems, and $\mathbf{u}^-$ be the minimal lower bound over all active systems. Based on the updated bounds on $\mathbf{u}$ for the interval $[v, v+h]$, we can update the range of the basic type $[\ell_0^-, \ell_0^+]$ for the whole interval $[v, v+h]$.

The program iterates the whole process to narrow down the interval $[\ell_0^-, \ell_0^+]$ until it cannot be improved further. The resulting improved bounds on all variables are used.

Also in the program, if $\ell_0^- < \ell_0^+$, then we halve the length of step $h := h/2$ in the iteration for narrowing down the range for the basic type. This halving is repeated until either $\ell_0^- = \ell_0^+$ or some predetermined fraction of the original h is reached.

On each iteration, the program tests the condition

$$\sum_{i,j}(d-i)(d-i-1)\binom{k}{j}u^+_{i,j} < 1, \tag{5.7}$$

which implies condition (ii) of Theorem 1.2 since $\mathbf{u} = \mathbf{y}/s$ and $\sum_{i,j}(d-i)\binom{k}{j}u_{i,j} = 1$. The main inequality in Theorem 1.2(i) is also checked at every step. The lower bound on $y_{0,0}$ does not need to be checked at every step since this variable has a positive derivative unless it becomes 'basic'. Therefore, to verify that this condition holds, it suffices to check that the range of possible basic variables never contains the type $(0,0)$.

The first point at which both this and (5.7) are seen to hold is denoted by v_0. If the condition in Theorem 1.2(i) is violated before (5.7) holds, no result is obtained for this value of d and k, and we set $v_0 = \infty$.

The program gives the rigorous results shown in Table 1. The largest priority of all the types is shown first. The next column shows the basic type or range of basic types $[\ell_0^-, \ell_0^+]$ that was obtained at the point v_0. We define the *density* to be the left-hand side of (5.7), and the density column shows an upper bound on the density at the point v_0. These results, together with the check (which is easily made) that type $(0,0)$ was not contained in the basic range at any point in time, complete the proof of Theorem 1.1.

https://doi.org/10.1017/S0963548306007954 Published online by Cambridge University Press

References

[1] Achlioptas, D. and Moore, C. (2004) The chromatic number of random regular graphs. In *Proc. RANDOM'04*, pp. 219–228.

[2] Achlioptas, D. and Naor, A. (2005) The two possible values of the chromatic number of a random graph. *Ann. of Math.* **162** 1335–1351.

[3] Billingsley, P. (1986) *Probability and Measure*, 2nd edn, Wiley, New York.

[4] Bollobás, B. (1985) *Random Graphs*, Academic Press, New York.

[5] Borodin, O. and Kostochka, A. (1977) On an upper bound of a graph's chromatic number depending on the graph's degree and density. *J. Combin. Theory Ser. B* **23** 247–250.

[6] Bressan, A. (1988) Unique solutions for a class of discontinuous differential equations. *Proc. Amer. Math. Soc.* **104** 772–778.

[7] Catlin, P. (1978) A bound on the chromatic number of a graph. *Discrete Math.* **22** 81–83.

[8] Cooper, C., Frieze, A., Reed, B. and Riordan, O. (2002) Random regular graphs of non-constant degree: Independence and chromatic number. *Combin. Probab. Comput.* **11** 323–342.

[9] Díaz, J., Serna, M. and Wormald, N. C. Computation of the bisection width for random *d*-regular graphs. *Theoret. Comput. Sci.*, to appear.

[10] Díaz, J., Grammatikopoulos, G., Kaporis, A. C., Kirousis, L. M., Perez, X. and Sotiropoulos, D. G. (2005) 5-regular multi-graphs are 3-colorable with positive probability independent of their size. To appear in *Proc. 13th Annual European Symposium on Algorithms.*

[11] Frieze, A. M. and Łuczak, T. (1992) On the independence and chromatic numbers of random regular graphs. *J. Combin. Theory Ser. B* **54** 123–132.

[12] Kemkes, G. and Wormald, N. (2005) On the chromatic number of a random 5-regular graph. Department of C&O research report series CORR 2005-22, University of Waterloo.

[13] Kostochka, A. V. (1978) Degree, girth and chromatic number. In *Combinatorics* (Proc. Fifth Hungarian Colloq., Keszthely, 1976), Vol. II, *Colloq. Math. Soc. Janos Bolyai*, Vol. 18, North-Holland, Amsterdam/New York. pp. 679–696,

[14] Krivelevich, M., Sudakov, B., Vu, V. H. and Wormald, N. C. (2001) Random regular graphs of high degree. *Random Struct. Alg.* **18** 346–363.

[15] Lawrence, J. (1978) Covering the vertex set of a graph with subgraphs of smaller degree. *Discrete Math.* **21** 61–68.

[16] Molloy, M. (1992) Mathematics thesis, University of Waterloo.

[17] Shi, L. and Wormald, N. (2007) Colouring random 4-regular graphs. *Combin. Probab. Comput.* **16** 309–344.

[18] Wormald, N. C. (1995) Differential equations for random processes and random graphs. *Ann. Appl. Probab.* **5** 1217–1235.

[19] Wormald, N. C. (1999) The differential equation method for random graph processes and greedy algorithms. In *Lectures on Approximation and Randomized Algorithms* (M. Karoński and H. Prömel, eds), PWN, Warsaw, pp. 73–155.

[20] Wormald, N. C. (1999) Models of random regular graphs. In *Surveys in Combinatorics, 1999* (J. D. Lamb and D. A. Preece, eds), pp. 239–298.

[21] Wormald, N. C. (2003) Analysis of greedy algorithms on graphs with bounded degrees. *Discrete Math.* **273** 235–260.

https://doi.org/10.1017/S0963548306007954 Published online by Cambridge University Press