The freezing threshold for k-colourings of a random graph.*

Michael Molloy†

April 28, 2016

Abstract

We determine the exact value of the freezing threshold, r_k, for k-colourings of a random graph when $k \geq 14$. We prove that for random graphs with density above r_k, almost every colouring is such that a linear number of vertices are frozen, meaning that their colours cannot be changed by a sequence of alterations whereby we change the colours of $o(n)$ vertices at a time, always obtaining another proper colouring. When the density is below r_k, then almost every colouring is such that every vertex can be changed by a sequence of alterations where we change $O(\log n)$ vertices at a time.

Frozen vertices are a key part of the clustering phenomena discovered using methods from statistical physics. The value of the freezing threshold was previously determined by the non-rigorous cavity method.

1 Introduction

Over the past decade, some groundbreaking hypotheses arising from statistical physics have driven much of the progress on random constraint satisfaction problems (CSP's). In particular, the *1-Step Replica Symmetry Breaking* hypothesis (1RSB) (see eg. [51]) says that, at a certain constraint density, called the *clustering threshold*, w.h.p.[1] the solution space shatters into an exponential number of *clusters of solutions*, where each cluster is well-connected and any two clusters are well-separated. Furthermore, at a higher density, called the *freezing threshold*, there are a linear number of *frozen variables* in *almost every* cluster; i.e. variables that are fixed throughout the cluster.

These hypotheses have had an enormous impact on study of random CSP's in the math and computer science communities. An understanding of these hypotheses has led to substantial new results, eg [21, 50, 14, 67, 37, 61, 22, 24, 34, 1, 46, 36]. Furthermore, much work has gone towards rigorously proving aspects of these hypotheses, eg [2, 3, 61, 6, 39, 29, 72, 25, 26, 28, 23, 12, 30, 27]. The main contribution of this paper is of the latter type.

In this paper, we rigorously prove a major hypothesis concerning frozen variables for k-COL; i.e. k-colourability of the Erdős-Rényi random graph $G_{n,M}$. This is one of the two most widely studied random CSP's, the other being k-SAT. We prove, for k sufficiently large, that frozen variables (i.e. vertices) do, indeed, arise. Furthermore, we prove the *exact* location of the freezing threshold; this had previously been estimated non-rigorously using the cavity method. We also determine the number of frozen variables, up to a $o(n)$ term.

Our main tool is the *planted model* which Achlioptas and Coja-Oghlan[21] proved could be used to analyze certain random CSP's (see also [61]), and Babst etal.[11] refined for k-colouring. Our approach should apply to determine the freezing threshold of any random CSP for which we can use the planted model. We chose to begin with k-COL, because it is the most well-studied such CSP. Subsequently to this work, with Restrepo we were able to apply the same approach to NAE-SAT and hypergraph 2-colourability[59].

*A shorter version of this paper appeared in the proceedings of STOC 2012.

†Dept of Computer Science, University of Toronto, molloy@cs.toronto.edu. Research supported by an NSERC Discovery Grant.

[1]We say that a property holds *with high probability (w.h.p.)* if it holds with probability tending to one as the number of variables tends to infinity.

To prove our theorem, we strip the random graph down to what we call a *Kempe core*, and prove that almost all of the vertices in the Kempe core are frozen. This is very much like the approach [3] took for random k-SAT, and that paper was an important inspiration for this one.

It has long been observed that most random CSP's appear to be very difficult to solve for a wide range of constraint densities. This was first observed for k-SAT in [20, 53]. For many CSP's, there is an "algorithmic barrier" substantially lower than the density at which they are w.h.p. unsatisfiable. For example: Random instances of k-SAT are known to pass from being w.h.p. satisfiable to w.h.p. unsatisfiable at constraint density $2^k \ln 2 + O(k)$[9, 26, 30], but no algorithm has been proven to w.h.p. find a satisfying solution for problems of density higher than $O(\frac{2^k \ln k}{k})$[21]. The random graph $G_{n,M}$ is known to pass from being w.h.p. k-colourable to w.h.p. not k-colourable at edge-density $k \ln k + O_k(1)$[8, 28, 23], but no algorithm has been proven to w.h.p. find a k-colouring of a random graph with edge-density higher than $\frac{1}{2} k \ln k(1 + o_k(1))$[5, 38]. These barriers are asymptotically (in k) equal to the hypothesized location of the clustering threshold, and this was given rigorous grounding in [21]. To be clear: this is not a barrier for all efficient algorithms; eg. Survey Propogation appears to succeed past this point (see below). But it appears to be a barrier for algorithms that are simple enough for rigorous analysis using current techniques, specifically simple greedy algorithms.

In [73, 71, 47] it is argued that these algorithmic difficulties are not brought on by the clustering threshold, but rather by the freezing threshold. In other words, the clusters do not pose significant difficulties until they have frozen variables. In fact, Achlioptas and Moore[7] prove that a simple greedy algorithm finds 3-colourings of random graphs above the hypothesized clustering threshold but below the hypothesized freezing threshold. While the clustering threshold is hypothesized to be strictly less than the freezing threshold, the gap tends to zero as k grows. In particular, the freezing threshold is also asymptotic to the observed algorithmic barrier.

For small values of k, the algorithm Survey Propogation[14] seems empirically to be able to find solutions for random k-SAT and k-colourability when the density is very close to the satisfiability threshold. It appears that this algorithm can only find solutions which have no frozen variables, despite the fact that the density is far above the freezing threshold and so almost all solutions have frozen variables. It appears that the second freezing threshold, above which *every* solution has frozen variables, is a second algorithmic barrier above which no known algorithm can find a solution, even empirically. Recent estimates[13] place this barrier substantially higher than the clustering/freezing thresholds, but still much lower than the satisfiability threshold for large k.

We close this section by noting that the *cavity method* has been used to predict many thresholds and other important results concerning random CSP's, including satisfiability thresholds (see eg. [51] for many examples). The quest to "rigourize" applications of the cavity method has been one of the most important trends in the study of random structures over the past decade. Very roughly speaking, the cavity method focuses on analyzing the distance-d neighbourhood of a randomly selected vertex for arbitrarily large, but constant, d, making use of the fact that this neighbourhood is w.h.p. a tree. It then hypothesizes the manner in which the remainder of the graph should affect the analysis; this is typically the point which is very difficult to do rigorously as it concerns the long-range dependencies between vertices in the graph. In this paper, we effectively show that as far as freezing is concerned, *the effect of the long-range dependencies is negligible*; the freezing threshold is exactly what the tree-analysis predicts. It is hoped that our techniques will lead to other results along this line.

2 Clusters and Frozen Variables

We study $G_{n,M}$, the random graph with n vertices and M edges, where each such graph is equally likely. We are interested in the range $M = rn$ where r is constant. This model was introduced by Erdos and Renyi in two seminal papers[31, 32]. In these papers, they posed several natural questions about random graphs. All but one have since been answered; the remaining question is: What is the chomatic number of $G_{n,M=rn}$ for $r > \frac{1}{2}$? It is widely believed that for each $k \geq 3$, there is a constant ϕ_k such that for $r < \phi_k$, $G_{n,M=rn}$ is w.h.p. k-colourable while for $r > \phi_k$, $G_{n,M=rn}$ is w.h.p. not k-colourable. The determination of ϕ_k is one of the most important open problems, and indeed the oldest open problem, in random graph theory. Thus far, we do not even know whether ϕ_k exists. Achlioptas and Friedgut[4] proved something close - the existence of a function $\phi_k(n)$. Achlioptas and Naor[8] proved that $\phi_k(n) = k \ln k + O(\ln k)$. Coja-Ohlan and Vilenchik[28] and Coja-Oghlan[23] improved this to $\phi_k(n) = (k - \frac{1}{2}) \ln k + O_k(1)$. Coja-Oghlan et al.[25] determine the k-colourability threshold for random regular graphs (when k is large); note that this threshold is not sharp in the Erdos-Renyi sense, as the edge-density parameter is integer-valued.

The 1-RSB hypothesis arose from statistical physics and provides a very strong picture of the solutions of random constraint satisfaction problems. It was first applied to k-COL by Mulet et.al.[63] (see eg. [66, 71] for further work). One of the central concepts is that when the density is above the *freezing threshold* almost all solutions will have *frozen variables*. In the setting of k-COL, these are vertices whose colours cannot be changed using a sequence of local alterations where we change the colours of a small number of variables at a time. Instead, to change the colour of a frozen vertex requires a global alteration where we change the colours of a linear proportion of the vertices at once. Formally:

Definition 2.1. *An ℓ-path of k-colourings of a graph G is a sequence $\sigma_0, \sigma_1, ..., \sigma_t$ of k-colourings of G, where for each $0 \leq i \leq t-1$, σ_i and σ_{i+1} differ on at most ℓ vertices. We say that two k-colourings σ, σ' are ℓ-connected if they can be joined by an ℓ-path $\sigma = \sigma_0, ..., \sigma_t = \sigma'$ for some $t \geq 0$.*

We emphasize that there is no restriction on the length of the path. So two ℓ-connected colourings might differ on arbitrarily many vertices, and we may require an arbitrarily long ℓ-path to join them.

Definition 2.2. *Given a k-colouring σ of a graph G, we say that a vertex v is* ℓ-frozen *with respect to σ if for every ℓ-path $\sigma = \sigma_0, \sigma_1, ..., \sigma_t$ of k-colourings of G, we have $\sigma_t(v) = \sigma(v)$.*

In other words, it is not possible to change the colour of v by changing at most ℓ vertices at a time. Usually, when we say that a vertex is frozen we mean that it is ℓ-frozen for some $\ell = \Theta(n)$.

We define

$$r_k^f = \min_{x>0} \frac{(k-1)x}{2(1-e^{-x})^{k-1}}. \tag{1}$$

For any $r > r_k^f$ we let $x_k(r)$ denote the largest positive solution to $r = \frac{(k-1)x}{2(1-e^{-x})^{k-1}}$.

Our main theorem is that, for k sufficiently large, r_k^f is the precise threshold for most colourings to have a linear number of ℓ-frozen vertices, where ℓ is linear in n:

Theorem 2.3. *For any $k \geq 14$, let σ be a uniformly random k-colouring of $G_{n,M=rn}$.*

(a) For any $r_k^f < r < (k-1)\ln(k-1)$ there exists positive constants $Q = Q(r,k)$ and $\alpha = \alpha(r,k)$ such that:

(i) w.h.p. there are $\frac{(k-1)x_k(r)}{2r} n + o(n)$ vertices that are αn-frozen with respect to σ.

(ii) w.h.p. there are $(1 - \frac{(k-1)x_k(r)}{2r})n + o(n)$ vertices that are not $Q \log n$-frozen with respect to σ.

(b) For any $r < r_k^f$, there exists positive constant $Q = Q(r,k)$ such that w.h.p. no vertices are $Q \log n$-frozen with respect to σ.

Remark: When $k \geq 14$, we have $r_k^f < (k-1)\ln(k-1)$ so part (a) is meaningful. The upper bound $(k-1)\ln(k-1)$ on r comes from Theorem 4.3 below; it is possible that a more careful analysis of the arguments from [11, 8] could allow us to weaken that upper bound and so extend the result to smaller values of k. Analysis from statistical physics shows (non-rigorously) that the freezing theshold is equal to r_k^f for $k \geq 9$ and is less than r_k^f for $k \leq 8$. So we will not be able to extend the result below $k = 9$.

Hypothesized values for r_k^f are provided in [71, 66] for $3 \leq k \leq 10$, using the cavity method to determine an expression for r_k^f and using population dynamics to estimate the value of that expression[2]. They begin

[2] Those papers report the threshold in terms of the average degree, rather than edge-density and so their values are exactly twice ours. Also note that what we call the freezing threshold is called the rigidity threshold in [71].

by obtaining a formula for the freezing threshold on the "tree factor graph", which is hypothesized to be equal to the freezing threshold in $G_{n,p}$ so long as one is below the condensation threshold. For $3 \le k \le 8$ the freezing threshold appears to be greater than the condensation threshold, and so they adjust their formula accordingly. Their formula for the tree factor graph is different than our expression (1), but can be shown to be equivalent[3].

Asymptotically, we have:

$$r_k^f = \frac{1}{2}k(\ln k + \ln\ln k + 1 + o(1)). \tag{2}$$

This rigorously confirms the asymptotics obtained using the cavity method; see (44) of [71] and (78) of [66].

In fact, we prove something stronger than Theorem 2.3. In Section 5, we define a subset of the vertices which we call the *Kempe core.* r_k^f is the threshold for the appearance of a Kempe core. We will prove that w.h.p. all but $o(n)$ vertices of the Kempe core are frozen and at most $o(n)$ vertices outside of the Kempe core are frozen. Thus, given a uniform k-colouring σ of $G_{n,M=rn}$, we w.h.p. specify precisely which vertices are frozen with respect to σ up to an error of $o(n)$ vertices.

2.1 Frozen Clusters

The key feature of the 1-RSB hypothesis is that when the density exceeds the *clustering threshold*, almost all of the solutions can be partitioned into *clusters*. One can travel amongst the solutions in a cluster by making local changes where one changes a small number of variables at a time. But to travel to a solution outside the cluster, requires a global change.[4]

More specifically, in the context of k-COL: Let $\Omega_k(G)$ denote the set of k-colourings of a graph G. It is believed that at some density $r \approx \frac{1}{2}k\ln k$, i.e. roughly half the k-colourability threshold, w.h.p. $\Omega_k(G)$ can be partitioned into sets $S_1, ..., S_x$ such that one can move within S_i by changing the colours of only $o(n)$ vertices at a time, but to move from S_i to S_j requires changing a linear number of vertices. More formally:

Definition 2.4. *For a partition $\Omega_k(G) = S_1 \cup ... \cup S_x$, we call the parts (a, b)-clusters if*

- *(a) for all $i \neq j$, no pair $\sigma \in S_i, \sigma' \in S_j$ is a-connected; and*
- *(b) for all i, every pair $\sigma, \sigma' \in S_i$ is b-connected.*

Condition (a) says that the clusters are *well-separated.* Condition (b) says that the clusters are *well-connected.*

If $a = b+1$ then (a, b)-clusters exist trivially in every graph. Remarkably, it appears that in $G_{n,M=cn}$ we have (a, b)-clusters when a is much greater than b: $a = \Theta(n), b = o(n)$. The clustering hypothesis[63, 51, 71] is:

Hypothesis A: *For $r > r_k^c \approx \frac{1}{2}k\ln k$, there exists a constant $\alpha > 0$ and a function $\beta(n) = o(n)$ such that w.h.p. almost all of $\Omega_k(G_{n,M=rn})$ can be partitioned into an exponential (in n) number of $(\alpha n, \beta(n))$-clusters. Furthermore, each cluster contains an exponential number of colourings. This does not happen for $r < r_k^c$.*

We note that further details are also hypothesized; eg. the clusters change substantially after the *condensation threshold*[48]. See [51] for a good overview. A freezing hypothesis[66, 71] is:

Hypothesis B: *For $r > r_k^f \approx \frac{1}{2}k\ln k$: W.h.p. almost all*[5] *clusters S_i have a linear number of* frozen vertices *v, with the property that for all $\sigma, \sigma' \in S_i$ we have $\sigma(v) = \sigma'(v)$. This does not happen for $r < r_k^f$*

Note that if Hypothesis A holds then for every $\beta(n) < \ell \le \alpha n$, the frozen vertices in the cluster containing σ are exactly the vertices that are ℓ-frozen with respect to σ. In particular, every vertex that is αn-frozen according to Definition 2.2, is also frozen in the sense of Hypothesis B, assuming Hypothesis A. So if Hypothesis A holds then Theorem 2.3 implies Hypothesis B for $k \ge 14$.

[3]Our thanks to a referee of a preliminary version of this paper[56] for showing this.

[4]Between the clustering and freezing thresholds, the picture is more subtle than this. There may be a path of small local changes from one cluster to another, but such paths are very "thin", and a random walk would require exponential time to find one.

[5]Here, "almost all" means for all but a vanishing proportion of the clusters when they are weighted by their size.

3 Related work

A previous version of this paper appeared in the Proceedings of STOC 2012[56]. The theorems here are somewhat stronger, mainly because we are able to make use of the recent work concering the planted model in [11]. The $Q \log n$ terms in Theorem 2.3 were $o(n)$ terms in [56], and the phrase "no vertices" in Theorem 2.3(b) was "$o(n)$ vertices".

As discussed above, the solution space geometry for k-colourings of $G_{n,M}$ was first studied in [63]. The freezing threshold was studied in great depth in [71, 66, 73]. These studies were non-rigorous, but mathematically sophisticated. The results in this paper confirm predictions made in those papers, and does not contradict anything found there. [71] was the first paper to argue that freezing may be the cause of the algorithmic barrier.

Achlioptas and Ricci-Tersenghi[10] were the first to rigorously prove any form of freezing in a random CSP. They studied random k-SAT and showed that for $k \geq 8$, for a wide range of edge-densities below the satisfiability threshold and for *every* satisfying assignment σ, the vast majority of variables are 1-frozen w.r.t σ. Equivalently, such vertices are frozen in what they call 1-clusters, which are equivalent to $(2, 1)$-clusters of Definition 2.4. Such clusters are trivially connected, but they are not known to be $\Theta(n)$-separated and hence to satisfy Hypothesis A. However, it is plausible that they are in some sense close to being the clusters of Hypothesis A.

[3, 2, 61] prove the asymptotic value of the freezing threshold for various random CSP's[6]. For k-COL, [2] establishes that the threshold is $(\frac{1}{2} + o(1))k \ln k$, which agrees with (2).

As mentioned above, [10] studies 1-clusters for random k-SAT. [61] studies analogous clusters for other CSP's. Such clusters are connected, but are not know to be $\Theta(n)$-separated. [3, 2] also studies what they call *cluster-regions*, which are proven to be $\Theta(n)$-separated but are not shown to be well-connected. [2] proves that for $r < (\frac{1}{2} - \epsilon)k \ln k$ w.h.p. almost all k-colourings are in a single cluster region, while for $r > (\frac{1}{2} + \epsilon)k \ln k$ the solution space shatters into an exponential number of $\Theta(n)$-separated cluster-regions, each containing an exponential number of colourings. They, and also [61], prove analgous results for other CSP's. While these cluster-regions do not satisfy Hypothesis A, note that, intuitively, the well-connected property of clusters does not seem to be critical to the problems that they pose for algorithms. So these results help to explain why the asymptotic order of the algorithmic barrier is $\frac{1}{2}k \ln k$.

After the preliminary version of this work[56], the author and Restrepo extended the techniques to prove analogous results for a general class of CSP's including hypergraph 2-colouring and NAE-SAT[59].

The clusters of k-XOR-SAT are very well-understood[6, 39, 36]. We know the clustering threshold which is also the freezing threshold, and have a very good description of the clusters and the frozen variables. The picture is much simpler here; for example, the same variables are frozen in every cluster. The simple linear algebraic characterization of the solution space is very helpful.

4 The planted model

Definition 4.1. *The* uniform model $U_{n,M}$ *is a random pair* (G, σ) *where* G *is taken from the* $G_{n,M=rn}$ *model and* σ *is a uniformly random* k*-colouring of* G*.*

The biggest hurdle to theorems such as Theorem 2.3 used to be that there was no representation of the uniform model that lends itself to analysis. This hurdle, along with the corresponding hurdles for a few other random CSP's, was overcome by Achlioptas and Coja-Oghlan[2] who proved that, under certain conditions, one can work instead with the much simpler planted model. Those conditions were weakened substantially by Babst etal[11].

Definition 4.2. *The* planted model $P_{n,M}$ *is a random pair* (G, σ) *chosen as follows: Take a uniformly random partition* σ *of* $\{1, ..., n\}$ *into* k *parts* $A_1, ..., A_k$*. Then choose* M *random edges, uniformly and without replacement, from all edges whose endpoints are in two different parts.*

[6]In fact, they focus on what [2] calls *rigid* variables, but it is simple to extend their argument to frozen variables.

In other words, $P_{n,M}$ is chosen by first choosing a uniformly random k-colouring σ of the vertices $\{1,...,n\}$, and then choosing a graph that is uniform amongst all graphs with that vertex set and with M edges, for which σ is a k-colouring. This clearly has a different distribution than what one obtains by carrying those steps out in the other order; i.e. first choosing $G_{n,M}$ and then taking a uniformly random k-colouring of G. But remarkably, [2, 11] proves that one can transfer w.h.p. properties:

Theorem 4.3. *[11] For every $k \geq 3$ and every $r < (k-1)\ln(k-1)$:*
Let $\mathcal{E}$ be any property of pairs (G,σ) where σ is a k-colouring of G. If $P_{n,M=rn}$ w.h.p. has $\mathcal{E}$ then $U_{n,M=rn}$ w.h.p. has $\mathcal{E}$.

Remark: The first transfer result of this type appeared in [2]. That result had the additional requirement that (essentially[7]) for any function $f(n) = o(n)$ the probabilty that $\mathcal{E}$ does not hold it $P_{n,M=rn}$ must be at most $e^{-f(n)}$. A previous version of this paper relied on [2] and, as a result, the statement corresponding to Theorem 2.3 was somewhat weaker; Most notably, part (b) only said that $o(n)$ vertices are frozen.

It will be more convenient to work in the $G_{n,p}$ version of the planted model, which we define as follows:

Definition 4.4. *The* planted model $P_{n,p}$ *is a random pair (G,σ) chosen as follows: Take a uniformly random partition σ of $\{1,...,n\}$ into k parts $A_1,...,A_k$. Each pair of vertices in two different parts is joined with an edge with probability p, where the edge-choices are independent.*

The following standard lemma permits us to work in $P_{n,p}$ rather than $P_{n,M}$. We say that a property $\mathcal{E}$ is *convex* if: for every three graphs G_1, G_2, G_3 on the same partition σ with $G_1 \subseteq G_2 \subseteq G_3$, if $(G_1,\sigma),(G_3,\sigma)$ both have $\mathcal{E}$ then (G_2,σ) has $\mathcal{E}$.

Lemma 4.5. *Consider any convex property $\mathcal{E}$ of pairs (σ, G) where σ is a k-colouring of G, and any constant r. Setting $c = \frac{2k}{k-1}r$, we have:*
If $P_{n,p=c/n}$ w.h.p. has $\mathcal{E}$ then $P_{n,M=rn}$ w.h.p. has $\mathcal{E}$.

The (omitted) proof is almost identical to the proof of Theorem 2.2(b.ii) in [16], which implies the same statement for $G_{n,p}$ and $G_{n,M}$.

We define

$$c_k = \min_{y>0} \frac{ky}{(1-e^{-y})^{k-1}}.$$

For any $c > c_k$ we let $y_k(c)$ denote the largest solution to $c = \frac{ky}{(1-e^{-y})^{k-1}}$. Note that $c_k = \frac{2k}{k-1}r_k$. We define:

$$\lambda_k(c) = y_k(c)/c.$$

Definition 4.6. *We say that v is an ℓ*-frozen variable of (G,σ) *if v is ℓ-frozen with respect to σ.*

So, roughly speaking, our goal is to prove that c_k is the threshold for $P_{n,p=c/n}$ to have a linear number of αn-frozen variables.

5 Kempe cores

Given a k-colouring σ of a graph G, with colour classes $A_1,...,A_k$, a *Kempe chain* is a component of the subgraph induced by two colour classes. Suppose C is a non-empty Kempe chain on colour classes A_i, A_j. Then exchanging the colours i,j on the vertices of C will result in a new k-colouring of G. Note that a single vertex of colour i will constitute a Kempe chain if it has no neighbours of colour j, for some $j \neq i$. Kempe chains were introduced by Kempe[44] in his work on the Four Colour Problem.

It is clear that a vertex that is in a Kempe chain of size at most ℓ is not ℓ-frozen. This inspires us to remove all "small" Kempe chains from our graph, in order to look for frozen vertices. A bit of thought will make it clear that w.h.p. most vertices in Kempe chains of size at most ℓ in the remaining graph are not

[7] In fact, this only has to hold for a specific $f(n)$, but we have no knowledge about $f(n)$ other than that it is $o(n)$.

ℓ-frozen either. This follows from branching properties of the random graph: if C is a small Kempe chain in the remaining graph, w.h.p. the small Kempe chains that were removed from the original graph each have at most one edge to C. Furthermore none of those chains adjacent to C are adjacent to each other. Thus we can flip the vertices on some subset of those chains without them interfering with each other, thus enabling C to be flipped. This intuition inspires us to remove small Kempe chains iteratively.

Of course, we need to specify what we mean by "small". It turns out that typically[8] there will be no Kempe chains of size between $O(\log n)$ and $\Theta(n)$; i.e. every Kempe chain will either be small or giant. To be specific, we will take small to mean: of size at most $\log^2 n$. Thus, we apply the following procedure:

Kempe-Strip
Input: a graph G and a k-colouring $\sigma = A_1, ..., A_k$ of G.
While there are any Kempe chains of size at most $\log^2 n$
    Remove the vertices of one such Kempe chain from G.

Definition 5.1. *The (possibly empty)* Kempe-core *is what remains after running Kempe-Strip.*

Note that, as with most core stripping procedures, the output does not depend on the order in which we choose to remove Kempe chains. So the Kempe-core is well-defined.

Clearly no vertex in the Kempe-core can have its colour changed by changing the vertices of a small Kempe chain. It is much less clear that almost every vertex in the Kempe-core cannot have its colour changed by changing a small subset of vertices which involve *more than two* colours.

To gain some intuition as to why this may be the case, note first that (almost) every very small subgraph, i.e. of size $O(1)$, is a tree. It is a simple exercise (see the proof of Lemma 7.6) to note that if we can change the colours of a subtree to obtain another colouring, then that tree must contain a subtree which is a Kempe chain. Thus, (almost) any valid change of $O(1)$ vertices can be simulated by a sequence of Kempe chain switches. It follows that we cannot obtain another colouring by changing $O(1)$ vertices of the Kempe-core (unless they do not form a subtree, in which case they must form a flippable unicycle). Much of the work in this paper is to establish that the same is true for changes of up to αn vertices for small $\alpha > 0$.

Key idea. We are now ready to present one of key ideas behind this paper. One natural way to analyze the Kempe core involves determining the effect of removing each individual Kempe chain, and then tracking those effects throughout Kempe-Strip. This is daunting, in part because of the large number of possible Kempe chains. Instead, we recall that the planted model is the union of $\binom{k}{2}$ random bipartite graphs, one on each pair of colours. Note that the Kempe chains are precisely the components of these bipartite graphs. W.h.p. each bipartite graph will have at most one giant component, and the remaining components will all have size $O(\log n)$. So, at least intuitively, Kempe-Strip is equivalent to the following procedure: repeatedly remove all but the giant component from each of the bipartite graphs (see STRIP in section 8.1). That procedure is much more amenable to analysis.

The following lemma is one of the main steps in this paper, and is proven in Section 8. (See also Lemma 6.3(a)).

Lemma 5.2. *For $k \geq 3$:*

(a) If $c < c_k$ then w.h.p. the Kempe-core of $P_{n,p=c/n}$ is empty.

(b) If $c > c_k$ then w.h.p. the Kempe-core of $P_{n,p=c/n}$ has size $k\lambda_k(c) + o(n)$.

Lemmas 5.2, 4.5 and Theorem 4.3 immediately yield:

Corollary 5.3. *For $k \geq 14$:*

(a) If $r < r_k^f$ then w.h.p. the Kempe-core of $U_{n,M=rn}$ is empty.

[8] This will always be true unless the density of a particular subgraph is equal to the *giant component threshold*.

(b) If $r_k^f < r < (k-1)\ln(k-1)$ then w.h.p. the Kempe-core of $U_{n,M=rn}$ has size $\frac{(k-1)x_k(r)}{2r}n + o(n)$.

Proof Straightforward but tedious calculus and computation show: (i) at $k = 14$, $r_k^f < (k-1)\ln(k-1)$ and (ii) for $k \geq 14$, r_k^f grows more slowly than $(k-1)\ln(k-1)$ (we omit the details). Note that the property of having a Kempe-core of size $k\lambda_k(c)+o(n)$ is convex since the size of the Kempe-core is monotone increasing under the addition of edges; more formally, for any $h(n) = o(n)$ the property of having a Kempe-core of size in the range $k\lambda_k(c) \pm h(n)$ is convex. So Theorem 4.3 and Lemma 4.5 allow us to translate our results from $P_{n,p=c/n}$ to $U_{n,M=rn}$, with $r = \frac{k-1}{2k}c$. Note that $x(r) = y(c)$. So the corollary follows from Lemmas 5.2 and 5.3, and the fact that

$$k\lambda_k(c) = \frac{ky_k(c)}{c} = \frac{kx_k(r)}{2kr/(k-1)} = \frac{(k-1)x_k(r)}{2r}. \tag{3}$$

□

Having analyzed the Kempe-core, the next step is to show that it is, essentially, the set of frozen vertices. In Section 9, we show:

Lemma 5.4. *For $k \geq 3$, in $P_{n,p=c/n}$:*

(a) If $c > c_k$ then w.h.p. at most $o(n)$ vertices outside of the Kempe-core are $\log n$-frozen.

(b) If $c < c_k$ then there exists $Q = Q(c,k)$ such that: w.h.p. no vertex is $Q\log n$-frozen.

In Section 7 we show that almost all of the Kempe-core is frozen. To be precise, we define:

Definition 5.5. *A* unicycle *is a connected graph with exactly one cycle. A* flippable unicycle *in (G, σ) is an induced subgraph[9] $U \subseteq G$ such that (i) U is a unicycle and (ii) there is a proper colouring σ' of G where $V(U)$ is the set of vertices on which σ, σ' differ.*

Lemma 5.6. *W.h.p., the expected total size of all flippable unicycles in the Kempe-core of $P_{n,p=c/n}$ is $O(1)$.*

It follows that w.h.p. all flippable unicycles in the Kempe-core are very small, and so any vertex on a flippable unicycle is not frozen. However, we will show that all other vertices in the Kempe-core are frozen. The main step is to show that the Kempe core is *internally rigid:*

Lemma 5.7. *For $k \geq 3$, $c > c_k$, there exists constant $\alpha = \alpha(c,k) > 0$ such that w.h.p. the Kempe-core K of $P_{n,p=c/n}$ has the following property:*
If $v \in K$ is not in a flippable unicycle then any k-colouring of K which differs from σ on v must differ from σ on at least $2\alpha n$ vertices of K.

This internal rigidity is enough to imply:

Corollary 5.8. *For $k \geq 3$, $c > c_k$, there exists constant $\alpha = \alpha(c,k) > 0$ such that w.h.p.: All vertices of the Kempe-core K of $P_{n,p=c/n}$, other than those in flippable unicycles, are αn-frozen.*

Proof Let $\Theta \subseteq K$ be the Kempe-core vertices that do not lie on flippable cycles. So by Lemma 5.6 and Markov's Inequality, w.h.p. $|K \backslash \Theta| = o(n)$. Lemma 5.7 says that every $v \in \Theta$ has the property that any k-colouring of K which differs from σ on v must differ from σ on at least $2\alpha n$ vertices of K, and thus on at least $2\alpha n - o(n) > \alpha n$ vertices of Θ. Consider any sequence of k-colourings of K, $\sigma = \sigma_0, \sigma_1, ..., \sigma_t$ such that

(i) for all $v \in \Theta$ and $0 \leq i \leq t-1$, we have $\sigma_i(v) = \sigma(v)$.

(ii) for some $v \in \Theta$ we have $\sigma_t(v) \neq \sigma(v)$.

[9] U is an *induced* subgraph if U contains every edge in G that joins two vertices of U.

In other words, t is the first step where a member of Θ changes colour.

By (ii), σ_t must differ from σ on at least αn vertices of Θ. Thus by (i), σ_t must differ from σ_{t-1} on those same αn vertices. Therefore, $\sigma = \sigma_0, \sigma_1, ..., \sigma_t$ is not a αn-path. But if at least one vertex of Θ is not αn-frozen, then there must be such a αn-path; consider the vertex $v \in \Theta$ whose colour can be changed by the shortest possible αn-path. So all of the vertices of Θ must be αn-frozen. □

This yields our main theorem:

Proof of Theorem 2.3: As in the proof of Corollary 5.8, Theorem 4.3 and Lemma 4.5 allow us to translate our results from $P_{n,p=c/n}$ to $U_{n,M=rn}$, with $r = \frac{k-1}{2k}c$. The theorem then follows from Lemmas 5.2, 5.4 and 5.6, Corollary 5.8 and (3). □

5.1 Whitening and magic subgraphs

In the random graph setting of this paper, the Kempe-core is essentially equivalent to what arises from the *whitening* procedure described in [71]. The whitening procedure for the case $k = 3$ was created independently in [69] where what remains at the end was called the *magic subgraph*. The procedure can be described as follows:

Begin with a properly k-coloured graph. Transform the graph into a directed graph by replacing each edge with two directed edges, one in each direction. We then iteratively delete edges as follows:

Consider an edge $u \to v$. If there is at least one colour $a \in \{1, ..., k\}$ such that (i) a is not on u; and (ii) a does not appear on any vertex $w \neq v$ such that the edge $w \to u$ remains, then we delete the edge $u \to v$.

(Rather than deleting $u \to v$, [71] colours it white, and condition (ii) is modified to say ".... such that the edge $w \to u$ is not white.)

Note that $u \to v$ is *deleteable* iff after removing v from the graph, it is possible to change the colour of u without changing any other vertex.

When no remaining edges are deletable, the set of vertices that have an inneighbour of every colour other than their own is (nearly) equal to the Kempe core.

To see this, first note that in our random graph models almost all Kempe chains that are removed by Kempe-Strip will be trees (this is made more explicit in Observation 8.3 below and some of the analysis that follows in Section 8). Next note that if a Kempe-chain is a tree then both directions of all its edges will be removed by the whitening procedure, starting with the leaves, working inwards and then back out to the leaves again. After the edge pointing to a leaf is removed then so are all edges pointing from that leaf to vertices outside of the Kempe chain. The vertices in this Kempe-chain now have outdegree zero, and so will have no effect on future steps. So one might as well delete them. Following this reasoning, we see that every Kempe chain removed by Kempe-Strip will have outdegree zero in the magic subgraph (ignoring the neglibile effects of the few small chains that are not trees).

With a bit more thought, one can see that vertices of the Kempe-core will all have outdegree at least one after the whitening procedure halts. In Section 8, we will see that a non-empty Kempe-core induces a single connected component between each pair of colour classes; this component is not a tree, and so has a 2-core. The reader can easily verify that no directed edges in each 2-core will be deleted. Amongst the pendant trees (see Definition 7.9 below), all edges pointing towards the 2-core will eventually be deleted, and all edges pointing away away from the 2-core will remain. Since this is true of every pair of colour classes, each vertex has an inneighbour of every colour (other than its own) and so no more edges can be deleted.

The whitening procedure was studied in [71] and other papers in statistical physics and is a key ingredient to much of their analysis of k-colourings of random graphs. The magic subgraph was analyzed in [69] to provide a non-rigorous analysis of planted 3-colourings in a random graph. Both papers provided a (unproven) formula for the freezing threshold in the planted model[71]/threshold for a magic subgraph[69] which is different from, but equivalent to, (1).

6 Kempe-cores in the planted model

6.1 Properties of the Kempe-core

Let K be the Kempe-core of $P_{n,p=c/n}$ for some $c > c_k$, and for each $1 \leq a \leq k$, we let $K_a = K \cap A_a$ be the vertices of K with colour a. For each $a \neq b$, we let $K_{a,b}$ denote the bipartite subgraph of K induced by (K_a, K_b).

Lemma 6.1. *Consider any two connected bipartite graphs H, H', each with vertex set (K_a, K_b), and with $|E(H)| = |E(H')|$. Then $\mathbf{Pr}(K_{a,b} = H) = \mathbf{Pr}(K_{a,b} = H')$.*

Proof Consider any (G, σ) for which the procedure Kempe-Strip yields a Kempe core with $K_{a,b} = H$. Form G' by replacing the subgraph H in G with H'. Then applying Kempe-Strip to (G', σ) will yield a Kempe core with $K_{a,b} = H'$. Furthermore, G, G' arise with the same probability in $P_{n,p=c/n}$, since they have the same number of edges. This gives a bijection from graphs which yield $K_{a,b} = H$ and graphs which yield $K_{a,b} = H'$, where each pair of graphs occur with the same probability. This implies the lemma. □

Definition 6.2. *The* 2-core *of a graph is what remains after iteratively deleting any vertices of degree less than 2.*

Remark: It is easy to see that the order in which we delete vertices does not affect what remains at the end, so the 2-core is well-defined.

For any $c \geq c_k$ we let $y_k(c)$ denote the largest positive solution y to $c = \frac{ky}{(1-e^{-y})^{k-1}}$. Recall from Section 4 that c_k is defined as the minimum c such that $y_k(c)$ exists. We define:

$$
\begin{aligned}
\lambda_k(c) &= y_k(c)/c \\
\xi_k(c) &= \frac{y_k(c)(1 - e^{-y_k(c)}(1 + y_k(c)))}{c(1 - e^{-y_k(c)})} \\
\mu_k(c) &= \frac{y_k(c)e^{-y_k(c)}}{c(1 - e^{-y_k(c)})} \sum_{i \geq 2} \frac{y_k(c)^i}{(i-1)!} \\
\tau_k(c) &= \frac{y_k(c)e^{-y_k(c)}}{c(1 - e^{-y_k(c)})} \frac{y_k(c)^2}{2}
\end{aligned}
$$

We prove the next lemma in Section 8:

Lemma 6.3. *For any $c > c_k$ w.h.p. we have that for every a, b, the subgraph induced by $K_{a,b}$ is connected and:*

(a) $|K_a| = \lambda_k(c)n + o(n)$;

(b) the 2-core of $K_{a,b}$ has $\xi_k(c)n + o(n)$ vertices in K_a and $\xi_k(c)n + o(n)$ vertices in K_b;

(c) the 2-core of $K_{a,b}$ has $\mu_k(c)n + o(n)$ edges;

(d) the 2-core of $K_{a,b}$ has $\tau_k(c)n + o(n)$ degree 2 vertices in K_a and $\tau_k(c)n + o(n)$ degree 2 vertices in K_b.

The following bounds are crucial to our analysis. Essentially, they establish that certain branching parameters are subcritical; those parameters concern (a) the proportion of non-2-core vertices in each K_a, and (b) the degree two vertices in the 2-core of $K_{a,b}$:

Lemma 6.4. *For every $c > c_k$, there is $\zeta = \zeta(c) > 0$ such that:*

(a) $1 - \frac{\xi_k(c)}{\lambda_k(c)} < \frac{1}{k-1}(1 - \zeta)$;

(b) $\frac{2\tau_k(c)}{\mu_k(c)} < \frac{1}{k-1}(1 - \zeta)$.

Proof At $c = c_k$, $y = y_k(c)$ is the point that minimizes $h(y) = \frac{ky}{(1-e^{-y})^{k-1}}$. Setting $\frac{\partial}{\partial y}h(y) = 0$ yields:

$(1-e^{-y})^{k-1} = (k-1)ye^{-y}(1-e^{-y})^{k-2}$, which yields $e^y - 1 = (k-1)y$. Thus $\frac{e^y-1}{y} = k-1 > e-1$ and so $y > 1$. Thus $e^y > k-1$ and so $e^y - 1$ grows faster than $(k-1)y$ for $y \geq y_k(c)$. Since $y_k(c)$ increases with c, we have that for every $c > c_k$:

$$e^{y_k(c)} - 1 > (k-1)y_k(c).$$

Part (a):

$$\frac{\xi(c)}{\lambda_k(c)} = \frac{1 - e^{-y_k(c)}(1 + y_k(c))}{1 - e^{-y_k(c)}} = \frac{e^{y_k(c)} - 1 - y_k(c)}{e^{y_k(c)} - 1} > \frac{e^{y_k(c)} - 1 - \frac{1}{k-1}(e^{y_k(c)} - 1)}{e^{y_k(c)} - 1} = 1 - \frac{1}{k-1}.$$

This implies (a) since the LHS and RHS do not change with n.

Part (b) follows similarly from:

$$\frac{2\tau_k(c)}{\mu_k(c)} = \frac{y_k(c)^2}{\sum_{i\geq 2} \frac{y_k(c)^i}{(i-1)!}} = \frac{y_k(c)}{\sum_{i\geq 1} \frac{y_k(c)^i}{i!}} = \frac{y_k(c)}{e^{y_k(c)} - 1} < \frac{1}{k-1}.$$

□

7 The Kempe-core is mostly frozen

In this section, we prove Lemma 5.7. Recall that we are working in the $P_{n,p}$ model. So we have a uniformly random partition σ of the vertices into $A_1, ..., A_k$, and a graph G formed by selecting each of the potential edges between different parts with probability $p = c/n$. Our focus will be on the Kempe-core, K, of (G, σ).

Definition 7.1. *A* Δ-set *is the symmetric difference of* σ *and another* k*-colouring of the Kempe-core,* K*. Specifically, given such a colouring* σ'*, the set of vertices* $u \in K$ *with* $\sigma(u) \neq \sigma'(u)$ *is a* Δ*-set, which we sometimes denote by* $\sigma\Delta\sigma'$*.*

Note that "$v \in \sigma\Delta\sigma'$" means the same thing as "$\sigma(v) \neq \sigma'(v)$".

Recall that the subgraph induced by a set S of vertices is the subgraph consisting of S and all edges whose endpoints are both in S. With this notation (and after ruling out flippable trees - see Lemma 7.6), Lemma 5.7 is equivalent to:

Lemma 7.2. *W.h.p. every* Δ*-set which induces a connected subgraph with more than one cycle has size at least* $2\alpha n$

The remainder of this section is devoted to the proof of Lemma 7.2.

7.1 The structure of Δ-sets

To prove Lemma 7.2 we first study the structure of Δ-sets. When we say the 2-core of a Δ-set, we mean the 2-core of the subgraph of the Kempe-core induced by that set. Similarly, when we say a component of a Δ-set, we mean a component of the subgraph of the Kempe-core induced by that set.

Observation 7.3. *Every component of a* Δ*-set is a* Δ*-set.*

Proof Let C be a component of $\sigma\Delta\sigma'$ and consider changing the colour of every $v \in C$ from $\sigma(v)$ to $\sigma'(v)$. Since C is a component of $\sigma\Delta\sigma'$, every vertex u adjacent to C has $\sigma(u) = \sigma'(u)$. It follows that for every edge uv, either u, v have colours $\sigma(u), \sigma(v)$ or u, v have colours $\sigma'(u), \sigma'(v)$. Thus we have a valid colouring, and C is the set of vertices on which it differs with σ; i.e. C is a Δ-set. □

Recalling Lemma 6.3, we suppose we have a Kempe-core K satisfying:

Property 7.4. *Each $K_{a,b}$ is connected and has a non-empty 2-core, and that 2-core is not a cycle.*

We start with a key observation about Δ-sets:

Proposition 7.5. *Let u be any vertex in a Δ-set $\sigma\Delta\sigma'$. Then every neighbour of u in $K_{\sigma(u),\sigma'(u)}$ is also in $\sigma\Delta\sigma'$.*

Proof Every neighbour w of u in $K_{\sigma(u),\sigma'(u)}$ has $\sigma(w) = \sigma'(u)$. Since σ' is a proper colouring, we cannot have $\sigma'(w) = \sigma'(u)$. Therefore $\sigma'(w) \neq \sigma(w)$. $\square$

Lemma 7.6. *Every component of a Δ-set has a non-empty 2-core.*

Proof If a component has an empty 2-core, then it is a tree. Consider a tree-component of a Δ-set $\sigma\Delta\sigma'$. We direct the edges of that tree as follows: each u has an edge directed to every neighbour that it has in $K_{\sigma(u),\sigma'(u)}$; by Proposition 7.5, all such neighbours must be in the tree.

Note that an edge uv will be directed in both directions iff $\sigma(u) = \sigma'(v)$ and $\sigma(v) = \sigma'(u)$; contract all such edges. The contracted tree is a tree, and each edge is directed in exactly one direction. So there must be a node which has no edges directed out of it. Our contraction rule implies that there is a pair of colours a, b such that every vertex u contracted into that node has $(\sigma(u), \sigma'(u)) = (a, b)$ or (b, a). Furthermore, u has no neighbours in $K_{a,b}$ that were not contracted into the node, else this would have produced a directed edge out of the node. Therefore, the vertices contracted into that node are a component of $K_{a,b}$. But since they form a tree, this violates Property 7.4. $\square$

Note the following simple fact:

Proposition 7.7. *If a graph H is connected, then the 2-core of H is empty or connected.*

Proof Strip to the 2-core by repeatedly removing vertices of degree 1. The removal of a degree 1 vertex cannot disconnect a graph. $\square$

Definition 7.8. *We say that a Δ-set is* complex *if its 2-core does not have any components that are cycles. We say that a Δ-set is* cyclic *if its 2-core is a cycle.*

In other words, *a Δ-set is cyclic iff it is a flippable unicycle.* So to prove Lemma 7.2 we must show that w.h.p. the 2-core of every complex Δ-set is large.

We now turn our attention to the structure of the vertices outside the 2-core of a graph:

Definition 7.9. *Consider any graph H such that every component of H has a non-empty 2-core. The edges not in the 2-core of H form a forest. We call a tree of that forest a* pendant tree*. By Proposition 7.7, the 2-core of each component is connected. It follows that each pendant tree T contains exactly one vertex of the 2-core; that vertex is the* vertex of attachment *for T. We consider a pendant tree to be rooted at its vertex of attachment; in particular, the* parent *of a vertex u not in the 2-core is its unique neighbour on the path from u to the 2-core.*

Lemma 7.6 implies that we can apply Definition 7.9 to the graph induced by any Δ-set, $\sigma\Delta\sigma'$. So for every vertex $v \in \sigma\Delta\sigma'$ that is not in the 2-core of $\sigma\Delta\sigma'$, we can talk about its parent in $\sigma\Delta\sigma'$. By Property 7.4 and Proposition 7.7, we can also talk about the parent of any non-2-core vertex in $K_{a,b}$.

Lemma 7.10. *Consider any non-2-core vertex u in a Δ-set $\sigma\Delta\sigma'$ and let w be the parent of u in $\sigma\Delta\sigma'$.*

(a) $\sigma'(u) = \sigma(w)$;

(b) u is not in the 2-core of $K_{\sigma(u),\sigma'(u)}$;

(c) w is the parent of u in $K_{\sigma(u),\sigma'(u)}$.

Proof Let T be the pendant tree of $\sigma\Delta\sigma'$ containing u. We proceed by induction.

Note that since $K_{\sigma(u),\sigma'(u)}$ is connected and has size greater than one (by Property 7.4 and Lemma 6.3), u has at least one neighbour in $K_{\sigma(u),\sigma'(u)}$.

Base case: u is a leaf of T. Then w is the only neighbour of u in $\sigma\Delta\sigma'$, and so by Proposition 7.5, w must be the only neighbour of u in $K_{\sigma(u),\sigma'(u)}$. Parts (a,b,c) now follow.

Now suppose (a,b,c) hold for every child of u in T.

By Proposition 7.5, every neighbour of u in $K_{\sigma(u),\sigma'(u)}$ must be in T. Consider any child w' of u in T with $\sigma(w') = \sigma'(u)$. By the inductive hypothesis, w' is not in the 2-core of $K_{\sigma(w'),\sigma'(w')}$, $\sigma'(w') = \sigma(u)$ and u is the parent of w' in $K_{\sigma(w'),\sigma'(w')}$. Thus we have $\{\sigma(u), \sigma'(u)\} = \{\sigma(w'), \sigma'(w')\}$ and so u is the parent of w' in $K_{\sigma(u),\sigma'(u)}$. So u has at most one non-child neighbour, w, in $K_{\sigma(u),\sigma'(u)}$. Parts (a,b,c) now follow. □

Lemma 7.11. *Consider any vertex v in the 2-core of a Δ-set $\sigma\Delta\sigma'$. Suppose v is not in the 2-core of $K_{\sigma(v),\sigma'(v)}$. Then the parent of v in $K_{\sigma(v),\sigma'(v)}$ is in the 2-core of $\sigma\Delta\sigma'$.*

Proof Let w be the parent of v in $K_{\sigma(v),\sigma'(v)}$; so $\sigma(w) = \sigma'(v)$. By Proposition 7.5, $w \in \sigma\Delta\sigma'$. If w is not in the 2-core of $\sigma\Delta\sigma'$, then because v is a neighbour of w and v is in the 2-core of $\sigma\Delta\sigma'$, v must be the parent of w in $\sigma\Delta\sigma'$. By Lemma 7.10(a,c), this implies that $\sigma'(w) = \sigma(v)$ and v is the parent of w in $K_{\sigma(w),\sigma'(w)}$. But now $\sigma(w) = \sigma'(v)$ and $\sigma'(w) = \sigma(v)$ so $K_{\sigma(w),\sigma'(w)} = K_{\sigma(v),\sigma'(v)}$; thus v is the parent of w and w is the parent of v in the same graph - contradiction. Therefore, w is in the 2-core of $\sigma\Delta\sigma'$. □

At this point, an intuitive outline of the next steps might be helpful. Our goal is to show that the 2-core, H, of a complex Δ-set must have linear size. If we could prove that we must have $E(H) \geq (1+\epsilon)V(H)$ for some constant $\epsilon > 0$ then we would win: short standard arguments show that w.h.p. all subgraphs of this density have size at least αn for some $\alpha = \alpha(\epsilon, c) > 0$. So most of the work is dealing with the case $E(H) < (1+\epsilon)V(H)$. In that case, the average degree in H is very close to 2 and so most of the vertices have degree 2. Since no component of H is a cycle (as the Δ-set is complex), it follows that most of the vertices of H must lie on very long paths of degree 2 vertices. We begin by studying the structure of those paths.

We start by focussing on two basic types of paths through vertices that have degree 2 in the 2-core of some Δ-set $\sigma\Delta\sigma'$. These are illustrated in Figures 1,2. We will see that every path through vertices of degree 2 in the 2-core can be decomposed into a small number of these two types of paths (Lemma 7.17).

Type A is a path $u_0, u_1, ..., u_t$ where there is some $a \neq b$ such that each u_i has degree 2 in the 2-core of $K_{a,b}$. It will follow that each u_i has $(\sigma(u_i), \sigma'(u_i)) = (a,b)$ or (b,a) depending on the parity of i (see Corollary 7.14).

Type B is a path $u_0, u_1, ..., u_t$ in which each u_i is not in the 2-core of $K_{\sigma(u_i),\sigma'(u_i)}$, and its parent in $K_{\sigma(u_i),\sigma'(u_i)}$ is u_{i+1}.

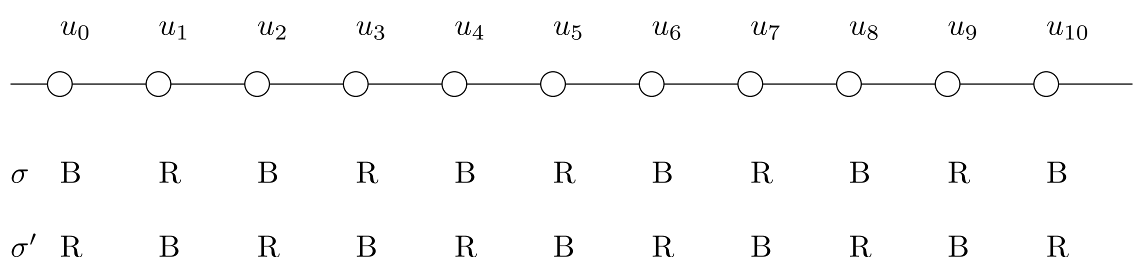


Figure 1: Type A path. Each u_i is in the 2-core of $K_{R,B}$.

To understand these two types of paths, it is important to recall Proposition 7.5. For any path $u_0, u_1, ..., u_t$ of degree 2 vertices in $\sigma\Delta\sigma'$, every neighbour of u_i in $K_{\sigma(u_i),\sigma'(u_i)}$ must also be in $\sigma\Delta\sigma'$. Type A: If u_i is in the 2-core of $K_{\sigma(u_i),\sigma'(u_i)}$ then it has 2 other neighbours in the 2-core - they are its neighbours on the path. Perhaps u_i has other neighbours in $K_{\sigma(u_i),\sigma'(u_i)}$ outside of the 2-core; in this case, they will be outside of

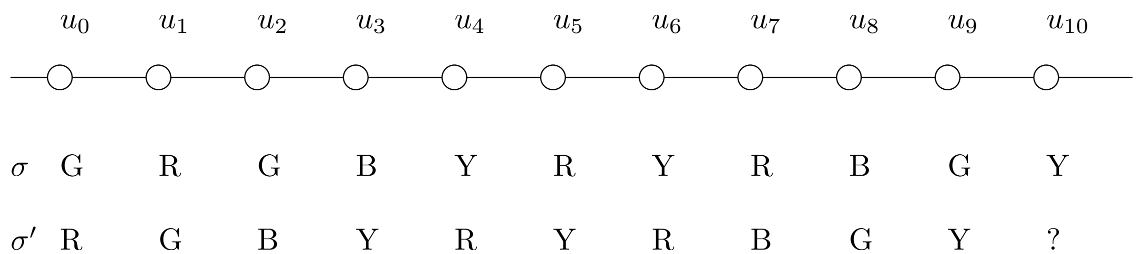


Figure 2: Type B path. For example, u_0 is not in the 2-core of $K_{G,R}$ and u_1 is the parent of u_0 in $K_{G,R}$.

the 2-core of $\sigma\Delta\sigma'$ and so are not on the path. Type B: If u_i is not in the 2-core of $K_{\sigma(u_i),\sigma'(u_i)}$ but is in the 2-core of $\sigma\Delta\sigma'$ then its parent in $K_{\sigma(u_i),\sigma'(u_i)}$ must also be in the 2-core of $\sigma\Delta\sigma'$ (by Lemma 7.11); that parent is u_{i+1}. Perhaps u_i has other neighbours in $K_{\sigma(u_i),\sigma'(u_i)}$; perhaps one of those neighbours is u_{i-1}; there may be others outside of the 2-core of $\sigma\Delta\sigma'$.

These two path types are defined more formally below. Note that this definition does not specify the colour pattern for the Type A path; that pattern will follow from Corollary 7.14.

Definition 7.12. *A* basic 2-path *in a Δ-set $\sigma\Delta\sigma'$ is a path $u_0, ..., u_x$ in the 2-core of $\sigma\Delta\sigma'$ such that*

(a) $x \geq 1$;

(b) each u_i has degree 2 in the 2-core of $\sigma\Delta\sigma'$;

(c) either
Type A: *every u_i is in the 2-core of $K_{\sigma(u_i),\sigma'(u_i)}$; or*
Type B: *every u_i is not in the 2-core of $K_{\sigma(u_i),\sigma'(u_i)}$ and, for $0 \leq i \leq x-1$, its parent in $K_{\sigma(u_i),\sigma'(u_i)}$ is u_{i+1}.*

We call u_0, u_x the endpoints*, and $u_1, ..., u_{x-1}$ the* internal vertices*. (So if $x = 1$ then there are no internal vertices.)*

We now prove that a Type A 2-path must have the colour pattern indicated in Figure 1.

Lemma 7.13. *Consider any vertex v in a Δ-set $\sigma\Delta\sigma'$ with 2-core H. Suppose that v has degree 2 in H, and let y, z be the two neighbours of v in H. If v is in the 2-core of $K_{\sigma(v),\sigma'(v)}$ then:*

(a) v has degree 2 in the 2-core of $K_{\sigma(v),\sigma'(v)}$;

(b) $\sigma(y) = \sigma(z) = \sigma'(v)$;

(c) y, z are both in the 2-core of $K_{\sigma(v),\sigma'(v)}$.

Proof Since v is in the 2-core of $K_{\sigma(v),\sigma'(v)}$, v has at least two neighbours in the 2-core of $K_{\sigma(v),\sigma'(v)}$. By Proposition 7.5, all of those neighbours are in $\sigma\Delta\sigma'$. We will argue that those neighbours must, in fact, be in the 2-core H of $\sigma\Delta\sigma'$ and hence must be y, z.

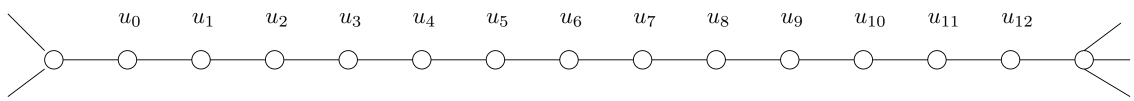
Suppose, to the contrary, that $w \in K_{\sigma'(v)}$ is a neighbour of v in the 2-core of $K_{\sigma(v),\sigma'(v)}$ and $w \notin H$. Since $v \in H$, and v is adjacent to w, v must be the parent of w in $\sigma\Delta\sigma'$. By Lemma 7.10(a), this implies that $\sigma'(w) = \sigma(v)$ and since $w \in K_{\sigma'(v)}$ we have $\sigma(w) = \sigma'(v)$. So $K_{\sigma(w),\sigma'(w)} = K_{\sigma(v),\sigma'(v)}$ and, since w was chosen to be in the 2-core of $K_{\sigma(v),\sigma'(v)}$, w is in the 2-core of $K_{\sigma(w),\sigma'(w)}$. This contradicts Lemma 7.10(b).

Therefore v has exactly two neighbours in the 2-core of $K_{\sigma(v),\sigma'(v)}$, and y, z are those neighbours; so $y, z \in K_{\sigma'(v)}$. Parts (a,b,c) follow immediately. $\square$

This immediately yields:

Corollary 7.14. *If $u_0, ..., u_x$ is a Type A 2-path in a Δ-set $\sigma\Delta\sigma'$, then there are colours a, b such that $(\sigma(u_i), \sigma'(u_i)) = (a, b)$ for even i, and $(\sigma(u_i), \sigma'(u_i)) = (b, a)$ for odd i; i.e. the sequences $\sigma(u_i)$ and $\sigma'(u_i)$ both alternate over the same two colours. Furthermore, if v, w are the other neighbours of u_0, u_x, respectively (i.e. $v \neq u_1, w \neq u_{x-1}$), then $\sigma(v) = \sigma'(u_0)$ and $\sigma(w) = \sigma'(u_x)$.*

Next we prove that every path through degree 2 vertices in the 2-core of a Δ-set must be composed from basic 2-paths. Figure 3 gives an example of a path composed of 3 basic 2-paths, which is the maximum number possible.



σ	R	Y	G	B	R	B	R	B	R	B	P	R	P
σ'	?	R	Y	G	B	R	B	R	B	P	R	P	?

Figure 3: u_4, u_5, u_6, u_7, u_8 form a Type A Path. u_3, u_2, u_1, u_0 and $u_9, u_{10}, u_{11}, u_{12}$ form two Type B paths.

Remark: The reader may have noticed an asymmetry between colourings σ, σ' in Figure 3, namely that vertices u_3, u_9 continue the R/B pattern in σ but not in σ'. This asymmetry arises from the fact that the sets $K_{a,b}$ are defined in terms of σ.

Setup for Lemmas 7.15 - 7.20: Let H be the 2-core of a Δ-set $\sigma\Delta\sigma'$. Consider any path $W, v_0, ..., v_r, Y$, $r \geq 1$, in H where each v_i has degree exactly 2 in H and W, Y each have degree at least 3 in H. It will be convenient to set $v_{-1} = W$ and $v_{r+1} = Y$. The next two lemmas concern this path.

Lemma 7.15. *Consider any $0 \leq i \leq r$. If v_i is not in the 2-core of $K_{\sigma(v_i),\sigma'(v_i)}$ then*

(a) one of v_{i-1}, v_{i+1} is the parent of v_i in $K_{\sigma(v_i),\sigma'(v_i)}$;

(b) if that parent is not W or Y, then either $v_i, v_{i+1}, ..., v_r$ or $v_i, v_{i-1}, ..., v_0$ is a Type B 2-path.

Proof v_i is in the 2-core, H, of $\sigma\Delta\sigma'$ but not in the 2-core of $K_{\sigma(v_i),\sigma'(v_i)}$. Let w be the parent of v_i in $K_{\sigma(v_i),\sigma'(v_i)}$. Lemma 7.11 says that w is in the 2-core of $\sigma\Delta\sigma'$. Since v_{i-1}, v_{i+1} are the only two neighbours of v_i in $\sigma\Delta\sigma'$, this establishes part (a).

WLOG, assume w is v_{i+1}. To prove part (b), we can assume that $v_{i+1} \neq Y$. We will show that v_{i+1} is not in the 2-core of $K_{\sigma(v_{i+1}),\sigma'(v_{i+1})}$, which will allow us to apply part (a) to v_{i+1}.

Suppose, to the contrary, that v_{i+1} is in the 2-core of $K_{\sigma(v_{i+1}),\sigma'(v_{i+1})}$. Then applying Lemma 7.13(b,c) to v_{i+1} implies that $\sigma(v_i) = \sigma'(v_{i+1})$ and v_i is in the 2-core of $K_{\sigma(v_{i+1}),\sigma'(v_{i+1})}$. But $v_{i+1} = w$ is adjacent to v_i in $K_{\sigma(v_i),\sigma'(v_i)}$ and so $\sigma(v_{i+1}) = \sigma'(v_i)$. Thus $K_{\sigma(v_i),\sigma'(v_i)} = K_{\sigma(v_{i+1}),\sigma'(v_{i+1})}$, and we have contradicted the fact that v_i is not in the 2-core of $K_{\sigma(v_i),\sigma'(v_i)}$.

Therefore, we have shown that v_{i+1} is not in the 2-core of $K_{\sigma(v_{i+1}),\sigma'(v_{i+1})}$ and so we can apply part (a) to v_{i+1} to show that the parent of v_{i+1} in $K_{\sigma(v_{i+1}),\sigma'(v_{i+1})}$ is either v_i or v_{i+2}. It is not possible for v_i to be that parent as this would require $\sigma(v_i) = \sigma'(v_{i+1})$, and so $K_{\sigma(v_{i+1}),\sigma'(v_{i+1})} = K_{\sigma(v_i),\sigma'(v_i)}$, and so v_i would be the parent of v_{i+1} in the same graph in which v_{i+1} is the parent of v_i. Therefore, v_{i+2} must be the parent of v_{i+1} in $K_{\sigma(v_{i+1}),\sigma'(v_{i+1})}$. Continuing this argument down the path establishes part (b). $\square$

Definition 7.16. *A* piece *of a path $v_0, ..., v_r$ is the subpath formed by a contiguous subsequence $v_i, ..., v_j$.*

Lemma 7.17. *$v_0, ..., v_r$ can be split into at most three pieces, each of which either has exactly one vertex or is a basic 2-path.*

Proof If $r \leq 2$ then it is trivial. So we assume $r \geq 3$.

Case 1: Some v_i is in the 2-core of $K_{\sigma(v_i),\sigma'(v_i)}$. Let $a = \sigma(v_i)$ and $b = \sigma'(v_i)$. Let j_1 be the largest $0 \leq j_1 < i$ such that $(\sigma(v_{j_1}), \sigma'(v_{j_1}))$ is not either (a,b) or (b,a); if no such j_1 exists then we set $j_1 = -1$. Similarly, let j_2 be the largest $i < j_2 \leq r$ such that $(\sigma(v_{j_2}), \sigma'(v_{j_2}))$ is not either (a,b) or (b,a); if no such j_2 exists then we set $j_2 = r+1$.

For all $j_1 + 1 \leq j \leq j_2 - 1$, $(\sigma(v_j), \sigma'(v_j))$ is either (a,b) or (b,a). This allows us to apply Lemma 7.13 inductively from v_i to v_{j_1+1} and from v_i to v_{j_2-1} and show that each such v_j is in the 2-core of $K_{a,b}$. Therefore the subpath $v_{j_1+1}, ..., v_{j_2-1}$ either has exactly one vertex (v_i) or is a Type A 2-path.

If $j_2 \leq r$ then Lemma 7.13, applied to v_{j_2-1}, implies that $\sigma(v_{j_2}) = \sigma'(v_{j_2-1})$. If v_{j_2} were in the 2-core of $K_{\sigma(v_{j_2}),\sigma'(v_{j_2})}$ then Lemma 7.13 applied to v_{j_2} would imply that $\sigma'(v_{j_2}) = \sigma(v_{j_2-1})$, and so $(\sigma(v_{j_2}), \sigma'(v_{j_2}))$ is either (a,b) or (b,a) thus contradicting our choice of j_2. So we can apply Lemma 7.15 to v_{j_2} to show that the subpath $v_{j_2}, ..., v_r$ either has exactly one vertex (v_r) or is a Type B 2-path. (Note that $v_{j_2}, ..., v_0$ cannot be the Type B 2-path since $v_{j_2-1} \in K_{\sigma(v_{j_2-1}),\sigma'(v_{j_2-1})}$.)

Similarly, if $j_1 \geq 0$ then the subpath $v_{j_1}, ..., v_0$ either has exactly one vertex or is a Type B 2-path. This provides our split into at most three pieces.

Case 2: No v_i is in the 2-core of $K_{\sigma(v_i),\sigma'(v_i)}$. Recall we assume that $r \geq 3$, and pick some $1 \leq \ell \leq r-1$. Since we are in Case 2, v_ℓ is not in the 2-core of $K_{\sigma(v_\ell),\sigma'(v_\ell)}$. So Lemma 7.15 and the fact that $\ell \notin \{0, r\}$ implies that v_ℓ lies in a Type B 2-path extending to either v_0 or v_r; WLOG assume it is v_r. Let $j \leq \ell$ be the smallest value such that $v_j, ..., v_r$ is a Type B 2-path. If $j = 0$ then we have one piece. If $j \geq 1$ then, since we are in Case 2, v_{j-1} is not in the 2-core of $K_{\sigma(v_{j-1}),\sigma'(v_{j-1})}$. By Lemma 7.15 and our choice of j, $v_{j-1}, v_{j-2}, ..., v_0$ either has exactly one vertex ($j - 1 = 0$) or is a Type B 2-path. Thus we can split into two pieces. □

As we discussed above, if we could prove that the 2-core of $\sigma\Delta\sigma'$ must have edge-density at least $1+\epsilon$, then we would win - a standard first-moment argument shows that every subgraph of that density must be large. But we cannot show this; long paths of degree 2 vertices can bring the edge-density arbitrarily close to 1. So we use an approach first applied in [60]: we contract those paths to obtain a dense graph, and then adapt that first-moment argument to show that this contracted graph must be large. One of the keys to this argument is a good understanding of those contracted paths - so we only contract basic 2-paths (which we understand well). The next lemma will apply Lemma 7.17 to show that we can obtain a useful contraction. First we define a contractable collection of basic 2-paths.

Definition 7.18. *We choose $\mathcal{P}(\sigma\Delta\sigma')$ to be a vertex-disjoint collection of basic 2-paths in the 2-core of $\sigma\Delta\sigma'$ such that: For every path $W, v_0, ..., v_r, Y$ in the 2-core of $\sigma\Delta\sigma'$, where each v_i has degree exactly 2 in the 2-core of $\sigma\Delta\sigma'$ and W, Y each have degree at least 3 in the 2-core of $\sigma\Delta\sigma'$, we can split $v_0, ..., v_r$ into at most three pieces, each of which either has exactly one vertex or is a member of $\mathcal{P}(\sigma\Delta\sigma')$.*

Lemma 7.17 implies that $\mathcal{P}(\sigma\Delta\sigma')$ exists. $\mathcal{P}(\sigma\Delta\sigma')$ might not be uniquely defined. It is possible that there are different ways to partition some $v_0, ..., v_r$ as in Lemma 7.17, thus yielding different choices for $\mathcal{P}(\sigma\Delta\sigma')$. If there are multiple choices for $\mathcal{P}(\sigma\Delta\sigma')$ then we arbitrarily specify one of them.

We partition the vertices of the 2-core of any Δ-set $\sigma\Delta\sigma'$ as follows:

- $V_1(\sigma\Delta\sigma')$ - the internal vertices of the basic 2-paths in $\mathcal{P}(\sigma\Delta\sigma')$;
- $V_2(\sigma\Delta\sigma')$ - the vertices of the 2-core of $\sigma\Delta\sigma'$ that are not in $V_1(\sigma\Delta\sigma')$; i.e. the endpoints of the basic 2-paths in $\mathcal{P}(\sigma\Delta\sigma')$ and the vertices that are not in those basic 2-paths.

Observation 7.19. *If $\sigma\Delta\sigma$ is a complex Δ-set then $|V_2(\sigma\Delta\sigma')| \geq 1$.*

Proof By Lemma 7.6, the 2-core of $(\sigma\Delta\sigma')$ is not empty. Since $\sigma\Delta\sigma'$ is complex, its 2-core has no cycle-components. Therefore the 2-core must have at least one vertex of degree at least 3. That vertex cannot be in $V_1(\sigma\Delta\sigma')$, so it must be in $V_2(\sigma\Delta\sigma')$. □

Lemma 7.20. *For any complex Δ-set $\sigma\Delta\sigma'$ with $|\mathcal{P}(\sigma\Delta\sigma')| = t$, the 2-core of $\sigma\Delta\sigma'$ has at least $\frac{201}{200}|V_2(\sigma\Delta\sigma')| - t$ edges with both endpoints in $V_2(\sigma\Delta\sigma')$.*

Proof Let H be the 2-core of $\sigma\Delta\sigma'$. Form H' by contracting every basic 2-path $u_0, ..., u_x$ in $\mathcal{P}(\sigma\Delta\sigma')$ into a single edge (u_0, u_x). For example, the path of degree 2 vertices in Figure 3 becomes the path $u_0, u_3, u_4, u_8, u_9, u_{12}$.

Every vertex in H' has the same degree in H' as in H. Since no component of H is a cycle (as $\sigma\Delta\sigma'$ is complex), every degree 2 vertex of H lies in a path $W, v_0, ..., v_r, Y$ as in Definition 7.18. Every such path is contracted into a path with at most six degree two vertices. Therefore, H' does not contain any path $v_0, v_1, v_2, v_3, v_4, v_5, v_6$ of seven degree 2 vertices. Furthermore, H' has minimum degree at least 2. From that, it is easy to argue that H' has at least $\frac{201}{200}|V(H')|$ edges (this also follows from Lemma 11 of [60]). The lemma now follows since $V(H') = V_2(\sigma\Delta\sigma')$, there are exactly t contracted edges in H', and each of the $\frac{201}{200}|V_2(\sigma\Delta\sigma')| - t$ non-contracted edges is an edge of H. $\square$

We will use Lemma 7.20 to prove that every complex Δ-set is large in the next subsection. We close this subsection by analyzing the structure of cyclic Δ-sets.

Lemma 7.21. *If $u_1, ..., u_r$ is the cycle forming the 2-core of a cyclic Δ-set $\sigma\Delta\sigma'$, then (after possibly reversing the order of the labels): Every u_i is not in the 2-core of $K_{\sigma(u_i),\sigma'(u_i)}$ and its parent in $K_{\sigma(u_i),\sigma'(u_i)}$ is u_{i+1} (addition mod r).*

Thus, we can view this 2-core as the cycle analogue of a Type B 2-path.

Proof If at least one u_i is not in the 2-core of $K_{\sigma(u_i),\sigma'(u_i)}$, then the same reasoning as in the proof of Lemma 7.15(a) implies that its parent in $K_{\sigma(u_i),\sigma'(u_i)}$ is either u_{i-1} or u_{i+1}; WLOG assume it is u_{i+1}. The same reasoning as in the proof of Lemma 7.15(b) implies that u_{i+1} cannot be in the 2-core of $K_{\sigma(u_{i+1}),\sigma'(u_{i+1})}$. So we can repeat the argument inductively around the cycle to prove that the lemma holds.

So assume that every u_i is in the 2-core of $K_{\sigma(u_i),\sigma'(u_i)}$. Lemma 7.13 now implies that every vertex u_j has $(\sigma(u_j), \sigma'(u_j)) = (a, b)$ or (b, a) for the same two colours (b, a), just as it implied Corollary 7.14. Furthermore Lemma 7.13(a) implies that each u_j has no other neighbours in the 2-core of $K_{a,b}$, other than its neighbours in the cycle. It follows that this cycle is a component of the 2-core of $K_{a,b}$, contradicting Property 7.4 and Propostition 7.7. $\square$

We close by noting how Lemma 5.7 follows from Lemma 7.2:

Proof of Lemma 5.7: Let σ' be any k-colouring of K with $\sigma'(v) \neq \sigma(v)$; i.e. with $v \in \sigma\Delta\sigma'$. By Observation 7.3, we can assume that $\sigma\Delta\sigma'$ is connected; otherwise replace σ' with σ'' such that $\sigma\Delta\sigma''$ is the component of $\sigma\Delta\sigma'$ containing v.

Lemma 7.6 implies that $\sigma\Delta\sigma'$ is not a tree. Since v does not lie in a flippable unicycle, $\sigma\Delta\sigma'$ has more than one cycle. So this lemma follows from Lemma 7.2. $\square$

The next two subsections are devoted to proving Lemma 7.2.

7.2 A first moment bound for Δ-sets

Proof of Lemma 7.2

We will bound the expected number of complex Δ-sets in terms of various size-parameters. We will focus on the 2-cores of the Δ-sets.

Let $\sigma\Delta\sigma'$ be a complex Δ-set (Definition 7.18), and define:

- $a = |V_2(\sigma\Delta\sigma')|$
- $t = |\mathcal{P}(\sigma\Delta\sigma')|$
- $j_1, ..., j_t \geq 0$ are the number of internal vertices in the basic 2-paths of $\mathcal{P}(\sigma\Delta\sigma')$
- $J = j_1 + ... + j_t$

Let $X_{a,J}$ denote the number of 2-cores of complex Δ-sets with parameters a, J. We bound $E(X_{a,J})$ for all $a + J < 2\alpha n$ as follows:

First we choose the a vertices of $V_2(\sigma\Delta\sigma')$. We will overcount by choosing from all of $\{1, ..., n\}$ rather than from just the Kempe-core. So the number of choices is at most

$$\binom{n}{a}.$$

Next we choose the values of t and $j_1, ..., j_t$. Since the basic 2-paths of $\mathcal{P}(\sigma\Delta\sigma')$ are vertex-disjoint (by Definition 7.18) and each has two endpoints in $V_2(\sigma\Delta\sigma')$, we have $t \leq \frac{a}{2}$.

Recalling Lemma 7.20, we choose a set $\mathcal{E}$ of $\lceil \frac{201}{200}a \rceil - t$ edges within $V_2(\sigma\Delta\sigma')$; the number of choices for $\mathcal{E}$ is:

$$\binom{\binom{a}{2}}{\lceil \frac{201}{200}a \rceil - t}.$$

Then we choose, from amongst the vertices of $V_2(\sigma\Delta\sigma')$, the endpoints (v_i, w_i) of each of the basic 2-paths $P_1, ..., P_t \in \mathcal{P}(\sigma\Delta\sigma')$. If two choices result in the same set of paths, just with a permutation of the indices, then we consider those two choices to be equivalent. The number of choices is at most:

$$a^{2t}/t!$$

(Note that we do not divide by 2^t since the direction matters on a Type B path.)

We define the following events:

- E_1 - the event that the statements of Lemma 6.3(a,b,c,d) hold.
- E_2 - the event that all the edges of $\mathcal{E}$ are present.
- E_3 - the event that each pair (v_i, w_i) is joined by a basic 2-path.

For a random variable X and an event E, we use $X \wedge E$ to denote the variable that is equal to X if E holds and 0 if E does not hold. We will actually bound $E(X_{a,J} \wedge E_1)$, recalling from Lemma 6.3 that E_1 holds w.h.p.

We begin by noting that, since $t \leq \frac{a}{2}$, we have $\frac{201}{200}a - t > \frac{a}{2}$. This yields:

$$\binom{\binom{a}{2}}{\lceil \frac{201}{200}a \rceil - t} \times \mathbf{Pr}(E_2) = \binom{\binom{a}{2}}{\lceil \frac{201}{200}a \rceil - t} \left(\frac{c}{n}\right)^{\lceil \frac{201}{200}a \rceil - t} \leq \left(\frac{e\frac{a^2}{2}}{\lceil \frac{201}{200}a - t \rceil}\frac{c}{n}\right)^{\lceil \frac{201}{200}a \rceil} - t < \left(\frac{eca}{n}\right)^{\lceil \frac{201}{200}a \rceil - t}. \quad (4)$$

Recall the constant $\zeta = \zeta(c) > 0$ from Lemma 6.4. In Section 7.3, we will prove:

Lemma 7.22. *There is a constant $R = R(c, k)$ such that if $a + J < 2\alpha n$ then*

$$\mathbf{Pr}(E_3 \wedge E_1 | E_2) < R^a (1 - \frac{\zeta}{2})^J \left(\frac{1}{n}\right)^t.$$

This yields that for $a + J < 2\alpha n$:

$$\begin{aligned}
E(X_{a,J} \wedge E_1) \quad &\leq \quad \sum_{t, j_1 + ... + j_t = J} \binom{n}{a} \frac{a^{2t}}{t!} \times \binom{\binom{a}{2}}{\lceil \frac{201}{200}a \rceil - t} \times \mathbf{Pr}(E_2) \times \mathbf{Pr}(E_3 \wedge E_1 | E_2) \\
&\leq \quad \sum_{t, j_1 + ... + j_t = J} \left(\frac{en}{a}\right)^a \frac{a^{2t}}{t!} \left(\frac{eca}{n}\right)^{\lceil \frac{201}{200}a \rceil - t} R^a (1 - \frac{\zeta}{2})^J \left(\frac{1}{n}\right)^t \qquad \text{by (4) and Lemma 7.22} \\
&< \quad \sum_{t \geq 0} \left(\frac{en}{a}\right)^a \frac{a^{2t}}{t!} \left(\frac{a}{n}\right)^{\lceil \frac{201}{200}a \rceil - t} \left(R(ec)^2\right)^a \left(\frac{1}{n}\right)^t \sum_{j_1 + ... + j_t = J} (1 - \frac{\zeta}{2})^J \qquad \text{as } \lceil \tfrac{201}{200}a \rceil \leq 2a \\
&< \quad \left(Z_1 \frac{a}{n}\right)^{\lceil \frac{a}{100} \rceil} \sum_{t \geq 0} \frac{a^t}{t!} \sum_{j_1 + ... + j_t = J} (1 - \frac{\zeta}{2})^J \qquad \text{for some constant } Z_1 = Z_1(c, k) > 0.
\end{aligned}$$

The number of choices for $j_1, ..., j_t \geq 0$ that sum to J is $\binom{J+t-1}{t-1}$. It is straightforward to verify that there is a constant $Z_2 = Z_2(\zeta) = Z_2(c,k) > 1$ such that for any t and $J \geq Z_2 t$, $\binom{J+t-1}{t-1}(1-\frac{\zeta}{4})^J$ is monotone decreasing as J increases. Thus, for $J \geq Z_2 t$, we have $\binom{J+t-1}{t-1}(1-\frac{\zeta}{4})^J < \binom{Z_2t+t-1}{t-1}(1-\frac{\zeta}{4})^{Z_2 t} < \binom{Z_2 t+t}{t}$, and for $J < Z_2 t$, we have $\binom{J+t-1}{t-1}(1-\frac{\zeta}{4})^J < \binom{J+t-1}{t-1} < \binom{Z_2t+t}{t}$. This, along with the bound $1 - \frac{\zeta}{2} < (1-\frac{\zeta}{4})^2$, implies:

$$\sum_{j_1+\ldots+j_t=J} (1-\frac{\zeta}{2})^J < \binom{J+t-1}{t-1}(1-\frac{\zeta}{4})^{2J} < \binom{Z_2 t+t}{t}(1-\frac{\zeta}{4})^J < (e(Z_2+1))^t(1-\frac{\zeta}{4})^J. \tag{5}$$

Thus, for $a + J \leq 2\alpha n$, we have:

$$\begin{aligned}
E(X_{a,J} \wedge E_1) &< \left(Z_1\frac{a}{n}\right)^{\lceil \frac{a}{100} \rceil}(1-\frac{\zeta}{4})^J \sum_{t\geq 0}\frac{(ea(Z_2+1))^t}{t!} \\
&= \left(Z_1\frac{a}{n}\right)^{\lceil \frac{a}{100} \rceil}(1-\frac{\zeta}{4})^J e^{ea(Z_2+1)} \\
&< \left(Z\frac{a}{n}\right)^{\lceil \frac{a}{100} \rceil}(1-\frac{\zeta}{4})^J \qquad \text{for some constant } Z = Z(c,k) > 0 \qquad (6)\\
&< (1-\frac{\zeta}{4})^{a+J}, \qquad (7)
\end{aligned}$$

by applying $a < 2\alpha n$ and taking $\alpha = \alpha(c,k)$ to be sufficiently small that $2Z\alpha < (1-\frac{\zeta}{4})^{100}$.

We apply (6) to small values of a, say $a \leq \log^2 n$, and (7) to larger values of a. Recalling Observation 7.19, we sum from $a \geq 1$. This yields:

$$\begin{aligned}
E(X_{a,J} \wedge E_1) &< \sum_{a=1}^{\log^2 n}\sum_{J=0}^{\alpha n - a}\left(Z\frac{a}{n}\right)^{\lceil \frac{a}{100} \rceil}(1-\frac{\zeta}{4})^J + \sum_{a=\log^2 n}^{\alpha n}\sum_{J=0}^{\alpha n - a}(1-\frac{\zeta}{4})^{a+J} \\
&< O(1) \times \sum_{a=1}^{\log^2 n}\left(Z\frac{a}{n}\right)^{\lceil \frac{a}{100} \rceil} + O(1) \times \sum_{a=\log^2 n}^{\alpha n}(1-\frac{\zeta}{4})^a \qquad (8)\\
&< O(\frac{1}{n}) + e^{-\frac{\zeta}{8}\log^2 n} = o(1).
\end{aligned}$$

Since E_1 holds w.h.p., this yields that w.h.p. the Kempe-core of $P_{n,p=c/n}$ has no complex Δ-sets of size less than αn, thus proving Lemma 7.2. $\square$

7.3 Proof of Lemma 7.22

We begin by restating the lemma:

Lemma 7.22 *There is a constant $R = R(c,k)$ such that if $a + J < 2\alpha n$ then* $\mathbf{Pr}(E_3 \wedge E_1 | E_2) < R^a(1-\frac{\zeta}{2})^J\left(\frac{1}{n}\right)^t$.

Proof At this point, we have chosen the vertices of $V_2(\sigma\Delta\sigma')$, the endpoints v_i, w_i for $i = 1, ..., t$, and exposed the fact that the fewer than $2a$ edges of $\mathcal{E}$ are present. Next, we will expose the values of the parameters bounded by Lemma 6.3. If E_1 holds then the following hold for each a, b:

(Q1) $|K_a| = \lambda_k(c)n + o(n)$;
(Q2) the 2-core of $K_{a,b}$ has $\xi_k(c)n + o(n)$ vertices in K_a and $\xi_k(c)n + o(n)$ vertices in K_b;
(Q3) the 2-core of $K_{a,b}$ has $\mu_k n + o(n)$ edges;
(Q4) the 2-core of $K_{a,b}$ has $\tau_k n + o(n)$ degree 2 vertices in K_a and $\tau_k n + o(n)$ degree 2 vertices in K_b.

Let $P_1, ..., P_t$ be the basic 2-paths of $\mathcal{P}(\sigma\Delta\sigma')$, where P_i has j_i internal vertices plus endpoints $v_i, w_i \in V_2(\sigma\Delta\sigma')$.

Before heading into the details, we comment on our goal: Note that our desired bound is very loose in the R^a term - any constant R will do - but much more precise in the $(1-\frac{\zeta}{2})^J$ term - we need a constant less than 1. This corresponds to the intuition that the paths are very long and so the term corresponding to their total length, J, is more important than the term corresponding to their total number, t. So in what follows, we will be very careful about the constants raised to the Jth power, but we do not need to be careful with those raised to the Rth power.

First, we choose whether each P_i is Type A or Type B. We will consider the Type B paths first, so we relabel the indices so that the Type B 2-paths come first in the sequence $P_1, ..., P_t$.

Type B paths

We expose the Type B 2-paths one-at-a-time, and for each such P_i, we expose the vertices one-at-a-time beginning with the vertex after v_i.

Suppose P_i is a Type B 2-path in $\sigma\Delta\sigma'$. Denote the vertices of P_i by $v_i = u_0, u_1, ..., u_{j_i}, u_{j_i+1} = w_i$.

First, we choose the colour $\sigma(u_x)$ for each $1 \le x \le j_i$; note that $\sigma(u_0), \sigma(u_{j_i+1})$ were determined when we chose $u_0 = v_i, u_{j_i+1} = w_i$. Since we must have $\sigma(u_{x+1}) \neq \sigma(u_x)$, as those two vertices are adjacent and hence cannot have the same colour, there are:

$$(k-1)^{j_i} \text{ choices for these colours.}$$

By the definition of a Type B 2-path, our choice of $\sigma(u_1), ..., \sigma(u_{j_i+1})$ also determines $\sigma'(u_0), ..., \sigma'(u_{j_i})$ because $\sigma'(u_x) = \sigma(u_{x+1})$.

Next, we choose the verties $u_1, ..., u_{j_i}$. We begin with u_0, which was already chosen. We then choose u_1: the parent of u_0 in $K_{\sigma(u_0),\sigma'(u_0)}$ (by the definition of a Type B 2-path). We continue this way down the path.

Suppose that we have chosen vertex u_x and are now choosing u_{x+1}, for some $x < j_i$; the case $x = j_i$ is a special case, since $u_{j_i+1} = w_i$ has already been chosen. Prior to choosing u_{x+1}, we have exposed the following:

- the values of the parameters from Lemma 6.3; i.e. from (Q1), (Q2), (Q3), (Q4);
- the values of $a, t, j_1, ..., j_t$
- the vertices of $V_2(\sigma\Delta\sigma')$ including the endpoints of each basic 2-path;
- the edges of $\mathcal{E}$ (which all lie within $V_2(\sigma\Delta\sigma')$);
- the edges and vertices of the basic 2-paths $P_1, ..., P_{i-1}$;
- the colours $\sigma(u_0), ..., \sigma(u_{j_i+1}), \sigma'(u_0), ..., \sigma'(u_{j_i})$;
- the vertices $u_1, ..., u_x$ and the fact that each is adjacent to the preceding vertex;
- for $\ell = 1, ..., x$: u_ℓ is not in the 2-core of $K_{\sigma(u_{\ell-1}),\sigma'(u_{\ell-1})}$, and the parent of u_ℓ in $K_{\sigma(u_{\ell-1}),\sigma'(u_{\ell-1})}$ is $u_{\ell+1}$; the same is also true of all vertices in $P_1, ..., P_{i-1}$ other than the last endpoint of each path.

The only edges that have been exposed have both endpoints in: $V_2(\sigma\Delta\sigma'), P_1, ..., P_{i-1}, u_0, ..., u_x$; let Ψ be that set of vertices. Thus

$$|\Psi| \le a + J < 2\alpha n.$$

So u_x is not in the 2-core of $K_{\sigma(u_x),\sigma'(u_x)}$, and we are about to expose its parent, u_{x+1}. We will bound the probability that u_{x+1} is not in the 2-core of $K_{\sigma(u_{x+1}),\sigma'(u_{x+1})}$.

Note that if $u_{x+1} \in \Psi$, then we have failed to construct a Δ-set $\sigma\Delta\sigma'$ subject to the specified parameters. So to upper bound the probability that our choices yield such a $\sigma\Delta\sigma'$, we can assume $u_{x+1} \notin \Psi$.

Case 1: $\sigma'(u_{x+1}) \neq \sigma(u_x)$. We expose the parent of u_x in $K_{\sigma(u_x),\sigma'(u_x)}$, and set it to be u_{x+1}. This exposes nothing new about u_{x+1} in $K_{\sigma(u_{x+1}),\sigma'(u_{x+1})}$, as that is a different graph since we are in Case 1 and since $\sigma'(u_x) = \sigma(u_{x+1})$. So in the random graph $K_{\sigma(u_{x+1}),\sigma'(u_{x+1})}$, we have exposed nothing about the edges involving any vertices outside of Ψ.

Consider any graph H that, subject to what has already been exposed, could be $K_{\sigma(u_{x+1}),\sigma'(u_{x+1})}$. Consider any H' formed from H by permuting the vertices in $K_{\sigma(u_{x+1})}\backslash\Psi$. So H' could be $K_{\sigma(u_{x+1}),\sigma'(u_{x+1})}$, subject to what has been exposed, and $|E(H')| = |E(H)|$. The same argument used in the proof of Lemma 6.1 yields that conditional on what has already been exposed:

$$\mathbf{Pr}(K_{\sigma(u_{x+1}),\sigma'(u_{x+1})} = H) = \mathbf{Pr}(K_{\sigma(u_{x+1}),\sigma'(u_{x+1})} = H').$$

Therefore, the probability that u_{x+1} is not in Ψ and is not in the 2-core of $K_{\sigma(u_{x+1}),\sigma'(u_{x+1})}$ is at most the number of non-2-core vertices in $K_{\sigma(u_{x+1})}$ divided by $|K_{\sigma(u_{x+1})}\backslash\Psi|$. Using the fact that (Q1), (Q2) hold, applying Lemma 6.4(a), using $|\Psi| < 2a$, and taking α sufficiently small in terms of ζ, this ratio is at most:

$$\frac{\lambda_k(c) - \xi_k(c)}{\lambda_k(c) - 2\alpha} + o(1) < \frac{1}{k-1}(1 - \frac{\zeta}{2}). \tag{9}$$

Case 2: $\sigma'(u_{x+1}) = \sigma(u_x)$. We argue as in Case 1, except this case is more delicate since $K_{\sigma(u_x),\sigma'(u_x)} = K_{\sigma(u_{x+1}),\sigma'(u_{x+1})}$. When we expose the parent of u_x in this graph, we need to bound the probability of that parent being outside the 2-core.

We have exposed that u_x is not in the 2-core of $K_{\sigma(u_x),\sigma'(u_x)}$. So if we remove the edge from u_x to its parent, $K_{\sigma(u_x),\sigma'(u_x)}$ will be disconnected into two components: one containing u_x and one containing the 2-core. Let H_1, H_2 be any two graphs which, subject to what has already been exposed, could be these two components.

Let $\Upsilon \subset H_2$ denote the set of vertices $w \in H_2$ such that, if we add an edge from u_x to w, then the graph H_w formed by H_1, H_2 and the edge (u_x, w), could be $K_{\sigma(u_x),\sigma'(u_x)}$ (subject to what has already been exposed). Consider any two vertices $w, w' \in \Upsilon$; we will argue that $H_w, H_{w'}$ are each equally likely to be $K_{\sigma(u_x),\sigma'(u_x)}$.

So consider any k-partite graph G with parts $A_1, ..., A_k$ (from Definition 4.2) for which, in the Kempe-core of G, we have $K_{\sigma(u_x),\sigma'(u_x)} = H_w$. Let G' be the graph obtained from G by replacing the edge (u_x, w) by (u_x, w'). So, in the Kempe-core of G', we have $K_{\sigma(u_x),\sigma'(u_x)} = H_{w'}$. Note further that swapping these two edges does not alter any of the information that has been exposed thus far. Therefore G could possibly be the original random graph drawn from $P_{n,p}$, subject to what has been exposed up to this point, iff G' could.

So this is a bijection from the possible graphs with $K_{\sigma(u_x),\sigma'(u_x)} = H_w$ and those with $K_{\sigma(u_x),\sigma'(u_x)} = H_{w'}$. For each pair (G, G') in that bijection, G, G' have the same number of edges and so are equally likely to be chosen as $P_{n,p}$. This implies that $H_w, H_{w'}$ are each equally likely to be $K_{\sigma(u_x),\sigma'(u_x)}$, conditional on what has been exposed thus far.

Now we note that Υ contains all vertices in $K_{\sigma'(u_x)}\backslash\Psi$ that are in the 2-core of H_2 (as well as, perhaps, some other vertices). By (Q2) and since $|\Psi| \le \alpha n$, there are at least $(\xi_k(c) - 2\alpha)n + o(n)$ such vertices and, by (Q1), $|\Upsilon| \le |K_a| = \lambda_k(c)n + o(n)$. So, conditioning on the event that H_1, H_2 are the components created by removing the edge from u_x to its parent, the probability that the parent of u_x is in the 2-core of $K_{\sigma(u_x),\sigma'(u_x)}$ is at least $(\xi_k(c) - 2\alpha)/\lambda_k(c) + o(1)$. Since this holds for any possible choice of H_1, H_2, the probability that the parent of x in H' is *not* in the 2-core is, by Lemma 6.4(a), at most

$$1 - \frac{\xi_k(c) - 2\alpha}{\lambda_k(c)} + o(1) < \frac{1}{k-1}(1 - \frac{\zeta}{2}), \tag{10}$$

if α is sufficiently small in terms of ζ.

So in both Case 1 and Case 2, we find that the probability that u_{x+1} is not in the 2-core of $K_{\sigma(u_{x+1}),\sigma'(u_{x+1})}$, conditional on what has been exposed thus far, is at most $\frac{1}{k-1}(1 - \frac{\zeta}{2})$.

To be clear: having bounded the *probability* that u_{x+1} is in the 2-core of $K_{\sigma(u_{x+1}),\sigma'(u_{x+1})}$, we now expose the vertex u_{x+1} and whether it is in that 2-core. If it is, then we halt our process having failed to produce the set of paths $P_1, ..., P_t$. If we continue, we have exposed that u_{x+1} is in that 2-core. We have not exposed any of the information used in the *analysis* of the probability that u_{x+1} is in the 2-core; eg. the components H_1, H_2.

The final edge: After exposing $u_1, ..., u_{j_i}$, we turn our attention to the edge (u_{j_i}, u_{j_i+1}). We have already exposed the vertices u_{j_i} and $u_{j_i+1} = w_i$, and we have exposed that u_{j_i} is not in the 2-core of $K_{\sigma(u_{j_i}),\sigma'(u_{j_i})}$. We will bound the probability that w_i is the parent of u_{j_i} in $K_{\sigma(u_{j_i}),\sigma'(u_{j_i})}$.

We consider H_1, H_2, Υ as in Case 2 above. Again, the size of Υ is at least the number of vertices in $K_{\sigma'(u_{j_i})} \backslash \Psi$ that are in the 2-core of H_2. So, by the same reasoning as in Case 2, if $w_i \in \Upsilon$ then the probability (conditional on what has been exposed so far, including H_1, H_2) that w_i is the parent of u_{j_i} is $1/|\Upsilon| \le 1/(\xi_k(c)n - 2\alpha n + o(n))$ and if $w_i \notin \Upsilon$ then the probability is zero. Since this is true for every choice of H_1, H_2, the conditional probability that w_i is the parent of u_{j_i} is at most:

$$\frac{1}{\xi_k(c)n - 2\alpha n + o(n)} < \frac{2}{\xi_k(c) - 2\alpha} \times \frac{1}{n}.$$

Putting this all together, each Type B 2-path P_i contributes to $\mathbf{Pr}(E_3 \wedge E_1 | E_2)$ a factor of at most

$$(k-1)^{j_i} \times (\frac{1}{k-1}(1 - \frac{\zeta}{2}))^{j_i} \times \frac{2}{\xi_k(c) - 2\alpha} \times \frac{1}{n}. \tag{11}$$

Type A paths

Next, we consider the Type A 2-paths. We will process them all at once.

At this point, we have chosen the following:

- the values of the parameters from Lemma 6.3;
- the values of $a, t, j_1, ..., j_t$
- the vertices of $V_2(\sigma\Delta\sigma')$ including the endpoints of each basic 2-path;
- the edges of $\mathcal{E}$ (which all lie within $V_2(\sigma\Delta\sigma')$);
- the edges and vertices of the Type B 2-paths;
- each vertex u in those Type B 2-paths (other than the last endpoint of each path) is not in the 2-core of $K_{\sigma(u_{\ell-1}),\sigma'(u_{\ell-1})}$, and its parent is the vertex that follows it on the path.

So again, the only edges that have been exposed have both endpoints in $V_2(\sigma\Delta\sigma')$ and the Type B 2-paths; we let Ψ be that set of vertices, and we have

$$|\Psi| \le a + J < 2\alpha n.$$

Now, we will choose the entire 2-core of every $K_{a,b}$. Recall that the edge-sets of each $K_{a,b}$ are independent of each other. Our first step is to expose the vertices of each 2-core, and the degree that each vertex has in the 2-core. As we are upper bounding the probability of $E_3 \wedge E_1$, we will assume that E_1 holds; in particular, we assume that properties (Q2), (Q3), (Q4) regarding the number of vertices, edges, and degree two vertices in the 2-core all hold.

We are conditioning on E_2, the event that the edges of $\mathcal{E}$ all appear; having chosen the vertices of the 2-core we may find that some of the edges of $\mathcal{E}$ are in the 2-core of $K_{a,b}$. None of the edges in the Type B 2-paths can be in the 2-core, so the edges of $\mathcal{E}$ are the only 2-core edges that have been exposed. We claim that every subgraph on the chosen vertices with the chosen degree sequence and containing those edges of $\mathcal{E}$ is equally likely to be the 2-core of $K_{a,b}$. To see this, consider two potential such 2-cores H_1, H_2, and any graph G satisfying everything exposed thus far such that H_1 is the 2-core of $K_{a,b}$ in G. Form G' by deleting the edges of H_1 and adding the edges of H_2. Note that H_2 is the 2-core of $K_{a,b}$ in G' and that G, G' are equally likely to be chosen as our random graph, as they have the same number of edges. It follows that we can use the configuration model[16] to choose our 2-core:

For each vertex v, we take $\deg(v)$ copies of v. For each edge $e \in \mathcal{E}$ in the 2-core, we take a copy of each of the two endpoints of e and pair them. Then we take a uniformly random bipartite matching of the remaining vertex-copies. We do this for the 2-core of every $K_{a,b}$. Standard arguments show that the probability that the resulting graph is simple is bounded away from zero. So if a property holds w.h.p. in this model, it holds w.h.p. upon conditioning on the graph being simple; i.e. for the correct model.

For each Type A 2-path P_i, following Lemma 7.14, we select the two colours on that path; i.e. the colours a, b such that for each $u \in P_i$ we have $(\sigma(u), \sigma'(u)) = (a, b)$ or (b, a). We have already chosen the endpoints of each path and have thus specified one or both of the colours a, b. So there are at most $k - 1$ choices for the second colour on each path. (To upper bound the probability of success, we will assume that the colours of the selected endpoints of P_i along with the parity of j_i do not conflict each other; i.e. the endpoints have the same colour iff the path has even length.)

For each a, b, we let $t_{a,b}$ denote the number of Type A 2-paths for which we selected the colours a, b in the preceding paragraph, and we let $J_{a,b}$ denote the total number of internal vertices on those $t_{a,b}$ paths. We let $J^a_{a,b}, J^b_{a,b}$ be the number of such vertices u for which we determined that $\sigma(u) = a, \sigma(u) = b$ resp.

We now choose the interior vertices for each Type A 2-path P_i. For each $K_{a,b}$, we must select $J^a_{a,b}, J^b_{a,b}$ vertices of colour a, b that have degree two in the 2-core of $K_{a,b}$ (by Definition 7.12) and are not incident with an edge of $\mathcal{E}$ (as the edges of $\mathcal{E}$ cannot be in basic 2-paths). Let $L^a_{a,b}, L^b_{a,b}$ be the number of such vertices to choose from in $K_{a,b}$. (To overcount, we'll consider all vertices of degree two in the 2-core of $K_{a,b}$, even though some are already known to be ineligible, eg. if they are in a previously exposed edge.) Since (Q4) holds, we have

$$L^a_{a,b}, L^b_{a,b} \le \tau_k(c)n + o(n).$$

Since the basic 2-paths are disjoint (by Definition 7.18), the number of choices for the degree 2 vertices is at most:

$$L^a_{a,b}(L^a_{a,b} - 1)...(L^a_{a,b} - J^a_{a,b} + 1)L^b_{a,b}(L^b_{a,b} - 1)...(L^b_{a,b} - J^b_{a,b} + 1).$$

Now we choose which vertex-copies of the vertices of each path, including the endpoints, will be matched with each other. The number of choices is at most $2^{J_{a,b}+2t_{a,b}}$. Finally, we bound the probability that these copies will be paired up. Because every edge in $K_{a,b}$ contains a vertex of each colour, the total number of vertex-copies from K_a in the 2-core of $K_{a,b}$ which do not lie in edges of $\mathcal{E}$, is the same as the total number from K_b; let $X_{a,b}$ be that number.

We proceed along the paths one-vertex-at-a-time, each time exposing whether the selected copy of that vertex is paired with the selected copy of the next vertex on the path. Every success removes a vertex-copy of each colour from the 2-core of $K_{a,b}$. So the probability that all of these $J_{a,b} + t_{a,b}$ pairings occur is:

$$\frac{1}{X_{a,b}(X_{a,b} - 1)...(X_{a,b} - (J_{a,b} + t_{a,b}) + 1)}.$$

This leads to the following bound on the probability that the $t_{a,b}$ Type A 2-paths that use edges from $K_{a,b}$ are formed:

$$\begin{aligned} &\frac{2^{J_{a,b}+2t_{a,b}} L^a_{a,b}...(L^a_{a,b} - J^a_{a,b} + 1)L^b_{a,b}...(L^b_{a,b} - J^b_{a,b} + 1)}{X_{a,b}...(X_{a,b} - (J_{a,b} + t_{a,b}) + 1)} \\ < \quad &\left(\frac{4}{X_{a,b} - J_{a,b} - t_{a,b}}\right)^{t_{a,b}} \frac{2^{J_{a,b}} L^a_{a,b}...(L^a_{a,b} - J^a_{a,b} + 1)L^b_{a,b}...(L^b_{a,b} - J^b_{a,b} + 1)}{X_{a,b}...(X_{a,b} - J_{a,b} + 1)}. \end{aligned} \tag{12}$$

Since (Q3) holds, the 2-core of $K_{a,b}$ has a total of $\mu_k(c)n + o(n)$ vertex-copies in K_a and $\mu_k(c)n + o(n)$ vertex-copies in K_b. $\mathcal{E}$ contains at most $2\alpha n$ edges, each using one vertex-copy on each side. So

$$X_{a,b} \ge \mu_k(c)n - 2\alpha n + o(n).$$

Therefore, if we choose α to be sufficiently small in terms of ζ, then by Lemma 6.4(b), we have $\frac{2L^a_{a,b}}{X_{a,b}-1}, \frac{2L^a_{a,b}}{X_{a,b}-1} \le$

$\frac{2\tau_k(c)n+o(n)}{\mu_k(c)n-2\alpha n+o(n)} < 1-\frac{\zeta}{2}$. It follows that for every $x>0$ we have:

$$\frac{2(L^a_{a,b}-x)}{X_{a,b}-2x-1}, \frac{2(L^b_{a,b}-x)}{X_{a,b}-2x-1} < 1-\frac{\zeta}{2}.$$

If $L^a_{a,b}=L^b_{a,b}$ then this would yield that the bound of (12) is at most

$$\left(\frac{4}{X_{a,b}-J_{a,b}-t_{a,b}}\right)^{t_{a_b}}(1-\frac{\zeta}{2})^{J_{a,b}}.$$

However, we must multiply by a corrective factor if $J^a_{a,b}\neq J^b_{a,b}$. Noting that $|J^a_{a,b}-J^b_{a,b}|<t_{a,b}$, and that $X_{a,b}-J_{a,b}>X_{a,b}-2\alpha n>\frac{1}{2}X_{a,b}$ for α sufficiently small, we find that the corrective factor is at most $2^{t_{a,b}}$. Similarly, we have $X_{a,b}-J_{a,b}-t_{a,b}>\frac{1}{2}\mu_k(c)n$, and the bound of (12) is at most:

$$\left(\frac{16}{\mu_k(c)n}\right)^{t_{a,b}}(1-\frac{\zeta}{2})^{J_{a,b}}.$$

We multiply this bound over all a,b. Then we multiply by the contribution from (11) for each Type B 2-path. We also multiply by the 2 choices for whether each P_i is Type A or Type B, and if it is Type B, the at most $k-1$ choices for its colours - a total of at most k choices for each path.

Setting $R=k\times\max(\frac{2}{\xi_k(c)-2\alpha},\frac{16}{\mu_k(c)n})$, and recalling that $t\leq a$, this yields:

$$\mathbf{Pr}(E_3\wedge E_1|E_2)\leq(1-\frac{\zeta}{2})^J R^a\left(\frac{1}{n}\right)^t,$$

as required. □

7.4 Cyclic Δ-sets

Having shown that w.h.p. there are no small complex Δ-sets, we now show that there are few small cyclic Δ-sets (i.e. flippable unicycles).

Lemma 7.23. *The expected value of the total number of vertices on all cyclic Δ-sets in $P_{n,p=c/n}$ is $O(1)$.*

Proof: Recall from Lemma 7.21 that the 2-core of a cyclic Δ-set $\sigma\Delta\sigma'$ is a cycle $u_1,...,u_r$ such that every u_i is not in the 2-core of $K_{\sigma(u_i),\sigma'(u_i)}$ and its parent is u_{i+1} (addition mod r). We refer to such a cycle as a *Type B cycle*, as it is the cyclic analogue of a Type B path.

We let X_r denote the number of Type B cycles of length r.

Set $r'=\min(r,\sqrt{n})$. We will restrict our analysis to the first r' vertices on the cycle. First we choose the colours $\sigma(u_1),...,\sigma(u_{r'})$; there are fewer than $k(k-1)^{r'-1}$ choices. Next we choose u_1; there are fewer than n choices. Then we proceed around the cycle: after choosing u_i, we expose its parent in $K_{\sigma(u_i),\sigma'(u_i)}$ and set that vertex to be u_{i+1}. The same argument as in the proofs of (9) and (10) shows that the probability that u_{i+1} is not in the 2-core of $K_{\sigma(u_{i+1}),\sigma'(u_{i+1})}$ is less than $\frac{1}{k-1}(1-\frac{\zeta}{2})$. Finally if $r=r'$ (i.e. if $r\leq\sqrt{n}$) we bound the probability that u_1 is the parent of u_r in $K_{\sigma(u_r),\sigma'(u_r)}$. The argument used in the proof of Lemma 7.22 for the edge (u_{j_i},w_i) yields a bound of $\frac{1+o(1)}{\lambda_k(c)n}$.

At this point, we can rule out all cycle lengths $r>\sqrt{n}$ as follows. The count in the preceding paragraph shows that the expected number of such Type B cycles is

$$\mathbf{Exp}(\sum_{r>\sqrt{n}}X_r)\leq\sum_{r=\sqrt{n}}^{n}nk(k-1)^{\sqrt{n}-1}[\frac{1}{k-1}(1-\frac{\zeta}{2})]^{\sqrt{n}-1}=o(n^{-2}). \tag{13}$$

So for the remainder, we take $r \le \sqrt{n}$. The expected total size of all such Type B cycles is:

$$\mathbf{Exp}(\sum_{r\le\sqrt{n}} rX_r) \le \sum_{r\le\sqrt{n}} nrk(k-1)^{r-1}[\frac{1}{k-1}(1-\frac{\zeta}{2})]^{r-1} \times \frac{1+o(1)}{\lambda_k(c)n} = \sum_{r\le\sqrt{n}} nr(1-\frac{\zeta}{2})^{r-1}O(\frac{1}{n}) = O(1). \quad (14)$$

Given a Type B cycle $u_1, ..., u_r$, we explore the remainder of the cyclic Δ-set $\sigma\Delta\sigma'$ using a branching process. Initially, the vertices $\{u_1, ..., u_r\}$ are labelled as *unexplored*. At each step, we choose an unexplored vertex w. We expose all vertices v not on the Type B cycle such that: (i) v is not in the 2-core of $K_{\sigma(v),\sigma(w)}$; (ii) w is the parent of v in $K_{\sigma(v),\sigma(w)}$. We label each such v as *unexplored* and we change the label of w to *explored*; we say that each such v is a *child* of w. Lemma 7.10 implies that this process will reach every member of $\sigma\Delta\sigma'$. (It may also reach some additional vertices if $\sigma\Delta\sigma'$ is not a maximal cyclic Δ-set.)

At any point, we let Ψ_e, Ψ_u denote the set of vertices labelled explored, unexplored, respectively; we let $\Psi = \Psi_e \cup \Psi_u$. Initially $|\Psi| = r \le \sqrt{n}$.

For the purposes of this analysis, we expose the 2-core of $K_{a,b}$ for all a, b. By Lemma 6.3, w.h.p. each K_a has size $\lambda_k(c)n + o(n)$ and contains $\xi_k(c)n + o(n)$ vertices from the 2-core of $K_{a,b}$. As we carry out the branching process, at each step we will have exposed the following:

EXPOSED:

- the 2-core of $K_{a,b}$ for all a, b;
- the edges already discovered to be in the unicycle, and the fact that each such edge coresponds to a child/parent in a $K_{a,b}$;
- for each $v \in \Psi_e$, the fact that no vertex outside of Ψ has v as a parent in any $K_{a,b}$.

Suppose that $w \in \Psi_u$ is the vertex we are about to explore. We will prove that as long as $|\Psi| = o(n)$ we have that, conditioning on EXPOSED:

$$\text{The expected number of children of } w \text{ is at most } 1 - \frac{\zeta}{2}, \quad (15)$$

where $\zeta = \zeta(c)$ comes from Lemma 6.4.

We know $w \in K_a$ where $a = \sigma(w)$. Consider any $b \ne a$ and consider any vertex $x \in K_b \backslash \Psi$ that is not in the 2-core of $K_{a,b}$. For any non-2-core vertex $u \in K_{a,b}$, we use $p(u)$ to denote the parent of u in $K_{a,b}$. We will show that in the pendant tree of $K_{a,b}$ containing x,

$$\mathbf{Pr}(p(x) = w) \le \frac{1+o(1)}{|K_b\backslash\Psi|}. \quad (16)$$

There are $k-1$ choices for b. If $|\Psi| = o(n)$ then there are $(\lambda_k(c) - \xi_k(c))n + o(n)$ choices for x and $|K_a\backslash\Psi| \ge \lambda_k(c)n + o(n)$. Applying Lemma 6.4(a), and taking α sufficiently small in terms of ζ, the expected number of children of w is at most

$$(k-1) \times \frac{\lambda_k(c) - \xi_k(c)}{\lambda_k(c)} + o(1) < 1 - \frac{\zeta}{2}.$$

This establishes (15).

To prove (16), consider any vertex $y \in K_a\backslash\Psi$. Recall that we have already exposed the vertices of the 2-core of $K_{a,b}$. If y is in the 2-core of $K_{a,b}$, then the same argument as in Case 2 of the proof of Lemma 7.22 proves that, conditional on the information of EXPOSED,

$$\mathbf{Pr}(p(x) = w) \le \mathbf{Pr}(p(x) = y) \quad (17)$$

The key observation for this argument is: Consider any graph where $p(x) = w$; replacing the edge xw with xy yields an equally probable graph with $p(x) = y$ (note that the reverse statement does not hold). It is important to note that replacing this edge cannot alter the information in EXPOSED.

Suppose y is not in the 2-core of $K_{a,b}$. Let E^* be the event that y is a descendant of x in a pendant tree of $K_{a,b}$.

Claim 1: $\mathbf{Pr}[(p(x) = w) \wedge E^*] = O(n^{-2})$.

Proof: We can expose the unique path from y to the 2-core of $K_{a,b}$ as follows: Set $z := y$. While z is not in the 2-core, set $z := p(z)$. Consider any z' in the 2-core and not in Ψ. The same argument from Case 2 of Lemma 7.22 that is used to prove (17) also implies $\mathbf{Pr}(p(z) = x) \le \mathbf{Pr}(p(z) = z')$. So at each step, since there are $\Theta(n)$ choices for z', the probability of reaching x is $O(1/n)$ and the probability of reaching the 2-core is $\Theta(1)$. It follows that the probability that we reach x before the 2-core, i.e. that x is on the path from y to the 2-core, is $O(1/n)$.

If we do reach x before the 2-core, then we next expose $p(x)$. Again using the same argument as in (17), we obtain that every 2-core vertex is at least as likely as y to be the parent of x and so the probability that $p(x) = y$ is $O(1/n)$. This proves the claim. □

Claim 2: $\mathbf{Pr}(p(x) = w|\overline{E^*}) \le \mathbf{Pr}(p(x) = y|\overline{E^*})$.

Proof: This is equivalent to showing that $\mathbf{Pr}[(p(x) = w) \wedge \overline{E^*}] \le \mathbf{Pr}[(p(x) = y) \wedge \overline{E^*}]$. Consider any graph H for which (i) $K_{a,b} = H$; (ii) setting $K_{a,b} = H$ does not contradict anything in EXPOSED; and (iii) the event $(p(x) = w) \wedge \overline{E^*}$ holds. Let H' be the graph obtained by replacing the edge (x, w) with (x, y). Since $\overline{E^*}$ holds for H, we have that H' is connected and y is the parent of x in H'. Furthermore, H and H' have the same 2-core and the same number of edges. Finally, note that replacing this edge does not alter anything in EXPOSED since $y \notin \Psi$ and $w \notin \Psi_e$. So H, H' are both equally likely to be $K_{a,b}$ as argued repeatedly previously, eg. in the proof of Lemma 6.1. Since each such H' can arise from at most one such H, the Claim follows. □

Therefore, $\mathbf{Pr}(p(x) = w) \le \mathbf{Pr}(p(x) = y) + \mathbf{Pr}[(p(x) = w) \wedge E^*] = \mathbf{Pr}(p(x) = y) + O(n^{-2})$. Summing over all $y \in K_a \backslash \Psi$ yields

$$|K_a \backslash \Psi| \times \mathbf{Pr}(p(x) = w) \le \sum_{y \in K_a \backslash \Psi} \mathbf{Pr}(p(x) = y) + O(n^{-2}) \le 1 + |K_a| \times O(n^{-2}) = 1 + o(1).$$

This yields (16) and thus completes the proof of (15).

Having proved (15), we now return to our branching process. Recall that when we explore $w \in \Psi_u$, we expose every child of w in each $K_{\sigma(w),b} \backslash \Psi$. We then place each of those children in Ψ_e and move w to Ψ_u. Note that EXPOSED still lists all the information that we have exposed, and so we can apply (15) at each step, so long as $|\Psi| = o(n)$.

Thus, at each step the expected change in $|\Psi_u|$ is at most $-\zeta/2$. Recall that initially $|\Psi_u| = r < \sqrt{n}$, the size of the Type B cycle. It follows easily that the expected number of steps until $|\Psi_u| = 0$ is $O(r)$. Note that the size of the cyclic Δ-set is r plus this number of steps and thus has expectation $O(r)$. By (14) the total size of the Type B cycles of length at most $\sqrt{n}$ is $O(1)$ and so the expected total size of cyclic Δ-sets arising from such Type B cycles is $O(1)$. By (14) the total size of cyclic Δ-sets arising from Type B cycles of length greater than $\sqrt{n}$ is less than $n \times o(n^{-2}) = o(1)$, since each such Δ-set has size at most n. This proves the lemma. □

8 The Kempe-core threshold

We adapt the argument from [54], where we analyzed a very similar core, but in a simpler setting. In fact, the main motivation for [54] was to develop a technique that we could use here to analyze the Kempe-core. The reader might prefer to read [54] before reading this section.

We let $G_{n_1,n_2,p}$ denote the random bipartite graph whose parts have size n_1, n_2 and where each of the $n_1 n_2$ possible edges is present independently with probability p. We will need the following bound on the size of the components of $G_{n_1,n_2,p=c/n}$.

Lemma 8.1. *Consider any constants $a_1, a_2, c > 0$ and any functions $n_1 = n_1(n) = a_1 n + o(n), n_2 = n_2(n) = a_2 n + o(n)$. There is a function $Q = Q(a_1, a_2, c)$ such that in $G_{n_1,n_2,p=c/n}$, with probability at least $1 - o(n^{-2})$:*

(a) If $a_1 a_2 c^2 > 1$ then there is a unique positive solution (α_1, α_2) to

$$\begin{aligned}\alpha_1 &= a_2(1 - e^{-\alpha_2 c})\\ \alpha_2 &= a_1(1 - e^{-\alpha_1 c}).\end{aligned}$$

Furthermore,

(i) the two parts of the largest component have sizes $\alpha_1 n + o(n), \alpha_2 n + o(n)$;

(ii) every other component has size at most $Q \log n$.

(b) If $a_1 a_2 c^2 < 1$ then every component has size at most $Q \log n$.

(c) If $a_1 a_2 c^2 = 1$ then only $o(n)$ vertices lie in components of size greater than $\log n$.

Remark 8.2. *It is easily checked in the proof that we can take $Q(a_1, a_2, c)$ to be a function that decreases as any of a_1, a_2, c increase whenever $a_1 a_2 c^2 > 1$.*

Parts (a,b) are essentially proven by Johansson in Chapter 2 of [42], but his statement is not quite as precise as we need here; he had an $O(\log^2 n)$ bound on the sizes of the small components in part (a ii), and he shows that this holds w.h.p. rather than with probability $1 - o(n^{-2})$. But the same approach with minor changes yields our stronger statement - we outline this proof.

We make use of this common version of the Chernoff Bound (see eg. [58]). $BIN(n, p)$ is the binomial random variable. For any $0 < t \leq np$ we have

$$\Pr\left[|BIN(n,p) - np| > t\right] < 2e^{-t^2/3np}. \tag{18}$$

Proof We modify the approach from section 5.2 of [41] where they prove the corresponding result for $G(n, p)$. Given a vertex v, we expose the component C_v containing v by a search, eg. a breadth-first search; this standard approach is commonly referred to as a *branching process*. The main point on which we (and [42]) differ from [41] is that when processing a vertex $u \in C_v$, where [41] exposes the undiscovered neighbours of u, we expose the undiscovered vertices of distance 1 and 2 from u. On step i, we let Y_i denote the number of neighbours and X_i denote the number of neighbours of those Y_i neighbours; each time we only count vertices which were not encountered previously during this search. Thus, in each step, the vertex u that we process is always on the same side of the bipartition as v. If the process dies out, i.e. if we discover all of C_v, then we pick a uniform new vertex on the same side as v and continue, until no vertices remain on that side. Thus X_i, Y_i may be defined for i much larger than $|C_v|$.

We use $C_1(v), C_2(v)$ to denote the vertices in C_v that are on the same side, opposite side resp. of the bipartition as v.

We begin with part (b). Suppose WLOG that v is on the side of the bipartition containing n_1 vertices. Note that Y_i is distributed as $BIN(n_2', p)$ where n_2' is the number of vertices on the opposite side of v that have not yet been discovered. So if we choose Y_i^+ with distribution $BIN(n_2, p)$ then Y_i^+ dominates Y_i from above; i.e. we can couple the two so that $Y_i \leq Y_i^+$ always. Having chosen Y_i^+, we choose X_i^+ with distribution $BIN(n_1 Y_i^+, p)$; thus X_i^+ dominates X_i from above.

Note, as in [41, 42] that if $|C_1(v)| \geq \ell$ then we must have $\sum_{i=1}^{\ell} X_i \geq \ell - 1$. Choose some constant $\epsilon = \epsilon(a_1, a_2, c) > 0$ such that $(1+\epsilon)^2 a_1 a_2 c^2 < 1$. Regrouping the binomial trials in a convenient manner and applying the Chernoff Bound, we see:

$$
\begin{aligned}
\mathbf{Pr}\left[|C_1(v)| \geq \ell\right] &\leq \mathbf{Pr}\left[\sum_{i=1}^{\ell} X_i^+ \geq \ell - 1\right] \\
&\leq \mathbf{Pr}\left[BIN(\ell n_2, p) > (1+\epsilon)\ell a_2 c\right] + \mathbf{Pr}\left[BIN(((1+\epsilon)\ell a_2 c) \times n_1, p) \geq \ell - 1\right] \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(the first term bounds } \sum_{i=1}^{\ell} Y_i^+\text{)} \\
&\leq \mathbf{Pr}\left[BIN(\ell n_2, p) > (1+\epsilon)\ell a_2 c\right] + \mathbf{Pr}\left[BIN(((1+\epsilon)\ell a_2 c) \times n_1, p) \geq (1+\epsilon)^2 \ell a_1 a_2 c^2 - 1\right] \\
&\leq e^{-q\ell} \qquad\qquad\qquad\qquad \text{for a constant } q = q(\epsilon, a_1, a_2, c) = q(a_1, a_2, c) \\
&\leq n^{-5} \qquad\qquad\qquad\qquad \text{for } \ell > \tfrac{1}{2} Q \log n \text{ where } Q = Q(a_1, a_2, c) > 10/q.
\end{aligned}
$$

Multiplying by n_1 proves that w.h.p. none of the n_1 vertices has $C_1(v) \geq \frac{1}{2} Q \log n$. The same calculation shows that w.h.p. none of the n_2 vertices has $C_1(v) \geq \frac{1}{2} Q \log n$. So w.h.p. no component has more than $\frac{1}{2} Q \log n$ vertices on either side; this proves (b).

For part (a), consider any $\ell_- < \ell_+$. We will first prove that every component has size at most ℓ_- or at least ℓ_+. As noted in [41, 42], if $\sum_{i=1}^{\ell} X_i \geq \ell - 1 + \frac{1}{2}(a_1 a_2 c^2 - 1)\ell$ for all $\ell_- \leq \ell \leq \ell_+$, then the branching process does not die out between steps $\ell_-, ..., \ell_+$ and so either $|C_1(v)| \leq \ell_-$ or $|C_1(v)| \geq \ell_+$. We will choose $\ell_-, \ell_+ = o(n)$. We will halt our process if the number of discovered vertices ever reaches ℓ_+; this will indicate that either $|C_v| \geq \ell_+$ or we are no longer exploring C_v. We bound Y_i from below by Y_i^- with distribution $BIN(n_2 - \ell_+, p)$ and we bound X_i from below by X_i^- with distribution $BIN((n_1 - \ell_+)Y_i^-, p)$. Thus, the probability that $\ell_- < |C_1(v)| < \ell_+$ is at most:

$$
\sum_{\ell=\ell_-}^{\ell^+} \mathbf{Pr}\left[\sum_{i=1}^{\ell} X_i^- \leq \ell - 1 + \frac{1}{2}(a_1 a_2 c^2 - 1)\ell\right].
$$

The same arguments as above, this time choosing ϵ such that $(1-\epsilon)^2 a_1 a_2 c^2 > 1 + \frac{1}{2}(a_1 a_2 c^2 - 1)$, shows that this holds with probabilty $\ell_+ e^{-\Theta(\ell_-)} = o(n^{-5})$ when $\ell_- = Q \log n$ for sufficiently large $Q = Q(a_1, a_2, c)$. To be specific, we choose $\ell_+ = n^{2/3}$. Multiplying by the $n_1 + n_2$ choices for v proves that w.h.p. there are no components of size between ℓ_- and ℓ_+.

The proof of Lemma 8 of [42] shows that w.h.p. there are no two components of size at least ℓ_+, and the failure probability is easily computed to be much less than n^{-3}. So there is at most one large component.

Theorem 9 of [42] shows that w.h.p. the size of that largest component is as stated in part (a i). He sets Y to be the number of vertices on components of size less than ℓ_- and applies the second moment method to show that the variable Y is concentrated. We can instead apply Talagrand's Inequality to obtain a failure probability of much less than $o(n^{-2})$. Specifically, we work with $Z = n - Y$ and note that (i) the choice of whether to include a particular edge can affect Z by at most $2k_-$ (the extreme case is when that edge joins two small components) and (ii) for any $s \geq 0$, if $Z \geq s$ then there is a set of at most $\max(s, \ell_-)$ edges which certify that $Z \geq s$, namely the edges in large pieces of large components containing a total of at least s vertices. These are enough to apply the version labelled "Talagrand's Inequality I" in Chapter 10 of [58]. This establishes that part (a) holds with probability $1 - o(n^{-2})$. Finally, we note that the calculation of $\mathbf{Exp}(Y)$ and argument described here works if we define Y to be the number of vertices on components of size less than ℓ' for any $\ell' \leq \ell_+$ so long as ℓ' grows with n; in fact, [42] uses a different value of ℓ_- than we do here. Repeating the argument with $\ell' = \log n$ shows that with probability $1 - o(n^{-2})$ we have $o(n)$ vertices on components of sizes between $\log n$ and ℓ^+. This will be useful in the final part of this proof.

We prove part (c) by the continuity of α_1, α_2, and the easily verified fact that they tend to 0 as $a_1 a_2 c^2 \to 1$ from above. Suppose $a_1 a_2 c^2 = 1$, and choose any $\zeta > 0$. It is straightforward to verify that we can choose some $\delta > 0$ so that upon increasing a_1 to $a_1 + \delta$, $\alpha_1 + \alpha_2 < \frac{1}{2}\zeta$. Applying part (a) and the final sentences from the previous paragraph show that with probability $1 - o(n^{-2})$ the number of vertices in components of size greater than $\log n$ of this bigger graph is less than ζn and hence the same is true of $G_{n_1,n_2,p}$. Since this is true for every $\zeta > 0$, this establishes part (c). $\square$

8.1 The stripping process

We will consider a parallel version of Kempe-Strip from Section 5. At each iteration, we remove *all* Kempe-chains of size at most $\log^2 n$.

Recall that we analyze the planted model $P_{n,p}$ (Definition 4.4). The vertex set is partitioned uniformly at random into k parts $A_1, ..., A_k$, and between each pair $1 \le a < b \le k$ we have a copy of $G_{|A_a|,|A_b|,p}$. Note that the Kempe-chains are the components of the bipartite subgraphs between the pairs of colour classes. So, by Lemma 8.1, the first iteration removes all but the giant component from each of these $\binom{k}{2}$ random bipartite graphs. In fact, we will see that the same is true of each subsequent iteration:

Observation 8.3. *This procedure is equivalent to repeatedly removing all small components (i.e. components of size less than* $\log^2 n$*) from the bipartite random graph induced by each pair of parts* A_a, A_b.

Remark: One should think of this as being equivalent to leaving only the giant component of each bipartite random graph. For some rare initial densities c, we will reach a point where the density of a remaining bipartite graph is within the giant component threshold, and so it will have many components of size greater than $\log^2 n$; in that case, all such components remain.

Note: In this and other procedures, we often use the superscript i to denote the value of a parameter at the beginning of iteration i, rather than the ith power of that parameter. For example:

At the beginning of iteration i, V_a^i will be the vertices remaining from part A_a, for each $1 \le a \le k$.

STRIP
Initialize $V_1^1 = A_1, ..., V_k^1 = A_k$.
For $i \ge 1$
for all $a \ne b$, $K_{a,b}^i$ is the vertex-set of the union of all components of size at least $\log^2 n$
in the bipartite subgraph induced by (V_a^i, V_b^i)
(so if no component is larger than $\log^2 n$ then $K_{a,b}^i = \emptyset$).
for every $1 \le a \le k$,
set $V_a^{i+1} = \cap_{b \ne a}(K_{a,b}^i \cap V_a)$.
if $V_a^{i+1} = V_a^i$ for all $1 \le a \le k$ then HALT and return $V_1^i \cup ... \cup V_k^i$.
if $V_1^{i+1} = ... = V_k^{i+1} = \emptyset$ then HALT and return $\emptyset$.

Note that when carrying out an iteration of STRIP, we can expose the vertex set of $K_{a,b}^i$ without actually exposing the edges within $K_{a,b}^i$. Thus, at iteration $i+1$, any pair of subsets of V_a^{i+1}, V_b^{i+1} with the appropriate size are equally likely to form the vertex sets of $K_{a,b}^{i+1}$. Note also that for any $(a, b) \ne (a', b')$, the edges of the bipartite graph induced by (A_a, A_b) are independent of the edges of the bipartite graph induced by $(A_{a'}, A_{b'})$. This yields:

Observation 8.4. *Given* V_a^i, V_b^i *and* $x_a = |K_{a,b}^i \cap V_a^i|, x_b = |K_{a,b}^i \cap V_b^i|$*, the vertices of* $K_{a,b}^i \cap V_a^i$ *and* $K_{a,b}^i \cap V_b^i$ *can be treated as uniformly random subsets of* V_a^i, V_b^i *of sizes* x_a, x_b*, respectively. Furthermore, if* $(a, b) \ne (a', b')$ *then the random subsets selected for* $K_{a,b}$ *are independent of the subsets chosen for* $K_{a',b'}$.

Proof This is very similar to the proof of Lemma 6.1. Run STRIP on any (G, σ) and let H be the graph induced by (V_a^i, V_b^i). Consider any permutation ϕ on $V(H)$ and replace every edge $(u, v) \in (V_a^i, V_b^i)$ with $(\phi(u), \phi(v))$; let H' be the resulting graph on $V(H)$, and G' be the resulting graph on $V(G)$. If we apply STRIP to (G', σ) then (i) the vertex sets V_a^i, V_b^i will be unchanged (ii) they will induce the subgraph H', and (iii) $K_{a',b'}^i$ will remain unchanged for every $(a', b') \ne (a, b)$. Furthermore, G, G' arise with the same probability in $P_{n,p=c/n}$, since they have the same number of edges. This implies that every H' obtained in this way is just as likely as H, even when conditioning on the subgraphs induced by every other $(V_{a'}^i, V_{b'}^i)$.

Noting that a uniformly random ϕ chooses two uniformly random subsets of V_a^i, V_b^i of sizes x_a, x_b, this implies the lemma. $\square$

We will focus mainly on V_1^i, V_2^i; by symmetry, the other sets V_a^i evolve in a similar manner. It will be useful to focus, in particular, on $K_{1,2}^i$; i.e. the giant component of the bipartite subgraph induced by the vertices of V_1^i, V_2^i. It will be convenient to consider sets U^i, W^i, where we will have $V_1^i \subseteq U^i \subseteq A_1$ and $V_2^i \subseteq W^i \subseteq A_2$. Initially, $U^i = A_1$, $W^i = A_2$; throughout the procedure, vertices are removed from U^i and W^i at the same rate that vertices which lie in small components of any bipartite subgraphs *except for* the one on (A_1, A_2) are removed from V_1^i and V_2^i. The sets U^i, W^i will be very close to uniformly chosen from A_1, A_2.

To form U^{i+1}, we select some (but not all) of the vertices to be removed from V_1^i as follows:

(1) Expose the number that should be removed because they are in small components of the subgraphs induced by $(V_1^i, V_b^i), 3 \le b \le k$.

(2) Select that many vertices uniformly at random from V_1^i and delete them; Observation 8.4 permits a coupling by which this is legal (see below). To carry out (2), we actually remove vertices uniformly from U^i until the appropriate number of vertices have been removed from V_1^i, and this leaves U^{i+1}.

Thus, the sequence of sets U^i decreases at the same rate that (1) contributes to the rate at which the sequence V_1^i decreases. Of course, V_1^i actually decreases more quickly since we also remove from it vertices that are in small components of the subgraph induced by (V_1, V_2).

We form W^{i+1} in the analogous manner, removing vertices from V_2^i.

More formally, U^i, W^i are defined by the following modified procedure, which captures STRIP from the viewpoint of the bipartite subgraph on (A_1, A_2).

Throughout this procedure, $K_{a,b}^i$ is defined as in STRIP; i.e. it is the vertex set of the union of all components in (V_a^i, V_b^i) of size at least $\log^2 n$. Typically, $K_{a,b}^i$ will be either the empty set or the giant component.

STRIP1
Initialize $V_1^1 = A_1, ..., V_k^1 = A_k$.
Initialize $U^1 = A_1, W^1 = A_2$.
For $i \ge 1$
    Expose the vertex-set of $K_{1,2}^i$.
    For every $3 \le b \le k$,
        Expose $\ell_b^i = |V_1^i \backslash K_{1,b}^i|$, the number of vertices removed from V_1^i
            because they are not in $K_{1,b}^i$.
        Pick a sequence of vertices chosen uniformly from U^i without replacement
                until ℓ_b^i of them are chosen from V_1^i.
            This sequence is L_b^i.
            Do not remove these vertices from U^i yet; they are still eligible
            to be chosen for another value of b.
        Set $K_{1,b}^i \cap V_1^i = V_1^i \backslash L_b^i$.
        Expose $q_b^i = |V_2^i \backslash K_{2,b}^i|$, the number of vertices removed from V_2^i
            because they are not in $K_{2,b}^i$.
        Pick a sequence of vertices chosen uniformly from W^i without replacement
                until q_b^i of them are chosen from V_2^i.
            This sequence is Q_b^i.
            Do not remove these vertices from W^i yet; they are still eligible
            to be chosen for another value of b.
        Set $K_{2,b}^i \cap V_2^i = V_2^i \backslash Q_b^i$.
    Set $U^{i+1} = U^i \backslash \cup_{3 \le b \le k} L_b^i$.
    Set $V_1^{i+1} = (K_{1,2}^i \cap V_1^i) \backslash \cup_{3 \le b \le k} L_b^i$.
    Set $W^{i+1} = W^i \backslash \cup_{3 \le b \le k} Q_b^i$.

Set $V_2^{i+1} = (K_{1,2}^i \cap V_2^i) \backslash \cup_{3 \le b \le k} Q_b^i$.
For every $3 \le a < b \le k$,
expose the vertex set of $K_{a,b}^i$.
Update $V_3^{i+1}, ..., V_k^{i+1}$ as in STRIP
if $V_a^{i+1} = V_a^i$ for all $1 \le a \le k$ then HALT and return $V_1^i, ..., V_k^i$.
if $V_1^{i+1} = ... = V_k^{i+1} = \emptyset$ then HALT and return $\emptyset, ..., \emptyset$.

It is important to note that, in each iteration, we only expose the vertex sets of each $K_{a,b}^i$, not the edges. Note further that, for each i, b, the ℓ_b^i vertices of L_b^i that are in V_1^i are uniform members of V_1^i. Therefore, $K_{1,b}^i \cap V_1^i$ is a uniformly random subset of V_1^i of size ℓ_b^i. Similarly, $K_{2,b}^i \cap V_2^i$ is a uniformly random subset of V_2^i of size q_b^i. So by Observation 8.4 and the fact that the graphs $K_{a,b}$ are chosen independently, we can couple STRIP1 with STRIP so that they produce the same sets $V_1^i, ..., V_k^i$. More precisely:

Run STRIP (as the master) and STRIP1 at the same time, both on the same (G, σ) drawn from $P_{n,p}$. When exposing the vertex set of some $K_{a,b}^i$, STRIP1 makes the same choice as STRIP. When choosing L_b^i for STRIP1, we choose each vertex from U^i in the following manner: We first determine whether it is in V_1^i (choosing it to be in V_1^i with probability equal to the number of remaining unchosen vertices in V_1^i divided by the number of remaining unchosen vertices in U^i). If it is not in V_1^i then we choose a uniform vertex from the unchosen vertices in $U^i \backslash V_1^i$. If it is in V_1^i, then we take the next vertex in a uniform permutation of what STRIP chose to be $V_1^i \cap K_{1,b}^i$. By Observation 8.4, this choice has the same distribution as a uniform member of V_1^i and hence this is a valid choice for STRIP1. Choose the vertices of Q_b^i in the analogous manner. Note that, under this coupling, STRIP and STRIP1 produce the same sets $K_{a,b}^i$ for every a, b, i and hence produce the same sets V_a^i, for every a, i.

Observation 8.5. *For each i, the set of components of size greater than $\log^2 n$ in (U^i, W^i) is identical to the set of components of size greater than $\log^2 n$ in (V_1^i, V_2^i).*

Proof This follows by noting that for each iteration i, and every $b \ge 3$, all vertices in the small components of (V_1^i, V_b^i) are removed from U^i. So all vertices in $U^i \backslash V_1^i$ must have been removed from V_1^i because they were in small components of (V_1^i, V_2^i) and hence are in small components of the subgraph induced by (U^i, W^i). More carefully:

Consider any $v \in U^i \backslash V_1^i$. Let $j \le i$ be the first index such that $v \notin V_1^j$; so $v \in V^{j-1} \backslash V^j$. We know that v cannot be in any L_b^{j-1} as otherwise v would be in $U^{j-1} \backslash U^j$ which would contradict $v \in U^i$. Therefore, $v \in V_1^{j-1} \backslash K_{1,2}^{j-1}$; i.e. v is in a component of size at most $\log^2 n$ in (V_1^{j-1}, V_2^{j-1}). So v has no neighbours in $K_{1,2}^{j-1}$, and hence has no neighbours in $V_2^i \subseteq K_{1,2}^{j-1}$. This establishes that there are no edges from $U^i \backslash V_1^i$ to V_2^i. The same argument holds for every value of i, and in particular there are no edges from $U^{j-1} \backslash V_1^{j-1}$ to V_2^{j-1}. Similarly, there are no edges from $W^{j-1} \backslash V_2^{j-1}$ to V_1^{j-1}. Therefore, the component of size at most $\log^2 n$ in (V_1^{j-1}, V_2^{j-1}) containing v is also a component of (U^{j-1}, W^{j-1}). Since $(U^i, W^i) \subseteq (U^{j-1}, W^{j-1})$, the component containing v in (U^i, W^i) has size at most $\log^2 n$. So every vertex in $U^i \backslash V_1^i$ and $W^i \backslash V_2^i$ lies in a component of (U^i, W^i) of size at most $\log^2 n$. And since there are no edges from $U^i \backslash V_1^i$ to V_2^i or from $W^i \backslash V_2^i$ to V_1^i, every component of (V_1^i, V_2^i) is also a component of (U^i, W^i). This proves the observation. □

This implies that $K_{1,2}^i$ is the set of components of size at least $\log^2 n$ in (U^i, W^i). This is very convenient, as (U^i, W^i) is much easier to analyze than (V_1^i, V_2^i). The reason is that U^i, W^i are uniform sets of $|U^i|, |W^i|$ vertices from A_1, A_2. At first glance, one might hope that this would mean (U^i, W^i) is distributed like $G_{n_1, n_2, p=c/n}$ where $n_1 = |U^i|, n_2 = |W^i|$. Unfortunately, this is not the case. One issue is that $|U^i|, |W^i|$ are determined in part by the number of vertices in small components of $(V_1^j, V_2^j), j < i$ and so there is dependency between the component sizes of (U^i, V^i) and the values of $|U^i|, |W^i|$. Nevertheless, concentration of $|U^i|, |W^i|$ will allow us to use Lemma 8.1 to bound $|K_{1,2}^i|$.

We define recursively:

$$\rho_1 = \nu_1 = \frac{1}{k}.$$

For $i \geq 1$, β_i is the largest solution to:

$$\beta_i = \rho_i(1 - e^{-\beta_i c}). \tag{19}$$

For $i \geq 2$:

$$\nu_{i+1} = \nu_i \left(\frac{\beta_i}{\nu_i}\right)^{k-1}, \qquad \text{for } i \geq 2; \tag{20}$$

$$\rho_{i+1} = \rho_i \left(\frac{\beta_i}{\nu_i}\right)^{k-2}, \qquad \text{for } i \geq 2. \tag{21}$$

Recall from Section 4 that

$$c_k = \min_{y>0} \frac{ky}{(1 - e^{-y})^{k-1}}.$$

If $c > c_k$ then let $\beta = \beta(c)$ be the greatest solution to $\beta = \frac{1}{k}(1 - e^{-\beta c})^{k-1}$ (setting $y = \beta c$ shows that a positive solution exists iff $c > c_k$). Set $\rho = \rho(c) = \frac{1}{k}\left(1 - e^{-\beta c}\right)^{k-2}$. Recalling the definition of $\lambda_k(c)$ from Section 4, note that:

$$\beta(c) = \lambda_k(c). \tag{22}$$

Lemma 8.6. *(a) If $c < c_k$ then there exists I such that $\beta_I = \nu_I = \rho_I = 0$.*

(b) If $c > c_k$ then $\lim_{i\to\infty} \beta_i = \lim_{i\to\infty} \nu_i = \beta$, $\lim_{i\to\infty} \rho_i = \rho > \frac{1}{c}$.

Proof Applying an easy induction to (20) and (21) yields $\nu_i = \frac{1}{k}\left(\prod_{j=1}^{i-1} \frac{\beta_j}{\nu_j}\right)^{k-1}$ and $\rho_i = \frac{1}{k}\left(\prod_{j=1}^{i-1} \frac{\beta_j}{\nu_j}\right)^{k-2}$. Therefore $(k\nu_i)^{k-2} = (k\rho_i)^{k-1}; \nu_i^{k-2} = k\rho_i^{k-1}$. Substituting that into (21) yields:

$$\rho_{i+1} = \rho_i \frac{\beta_i^{k-2}}{k\rho_i^{k-1}} = \frac{1}{k}\left(\frac{\beta_i}{\rho_i}\right)^{k-2} = \frac{1}{k}\left(1 - e^{-\beta_i c}\right)^{k-2}. \tag{23}$$

Let β', ν', ρ' be a set of fixed points of the recursive equations. (23) implies $\rho' = \frac{1}{k}\left(1 - e^{-\beta' c}\right)^{k-2}$ and (20) implies $\nu' = \beta'$, so (19) implies

$$\beta' = \rho'(1 - e^{-\beta' c}) = \frac{1}{k}(1 - e^{-\beta' c})^{k-2}(1 - e^{-\beta' c}) = \frac{1}{k}\left(1 - e^{-\beta' c}\right)^{k-1}. \tag{24}$$

As described above, (24) has a positive solution iff $c \geq c_k$. So for $c < c_k$, there is no positive fixed point β' and it follows that β_i, ρ_i, ν_i tend to zero. Therefore, there is some i such that $\rho_i < \frac{1}{k}$. Since $c < c_k = \min_{y>0} ky/(1 - e^{-y})^{k-1} < \min_{y>0} y/[\rho_i(1 - e^{-y})^{k-1}]$, there is no positive y satisfying $\frac{y}{c} = \rho_i(1 - e^{-y})^{k-1}$. Setting $y = c\beta_i$, there is no $\beta_i > 0$ satisfying (19) and so $\beta_i = 0$. Part (a) follows with $I = i + 1$.

For $c > c_k$, it is easy to check that $\rho_1 = \frac{1}{k} > \rho$. To see that $\rho_i > \rho$ for all i, let $b(x)$ be the largest solution to $b = x(1 - e^{-bc})$ and set $f(x) = \frac{1}{k}(1 - e^{-b(x)c})^{k-2}$. So $\beta_i = b(\rho_i)$ and (23) yields that $\rho_{i+1} = f(\rho_i)$. Clearly $b(x)$ increases with x and thus $f(x)$ increases with x. This implies that ρ_i is decreasing, and if $\rho_i > \rho$ then $\rho_{i+1} = f(\rho_i) > f(\rho) = \rho$; so $\lim_{i\to\infty} \rho_i$ exists and is at least ρ. Since β is the largest solution to (24), it follows that for every fixed point (β', ν', ρ'), we must have $\beta' \leq \beta, \nu' \leq \nu, \rho' \leq \rho$. Therefore $\lim_{i\to\infty} \rho_i = \rho$. This implies the other two limits.

Setting $y = c\beta$ yields $\rho = \beta/(1 - e^{-\beta c}) = \frac{1}{c} y/(1 - e^{-y}) > \frac{1}{c}$ since $e^{-y} > 1 - y$ for all $y > 0$. □

Lemma 8.7. *For any constant I, w.h.p. we have for each $1 \leq i \leq I$, $1 \leq a < b \leq k$:*

(a) $|K_{a,b}^i \cap V_a^i|, |K_{a,b}^i \cap V_b^i| = \beta_i n + o(n)$;

(b) $|U^i|, |W^i| = \rho_i n + o(n)$;

(c) $|V_a^i| = \nu_i n + o(n)$;

(d) *every vertex in* $V_a^i \backslash V_a^{i+1}$ *lies in a component of size less than* $\log^2 n$ *in the graph induced by* (V_a^i, V_b^i) *for at least one* $b \neq a$.

Furthermore:

(e) *If* $c < c_k$ *then there exists a constant* $I = I(c)$ *such that w.h.p.* $V_a^I = \emptyset$ *for all* $1 \leq a \leq k$*; i.e. no vertices remain after* I *iterations of STRIP.*

Remark: In our proof, we analyze STRIP1. By our coupling, the same bounds hold for the sets produced by STRIP.

Proof We will bound the probability that (a) holds for $a = 1, b = 2$, (c) holds for $a \in \{1, 2\}$, and that (d) holds for $a = 1$. By symmetry the same bounds hold for every a, b. We proceed by induction.

We define the events $A(i), B(i), C(i), D(i)$ to correspond to parts (a,b,c,d) of the lemma. Specifically, for some $h_i(n), h'_i(n) = o(n)$ that will be implicit below, our events are:

- $A(i)$: $|K_{1,2}^i \cap V_1^i|, |K_{1,2}^i \cap V_2^i| \in \beta_i n \pm h'_i(n)$;
- $B(i)$: $|U^i|, |W^i| \in \rho_i n \pm h_i(n)$;
- $C(i)$: $|V_1^i|, |V_2^i| \in \nu_i n \pm h_i(n)$;
- $D(i)$: every vertex in $V_1^i \backslash V_1^{i+1}$ lies in a component of size less than $\log^2 n$ in the graph induced by (V_1^i, V_b^i) for at least one $2 \leq b \leq k$.

We will prove that $\mathbf{Pr}(\overline{A(i) \wedge B(i) \wedge C(i) \wedge D(i)}) = o(1)$.

Base Case: It is easy to establish that $B(1), C(1)$ both hold w.h.p. Indeed, $U^1 = V_1^1 = A_1, W^1 = V_2^1 = A_2$, $\rho_1 = \nu_1 = \frac{1}{k}$ and $A_1, ..., A_k$ is a uniformly random partition of $\{1, ..., n\}$. The Chernoff Bound (18) yields that $\mathbf{Pr}(||A_1| - n/k| > h_1(n)) \leq e^{-\Theta(h_1(n)^2/n)}$. So we can take any $h_1(n) >> \log n$.

Inductive step: By Observation 8.5, we can analyze $K_{1,2}^i$ by analyzing the giant component of the subgraph induced by (U^i, W^i).

Choose uniformly random permutations $\mathcal{A}_1$ of A_1 and $\mathcal{A}_2$ of A_2. Let $U(x), W(y)$ be the sets consisting of the last x vertices in $\mathcal{A}_1$ and the last y vertices in $\mathcal{A}_2$. $U(x), W(y)$ are uniform subsets of sizes x, y chosen independently of the edges in the graph, and so the subgraph induced by $(U(x), W(y))$ is distributed exactly like $G_{x,y,p=c/n}$. Since the failure probability in Lemma 8.1 is $o(n^{-2})$ that lemma implies that w.h.p. for every $\rho_i n - h_i(n) \leq x, y \leq \rho_i n + h_i(n)$:

- if $\rho_i > \frac{1}{c}$ then $(U(x), W(y))$ has one component with $\beta_i n + o(n)$ vertices in each part, and every other component has size less than $Q \log n < \log^2 n$ for any $Q > Q(\beta_i, \beta_i, c)$; we let $h'_i(n) = o(n)$ be an upper bound on the $o(n)$ term over all x, y, which is determined by $h_i(n)$.
- if $\rho_i < \frac{1}{c}$ then every component of $(U(x), W(y))$ has size less than $Q \log n < \log^2 n$ for any $Q > Q(\beta_i, \beta_i, c)$
- if $\rho_i < \frac{1}{c}$ then only $o(n)$ vertices lie on components of $(U(x), W(y))$ of size greater than $\log n$.

Now, we use the permutations $\mathcal{A}_1, \mathcal{A}_2$ to run STRIP1 as follows. Each time we select a uniform vertex $v \in U^i$ to be placed in L_b^i we first choose whether v has already been selected for some $L_{b'}^i$ with $b' < b$. (Of course we make this choice with the appropriate probability; i.e. the number of vertices in $\cup_{b'<b} L_{b'}^i$ not yet selected for L_b^i divided by the number of vertices in U^i not yet selected for L_b^i.) If the choice is YES then we expose v; if it is NO then we choose v to be the next vertex in $\mathcal{A}_1$. We choose the sets Q_b^i in the same manner, this time taking the next vertex in $\mathcal{A}_2$. Thus, (U^i, W^i) is simply $(U(x), U(y))$ where $x = |U^i|, y = |W^i|$. So the bounds above say that if $B(i)$ holds then $A(i), D(i)$ hold:

Specifically, if $\rho_i > \frac{1}{c}$ then $K^i_{1,2}$ is the giant component of (U^i, W^i) and contains $\beta_i n + o(n)$ vertices on each side. If $\rho_i < \frac{1}{c}$ then $K^i_{1,2} = \emptyset$ and $\beta_i = 0$. If $\rho_i = \frac{1}{c}$ then all but $o(n)$ vertices are in components of size less than $\log n$, implying $A(i), D(i)$.

Next, we turn our attention to $B(i+1), C(i+1)$. We define $A^+(i)$ to be the event that for all $2 \le b \le k$, both parts of each $K^i_{1,b}$ have size in $\beta_i n \pm h'_i(n)$. By symmetry and since $k = O(1)$, $A^+(i)$ holds w.h.p.

If $A^+(i)$ holds then for each $3 \le b \le k$, $\ell^i_b \in |V^i_1| - \beta_i n \pm h'_i(n)$. It follows that for any $u \in U^i$, $\mathbf{Pr}(u \notin L^i_b) = \frac{\beta_i n}{|V^i_1|} + o(1)$ where the $o(1)$ term depends on $h'_i(n)$. Therefore:

$$\mathbf{Pr}(u \notin \cup_{b \ge 3} L^i_b) = \left(\frac{\beta_i n}{|V^i_1|}\right)^{k-2} + o(1) = \left(\frac{\beta_i}{\nu_i}\right)^{k-2} + o(1).$$

It follows that if $A^+(i)$ and $C(i)$ hold then

$$\mathbf{Exp}(|U^{i+1}|) \in (\rho_i n \pm h_i(n)) \times \left(\left(\frac{\beta_i}{\nu_i}\right)^{k-2} + o(1)\right) = \rho_{i+1} n + o(n),$$

$$\mathbf{Exp}(|V^i_1|) = |K^i_{1,2} \cap V^i_1| \times \left(\left(\frac{\beta_i}{\nu_i}\right)^{k-2} + o(1)\right) = \nu_i n \left(\left(\frac{\beta_i}{\nu_i}\right)^{k-2} + o(1)\right) = \nu_{i+1} n + o(n),$$

The $o(n)$ terms depend on $h_i(n), h'_i(n)$. Because these sets are determined by the choices of $\Theta(n)$ vertices from U_i, W_i, a straightforward concentration argument, such as one using Azuma's Inequality (see eg. [58]), shows that

$$\mathbf{Pr}(||U^{i+1}| - \rho_{i+1} n| > h_{i+1}(n)), \mathbf{Pr}(||V^{i+1}| - \nu_{i+1} n| > h_{i+1}(n)) = o(1),$$

for $h_{i+1}(n) = o(n)$ defined in terms of $h_i(n), h'_i(n)$. The same argument applies to $W_{i+1}, V^i_2, K^i_{1,2} \cap V^i_2$, thus yielding:

$$\mathbf{Pr}(\overline{B(i+1)} \wedge A(i) \wedge B(i) \wedge C(i)), \mathbf{Pr}(\overline{C(i+1)} \wedge A(i) \wedge B(i) \wedge C(i)) = o(1).$$

By induction, $\mathbf{Pr}(A(i) \wedge B(i) \wedge C(i)) = 1 - o(1)$. Therefore,

$$\mathbf{Pr}(B(i+1)), \mathbf{Pr}(C(i+1)) = 1 - o(1),$$

as required. Note that since $I = O(1)$, there are $O(1)$ events and so their union holds w.h.p.

Finally, we prove part (e). If $c < c_k$ then Lemma 8.6(a) implies that there is an i with $\rho_i = 0$, and so $|U^i|, |W^i| \le h_i(n) = o(n) < \frac{1}{2c} n$. This implies that (U^i, W^i) is contained in $(U(\frac{1}{2c}n), W(\frac{1}{2c}n))$ and Lemma 8.1 says that w.h.p. every component of $(U(\frac{1}{2c}n), W(\frac{1}{2c}n))$ has size less than $Q \log n < \log^2 n$ for $Q = Q(\rho_i) < Q(\frac{1}{2c}, \frac{1}{2c}, c)$. Therefore every such component will be removed in the next iteration and so part (e) holds with $I = i+1$. (In fact, one can show that it holds with $I = i$ unless we had $\rho_{i-1} = \frac{1}{c}$.) $\square$

Lemmas 8.6 and 8.7 imply that by running STRIP for a sufficiently large constant number of iterations, we can ensure that each V^I_a has size within an arbitrarily small multiplicative constant of βn. We close this section by bounding the rate of change in $|V^i_a|$ after it gets close enough to that limit.

Lemma 8.8. *For any $r, c > 0$ with $cr > 1$, let b be the largest solution to*

$$b = r(1 - e^{-bc}).$$

If $c > c_k$ then there exists $\zeta = \zeta(c) > 0$ such that at $r = \rho(c)$ we have:

$$\frac{\partial b}{\partial r} < \frac{(k-1)\beta(c)}{(k-2)\rho(c)} - 2\zeta.$$

Proof Define $f(y) = ky/(1-e^{-y})^{k-1}$ and recall that c_k is the minimum of $f(y)$ and that for $c \geq c_k$, $y(c)$ is the largest solution to $f(y) = c$. Recall also the definitions: $\beta = \beta(c)$ is the greatest solution to $\beta = \frac{1}{k}(1-e^{-\beta c})^{k-1}$; $\rho = \rho(c) = \frac{1}{k}\left(1-e^{-\beta c}\right)^{k-2}$. Note that $f(\beta c) = c$ and so $y(c) = \beta c$. It is straightforward to check that f is increasing for $y > y(c_k)$ and so for $c > c_k$ we have $y(c) > y(c_k)$ and $f'(y(c)) > 0$. Differentiating f and simplifying $f'(y) > 0$ yields that for $y = y(c) = \beta c$ we have:

$$1 - e^{-y} > (k-1)ye^{-y}. \tag{25}$$

Noting that $1 - e^{-\beta c} = \beta/\rho$, this yields $\beta/\rho > (k-1)\beta c e^{-\beta c}$ and so:

$$\frac{1}{k-1} > \rho c e^{-\beta c}. \tag{26}$$

Rearranging $b = r(1-e^{-bc})$, we obtain $r = b/(1-e^{-bc})$ and so

$$\frac{\partial b}{\partial r} = 1/\frac{\partial r}{\partial b} = \frac{(1-e^{-bc})^2}{(1-e^{-bc}) - bce^{-bc}}.$$

When $r = \rho$ we have $b = \beta$. Substituting $(1-e^{-\beta c}) = \beta/\rho$ and applying (26) yields that at $r = \rho$:

$$\frac{\partial b}{\partial r} = \frac{(\beta/\rho)^2}{(\beta/\rho) - \beta c e^{-\beta c}} = \frac{\beta/\rho}{1 - \rho c e^{-\beta c}} < \frac{\beta/\rho}{1 - \frac{1}{k-1}} = \frac{(k-1)\beta}{(k-2)\rho}.$$

Since $\frac{\partial b}{\partial r}$ at $r = \rho(c)$ is a function of c, this yields the lemma. $\square$

Corollary 8.9. *For any $c > c_k$ there exists $\zeta = \zeta(c), \xi = \xi(c) > 0$ such that for all $\rho(c) - \xi \leq r_1, r_2 \leq \rho(c) + \xi$:*

(a) there is a unique positive solution (b_1, b_2) to

$$\begin{aligned} b_1 &= r_1(1-e^{-b_2 c}) \\ b_2 &= r_2(1-e^{-b_1 c}) \end{aligned}$$

(b)

$$\frac{\partial(b_1+b_2)}{\partial r_1} < \frac{(k-1)\beta(c)}{(k-2)\rho(c)} - \zeta.$$

Proof Part (a) follows easily from a continuity argument (and see Lemma 8.1(a)). For part (b): Recall the equation relating b, r in Lemma 8.8 and note that, at $r_1 = r_2 = r$, changing r by δ is equivalent to changing both r_1 and r_2 by δ. It follows that at $r_1 = r_2 = \rho(c)$, we have

$$\frac{\partial b_1}{\partial r_1} + \frac{\partial b_1}{\partial r_2} = \frac{\partial b}{\partial r}.$$

By symmetry, $\frac{\partial b_1}{\partial r_2} = \frac{\partial b_2}{\partial r_1}$ at $r_1 = r_2$. Applying Lemma 8.8, at $r_1 = r_2 = \rho(c)$ we have:

$$\frac{\partial b_1}{\partial r_1} + \frac{\partial b_2}{\partial r_1} = \frac{\partial b}{\partial r} < \frac{(k-1)\beta}{(k-2)\rho} - 2\zeta.$$

Part (b) now follows from part(a) and the continuity of $\frac{\partial(b_1+b_2)}{\partial r_1}$. $\square$

This leads to the following lemma, which measures the change in $K^i_{a,b}$ as we delete more vertices from V^i_a, V^i_b.

Lemma 8.10. *For any $c > c_k$ there exists $\zeta = \zeta(c), \xi = \xi(c) > 0, I^+ = I^+(c), Q = Q(\zeta, c) > 0$ such that for any $I \geq I^+$, w.h.p. the subgraph remaining after I iterations is such that for every $1 \leq a < b \leq k$:*

Choose a sequence of ξn uniformly random vertices from V_a^I and a sequence of ξn uniformly random vertices from V_b^I. W.h.p. for every $t_a \leq \xi n$ and $t_b \leq \xi n$, in the subgraph remaining after removing the first t_a vertices of the first sequence and the first t_b vertices of the second sequence from (V_a^I, V_b^I):

(a) the largest component has size at least $|V_a^I \cup V_b^I| - (t_a + t_b)(1 + \frac{1}{k-2} - \zeta) + o(n)$; and

(b) the second largest component has size less than $Q \log n$.

Proof We let $\rho = \rho(c), \beta = \beta(c)$. We will take ζ, ξ from Corollary 8.9, but rescale them. Specifically, let ζ', ξ' be the values from Corollary 8.9 and set $\zeta = \frac{1}{2}\zeta', \xi = \frac{\beta}{4\rho}\xi'$.

We will focus on $a = 1, b = 2$. The result extends to every pair a, b by symmetry and since $k = O(1)$. Fix any t_1, t_2. Removing the first t_1, t_2 vertices from the sequences is equivalent to removing t_1, t_2 uniform vertices from V_1^I, V_2^I.

The key to our analysis is to modify the random vertex-removal. Rather than removing t_1, t_2 vertices from V_1^I, V_2^I, we will remove $s_1 = t_1 \frac{|U^I|}{|V_1^I|}, s_2 = t_2 \frac{|W^I|}{|V_2^I|}$ uniformly random vertices from U^I, W^I. We will argue that it suffices to analyze the modified experiment. Let T_1, T_2 denote the number of vertices that are removed from V_1, V_2 resp. in the modified experiment; note that the original experiment is simply the modfied experiment conditional on the event that $T_1 = t_1 \wedge T_2 = t_2$. Let E^* be the event that the largest component in the subgraph induced by the remaining vertices in (V_a^I, V_b^I) has size at least $|V_a^I \cup V_b^I| - (t_a + t_b)(1 + \frac{1}{k-2} - \zeta)$ and that all others have size less than $Q \log n$. Noting that $\mathbf{Exp}(T_i) = t_i$ and examining the binomial distribution, it follows easily that $\mathbf{Pr}(T_1 \geq t_1 \wedge T_2 \geq t_2) \geq \epsilon$ for some constant $\epsilon > 0$. Therefore, if $\mathbf{Pr}(E^*) = 1 - o(1)$ then $\mathbf{Pr}(E^* | T_1 \geq t_1 \wedge T_2 \geq t_2) = 1 - o(1)$. The probability of E^* conditioned on the values of T_1, T_2 clearly decreases as T_1, T_2 increase, and so $\mathbf{Pr}(E^* | T_1 = t_1 \wedge T_2 = t_2) \geq \mathbf{Pr}(E^* | T_1 \geq t_1 \wedge T_2 \geq t_2)$. Therefore, if E^* occurs w.h.p. in the modified experiment then it occurs w.h.p. in the original experiment.

By Lemma 8.7, w.h.p. $|U^I|, |W^I| = \rho_I n + o(n), |V_1^I|, |V_2^I| = \nu_I n + o(n)$. So we are removing $s_1 = t_1(\frac{\rho_I}{\nu_I} + o(1)) < \frac{1}{3}\xi' n$ (since $t_1 < \xi n$) uniform vertices from U^I and $s_2 = t_2(\frac{\rho_I}{\nu_I} + o(1)) < \frac{1}{3}\xi' n$ uniform vertices from W^I.

Recall that U^I, W^I are each formed by the removal of a sequence of uniformly chosen vertices from V_1, V_2 resp. In the proof of Lemma 8.7 we defined $U(x), W(y)$ to be uniform subsets of V_1, V_2 of sizes x, y, resp. and so $(U^I, W^I) = (U(x), W(y))$ for some $x, y = \rho_I n + o(n)$. Furthermore, deleting s_1, s_2 vertices from (U^I, W^I) will leave $(U(x'), W(y'))$ for some $x' = \rho_I n - s_1 + o(n), y' = \rho_I n - s_2 + o(n)$.

We choose I^+ sufficiently large that for all $I \geq I^+$ we have $\rho_I - \rho < \frac{1}{3}\xi'$ and $\frac{\rho_I}{\nu_I} < \frac{\rho}{\nu}(1 + \zeta'/10)$ (by Lemma 8.6).

Applying $\xi = \frac{\beta}{4\rho}\xi'$, this implies that the intervals $[|U^I| - s_1, |U^I|], [|W^I| - s_2, |W^I|]$ lie entirely within $[(\rho - \xi')n, (\rho + \xi')n]$. Therefore, Corollary 8.9 implies that the expression from Lemma 8.1 for the size of the giant component in $(U(x'), W(y'))$ is at least

$$
\begin{aligned}
2\beta_I n - (s_1 + s_2)\left(\frac{(k-1)\beta}{(k-2)\rho} - \zeta'\right) &> 2\beta_I n - (t_1 + t_2)\left(\frac{\rho_I}{\nu_I} + o(1)\right)\left(\frac{(k-1)\beta}{(k-2)\rho} - \zeta'\right) \\
&> 2\beta_I n - (t_1 + t_2)\left(\frac{(k-1)}{(k-2)}(1 + \frac{\zeta'}{10}) - \zeta'\right) n \\
&\qquad \text{since } \frac{\rho_I}{\nu_I} < \frac{\rho}{\beta}\left(1 + \frac{\zeta'}{10}\right) \text{ and } \frac{\rho_I}{\nu_I} > 1 \\
&> 2\beta_I n - (t_1 + t_2)\left(1 + \frac{1}{k-2} - \zeta\right), \qquad \text{since } \zeta = \frac{1}{2}\zeta'.
\end{aligned}
$$

So Lemma 8.1 implies that the probability of the remaining component being too small is at most $1 - o(n^{-2})$. It is implicit in Corollary 8.9 that $a_1 a_2 c^2$ is bounded away from 1 for all a_1, a_2 in the interval $[(\rho - \xi')n, (\rho + \xi')n]$; using this, it is easy to see that there is a $Q = Q(\zeta, c)$ which is greater than $Q(a_1, a_2, c)$ from Lemma 8.1

for all a_1, a_2 in that interval and so the probability that the second largest component is greater than $Q \log n$ is $o(n^{-2})$.

Multiplying by the fewer than n^2 choices for x', y' shows that w.h.p. the largest component of $(U(x'), W(y'))$ is as large as required and the other components are as small as required for every $x' = \rho_I n - s_1 + o(n), y' = \rho_I n - s_2 + o(n)$ with $s_1, s_2 < \frac{1}{3}\xi' n$. This is in the modified experiment described above. It follows that w.h.p. in the original experiment the largest component is as large as required and the other components are as small as required for every $t_1 \leq \xi n$ and $t_2 \leq \xi n$. $\square$

8.2 Termination

In this subsection, we prove Lemma 5.2.

Lemma 5.2 *For $k \geq 3$:*

(a) If $c < c_k$ then w.h.p. the Kempe-core of $P_{n,p=c/n}$ is empty.

(b) If $c > c_k$ then w.h.p. the Kempe-core of $P_{n,p=c/n}$ has size $k\lambda_k(c) + o(n)$.

We start by noting how the case $c < c_k$ follows immediately from the previous subsection:

Proof of Lemma 5.2(a) If $c < c_k$ then by Lemma 8.7(e), there exists a constant $I = I(c)$ such that $V_a^I = \emptyset$ for all $1 \leq a \leq k$; thus the Kempe-core is empty. $\square$

The remainder of this subsection is devoted to proving Lemma 5.2(b). So we assume $c > c_k$ throughout.

We adapt the proof of Lemma 5.1 in [54], showing that with sufficiently high probability, STRIP will terminate and return sets $V_1, ..., V_k$ where each V_a has size $\beta n + o(n)$. These are the sets $K_1, ..., K_k$ of the Kempe core.

Given $c > c_k$, we let $\beta = \beta(c)$ and $\rho = \rho(c)$, as defined above Lemma 8.6.

Intuition: Run STRIP until the start of the I^{th} iteration, for a sufficiently large constant I. We will show that it halts shortly thereafter. For each $i \geq I$, consider a vertex u in some V_a^i which is to be removed in iteration i. To be specific, we choose $u \in V_1^i \backslash K_{1,2}^i$; i.e. u is a vertex that is in a small component of the subgraph induced by (V_1^i, V_2^i). We will argue that the removal of u produces, in expectation, fewer than 1 vertices to be removed in iteration $i + 1$. Thus the number of vertices removed in each iteration has a negative drift and so, with high probability, will quickly drift to zero, at which point STRIP halts.

Removing u causes w to be removed in the next iteration iff it causes w to leave the giant component of some (V_1, V_b); i.e. u is a cutvertex in the giant component separating w from the bulk of the giant component. Note that this cannot occur for $b = 2$ as u is not in the giant component of (V_1, V_2). For each bipartite graph induced by $(V_1, V_b), b > 2$, u can be treated as a uniformly random member of V_1 by Observation 8.4. When we delete a uniformly random vertex u from V_1^i, the same reasoning used to prove Lemma 8.10 implies that we expect the size of the giant component on (V_1, V_b) to decrease by at most $1 + \frac{1}{k-2} - \zeta$. We are interested in the number of vertices *other than* u which leave the giant component, as these are the vertices that will be deleted from (V_1, V_b) during the next iteration of STRIP. Now for I sufficiently large, the probability that our uniform vertex u is in that giant component is very close to 1; specifically, it is greater than $1 - \frac{1}{2}\zeta$. Thus, since we expect at most $1 + \frac{1}{k-2} - \zeta$ vertices to leave the giant component, the expected number of those vertices that are not u is at most $1 + \frac{1}{k-2} - \zeta - (1 - \frac{1}{2}\zeta) = \frac{1}{k-2} - \frac{1}{2}\zeta$. In other words, the removal of u results in an expection of at most $\frac{1}{k-2} - \frac{1}{2}\zeta$ vertices to be deleted in the next iteration of STRIP because they move into small components of (V_1, V_b). There are $k - 2$ choices for $b > 2$ and so the removal of u results in an expectation of at most $1 - \frac{k-2}{2}\zeta$ new vertices to be deleted in the next iteration. Since this expectation is bounded by a constant less than one, we expect STRIP to halt very soon (as argued in the previous paragraph).

There are a number of ways to formalize this intuition into a proof. The following is, at heart, the same approach used in [54], but is phrased somewhat differently.

Lemma 8.11. *For any $\delta > 0$ there exists I sufficiently large that: w.h.p. STRIP terminates before δn additional vertices are removed from $\cup_{a=1}^{k} V_a^I$.*

This will prove the remainder of Lemma 5.2 as follows:

Proof of Lemma 5.2(b): We choose any $\delta > 0$, and I large enough for Lemma 8.11 to hold and (applying Lemma 8.6(b)) to satisfy

$$\nu_I - \beta < \frac{1}{2}\delta. \tag{27}$$

By Lemma 8.7(c), at the beginning of iteration I, w.h.p. we have for all a: $|V_a^I| = \nu_I n + o(n)$ which is between βn and $(\beta + \delta)n$ by (27). By Lemma 8.11, w.h.p. STRIP terminates before removing more than δn additional vertices. So w.h.p. STRIP halts and produces a Kempe core where each part contains between $(\beta - \delta)n$ and $(\beta + \delta)n$ vertices. Since this holds for every sufficiently small constant $\delta > 0$, there exists $h(n) = o(n)$ such that w.h.p. STRIP produces a Kempe core where each of the k parts contains $\beta n \pm h(n)$ vertices. This proves the lemma, after recalling from (22) that $\beta = \lambda_k(c)$. $\square$

Proof of Lemma 8.11: We can assume δ is small enough to satisfy

$$\delta < \frac{\xi\beta\zeta}{20k}, \tag{28}$$

where $\zeta, \xi = \zeta(c), \xi(c)$ are from Lemma 8.10. We take I to be large enough for Lemma 8.10 to hold.

Suppose that STRIP continues to an iteration $I^* \geq I$ such that at least δn vertices have been removed during iterations $I, ..., I^*$; i.e.

$$\sum_{a=1}^{k} |V_a^I \backslash V_a^{I^*+1}| \geq \delta n. \tag{29}$$

We can assume that I^* is the first such iteration, and so

$$\sum_{a=1}^{k} |V_a^I \backslash V_a^{I^*}| < \delta n. \tag{30}$$

Note: I^* may grow with n.

For each pair $1 \leq a, b \leq k, a \neq b$ we let $\ell_{a,b}$ denote the total number of vertices removed from V_a during iterations $I, ..., I^*$ because they were in small components of the graph on (V_a, V_b). More formally,

$$\ell_{a,b} = \sum_{i=I}^{I^*} |V_a^i \backslash K_{a,b}^i|.$$

Note that we allow both $a < b$ and $a > b$, so we use $K_{a,b}^i$ and $K_{b,a}^i$ to denote the same subgraph. Every $u \in V_a^I \backslash V_a^{I^*+1}$ is in $V_a^i \backslash K_{a,b}^i$ for some $I \leq i \leq I^*$ and at least one $b \neq a$. So by our choice of I^*, we have:

$$\sum_{a \neq b} \ell_{a,b} \geq \sum_{a=1}^{k} |V_a^I \backslash V_a^{I^*+1}| \geq \delta n. \tag{31}$$

We let $r_{a,b}$ denote the total number of vertices that are removed from V_a during iterations $I, ..., I^*$ because they were in a small component of the graph on $(V_a, V_{b'})$ for some $b' \neq b$. We let $r_{a,b}^-$ denote those vertices that were removed duing iterations $I, ..., I^* - 1$. More formally,

$$\begin{aligned} r_{a,b} &= \sum_{i=I}^{I^*} | \cup_{b' \neq b} V_a^i \backslash K_{a,b'}^i | \leq \sum_{b' \notin \{a,b\}} \ell_{a,b'} \qquad (32) \\ r_{a,b}^- &= \sum_{i=I}^{I^*-1} | \cup_{b' \notin \{a,b\}} V_a^i \backslash K_{a,b'}^i | \end{aligned}$$

Recalling the **Intuition** described above: $\ell_{a,b}$ (essentially) counts the removable vertices produced by $K_{a,b}$ as a result of deleting from V_a the $r_{a,b}$ (actually, $r^-_{a,b}$) vertices that arise from other bipartite graphs $K_{a,b'}$. The intuition was that the expected number of such removable vertices produced per deletion is less than $\frac{1}{k-2}$. We formalize this in Lemma 8.13 below when we show it is, on average, less than roughly $\frac{1}{k-2} - \frac{1}{4}\zeta$. This will lead to a contradiction.

During iterations $I, ..., I^*$, V_a^I loses $\ell_{a,b}$ vertices because they are in small components on (V_a, V_b) and during iterations $I, ..., I^* - 1$ it loses $r^-_{a,b}$ vertices because they are in small components on the other bipartite graphs; we will show that the overlap between these two groups of vertices is small, thus obtaining:

Lemma 8.12. *W.h.p. for all a, b:*

$$|V_a^I \backslash K_{a,b}^{I^*}| \geq \ell_{a,b} + r^-_{a,b} - \frac{4\delta^2}{\beta} n.$$

Proof Each of the vertices counted by $r^-_{a,b}$ is removed from V_a during iterations $I, ..., I^* - 1$. Each of the vertices counted by $\ell_{a,b}$ is either removed from V_a during iterations $I, ..., I^* - 1$, or is in $V_a^{I^*} \backslash K_{a,b}^{I^*}$ (and so will be removed during iteration I^*). Let X denote the number of vertices that are counted by both $r^-_{a,b}$ and $\ell_{a,b}$. Thus:

$$|V_a^I \backslash K_{a,b}^{I^*}| \geq \ell_{a,b} + r^-_{a,b} - X.$$

We will prove the lemma by bounding X. Every vertex u counted by X must appear in both $V_a^i \backslash K_{a,b}^i$ and $V_a^i \backslash K_{a,b'}^i$ for some $b' \neq b$ and $I \leq i \leq I^* - 1$. By Observation 8.4, from the perspective of $K_{a,b}^i$, the vertex set of each $V_a^i \backslash K_{a,b'}^i$ is a uniform subset of $|V_a^i \backslash K_{a,b'}^i|$ vertices from V_a^i. So we can view the random process as follows:

For each $I \leq i \leq I^* - 1$, we: (i) Expose the vertices of $K_{a,b}^i$. (ii) Expose Y_i, the total number of vertices lying in $\cup_{b' \neq b}(V_a^i \backslash K_{a,b'}^i)$. (iii) Choose Y_i uniformly random members of V_a^i. (iv) Increase X by the number of those Y_i vertices that are not in $K_{a,b}^i$.

All vertices counted by Y_i are in $V_a^i \backslash V_a^{i+1}$. So (30) implies $\sum_{i=I}^{I^*-1} Y_i < \delta n$. For each $I \leq i \leq I^*$: Lemma 8.7 and (30) imply $|V_a^i| > |V_a^I| - \delta n = \beta_I n + o(n) - \delta n > \frac{1}{2}\beta n$ as $\delta < \frac{\beta}{2}$ by (28). Note that (30) also implies that the total number of vertices in $V_a^i \backslash K_{a,b}^i$ is less than δn. It follows that X is dominated by the binomial variable $BIN(\delta n, \frac{\delta n}{\frac{1}{2}\beta n})$. Thus

$$\mathbf{Exp}(X) \leq \delta n \times \frac{\delta n}{\frac{1}{2}\beta n},$$

and the Chernoff Bound (18) implies $\mathbf{Pr}(X > 2\mathbf{Exp}(X)) < e^{-\Theta(n)}$. This yields the lemma. □

We will use Lemma 8.10 to bound the size of $K_{a,b}^{I^*}$, the giant component in the subgraph induced by (V_a, V_b) in what remains after $I^* - 1$ iterations, in terms of the number of vertices removed. This will yield the following bound, which we will prove contradicts (32):

Lemma 8.13. *W.h.p. for every $1 \leq a, b \leq k$ we have*

$$\ell_{a,b} + \ell_{b,a} \leq (r^-_{a,b} + r^-_{b,a}) \left(\frac{1}{k-2} - \frac{1}{2}\zeta \right) + \frac{10\delta^2}{\beta} n.$$

Proof By Observation 8.4, from the perspective of $K_{a,b}$, the vertices removed from V_a because they are in small components of some other $K_{a,b'}$ can be viewed as uniformly random vertices from V_a. Specifically, letting $r_{a,b}^i = |\cup_{b' \neq b} V_a^i \backslash K_{a,b'}^i|$, we can view those $r_{a,b}^i$ vertices as being uniformly chosen from V_a^i.

We will use Lemma 8.10 to bound the size of the giant component in what remains of $K_{a,b}$ after removing those $r_{a,b}^i$ uniform vertices from every V_a^i as well as the corresponding $r_{b,a}^i$ vertices from every V_b^i, for $i = I, ..., I^* - 1$. We cannot apply Lemma 8.10 directly as it addresses the removal of uniform vertices from

V_a^I rather than V_a^i. Intuitively, this should make a negligible difference as each V_a^i differs from V_a^I by at most $|V_a^I \backslash V_a^{I^*-1}| < \delta n$ vertices. To formalize this intuition, we choose the $r_{a,b}^i$ vertices from V_a^i by choosing a possibly larger number of vertices from V_a^I, just as in STRIP1 where we chose vertices from V_1^i by choosing vertices from U^i. Specifically:

For each a, b, i, we choose $r_{a,b}^i$ vertices for deletion from V_a^i by repeatedly choosing vertices from V_a^I until $r_{a,b}^i$ of them are from V_a^i. Let $q_{a,b}$ denote the total number that we choose from V_a^I during iterations $I, ..., I^* - 1$. We will show that $q_{a,b}$ is not much larger than $r_{a,b}^-$.

By (30), fewer than δn vertices are deleted during iterations $I, ..., I^* - 1$. Therefore, for each $I \le i \le I^*$, we have $|V_a^i| > |V_a^I| - \delta n > |V_a^I|(1 - \delta/\beta)$ (since $|V_a^I| = \beta_I n + o(n) > \beta n$). Therefore, each time that we choose a random member of V_a^I, the probability that it is in V_a^i is at least $(1 - \delta/\beta)$. A straightforward application of, eg. standard tail bounds on binomial variables, implies that w.h.p. for every a, b we have

$$q_{a,b} < \frac{r_{a,b}^-}{1 - \delta/\beta} + \frac{1}{8}\delta^2 n. \tag{33}$$

In Observation 8.5, we argued that $K_{1,2}^i$ is the giant component of the subgraph induced by (U^i, W^i). For the same reasons, $K_{a,b}^{I^*}$ is the giant component of what remains in (V_a, V_b) after we remove the $r_{a,b}^-$ vertices of $\cup_{i=I}^{I^*-1} \cup_{b' \neq b} V_a^i \backslash K_{a,b'}^i$ from V_a^I and the $r_{b,a}^-$ vertices of $\cup_{i=I}^{I^*-1} \cup_{b' \neq b} V_a^i \backslash K_{a,b'}^i$ from V_b^I. To see this, note that each vertex of $V_a^I \backslash V_a^{I^*}$ is removed from V_a^I for one of two reasons: (1) it belongs to a small component of the graph induced by (V_a^i, V_b^i) for some $I \le i \le I^* - 1$; (2) it belongs to a small component of the graph induced by $(V_a^i, V_{b'}^i)$ for some $I \le i \le I^* - 1$ and $b' \neq a, b$. Vertices removed for reason (2) are counted by $r_{a,b}^-$. Vertices removed for reason (1) cannot affect any linear-sized component of the graph induced by (V_a^j, V_b^j) for any $j > i$. The same reasoning applies to vertices removed from V_b^I. (This argument is presented more formally in the proof of Observation 8.5.)

Removing those $r_{a,b}^-, r_{b,a}^-$ vertices from V_a^I, V_b^I cannot decrease the size of the largest component in $(V_a^{I^*}, V_b^{I^*})$ by more than removing the $q_{a,b}$ vertices from V_a^I, V_b^I decreases it, since the former set of vertices is a subset of the latter. (Indeed, a bit of thought shows that the removal of each set of vertices results in the *same* largest component, but this fact is not needed.) By (33), our choice of I^* and the fact that (28) implies $\delta < \frac{\zeta}{2}, \frac{\beta}{4}$, we have

$$q_{a,b} < \frac{r_{a,b}^-}{1 - \delta/\beta} + \frac{1}{8}\delta^2 n < \frac{\delta n}{3/4} + \frac{1}{8}\delta^2 n < 2\delta n < \xi n,$$

and similarly $q_{b,a} < \xi n$. This allows us to apply Lemma 8.10 with $t_a = q_{a,b}$ and $t_b = q_{b,a}$ to bound the size of the largest component of what remains after removing $q_{a,b}, q_{b,a}$ uniform vertices from V_a^I, V_b^I, which we have just argued to be no larger than $K_{a,b}^{I^*}$. This and (33) yield that w.h.p. for every a, b we have

$$\begin{aligned}
|K_{a,b}^{I^*}| &\ge |V_a^I \cup V_b^I| - (q_{a,b} + q_{b,a})(1 + \frac{1}{k-2} - \zeta) + o(n) \\
&> |V_a^I \cup V_b^I| - \left(\frac{r_{a,b}^- + r_{b,a}^-}{1 - \delta/\beta} + \frac{1}{4}\delta^2 n\right)\left(1 + \frac{1}{k-2} - \zeta\right) + o(n) \\
&> |V_a^I \cup V_b^I| - \frac{r_{a,b}^- + r_{b,a}^-}{1 - \delta/\beta}\left(1 + \frac{1}{k-2} - \zeta\right) - \frac{1}{2}\delta^2 n \\
&\ge |V_a^I \cup V_b^I| - (r_{a,b}^- + r_{b,a}^-)(1 + \frac{1}{k-2})\frac{1 - \zeta/2}{1 - \delta/\beta} - \frac{1}{2}\delta^2 n \\
&> |V_a^I \cup V_b^I| - (r_{a,b}^- + r_{b,a}^-)(1 + \frac{1}{k-2} - \frac{1}{2}\zeta) - \frac{1}{2}\delta^2 n \qquad \text{since } \frac{\delta}{\beta} < \frac{\zeta}{20k} \text{by (28).}
\end{aligned}$$

Lemma 8.12 yields that w.h.p. for every a, b we have

$$|V_a^I \cup V_b^I| - |K_{a,b}^{I^*}| \ge \ell_{a,b} + \ell_{b,a} + r_{a,b}^- + r_{b,a}^- - \frac{8\delta^2}{\beta} n.$$

Combining the two preceding inequalities yields that w.h.p.

$$\begin{aligned}\ell_{a,b}+\ell_{b,a}+r^-_{a,b}+r^-_{b,a}-\frac{8\delta^2}{\beta}n &\le (r^-_{a,b}+r^-_{b,a})(1+\frac{1}{k-2}-\frac{1}{2}\zeta)+\frac{1}{4}\delta^2 n\\ \ell_{a,b}+\ell_{b,a} &\le (r^-_{a,b}+r^-_{b,a})\left(\frac{1}{k-2}-\frac{1}{2}\zeta\right)+\frac{10\delta^2}{\beta}n.\end{aligned}$$

□

Proof of Lemma 8.11: We can assume that δ is sufficiently small for all bounds proven above. Suppose, as above, that STRIP continues to some iteration I^* at which point at least δn vertices have been removed from $\cup_{a=1}^k V_a^I$. W.h.p. the bound of Lemma 8.13 holds. This leads to a contradiction as follows:

$$\begin{aligned}\sum_{a\neq b}\ell_{a,b} &= \sum_{a<b}(\ell_{a,b}+\ell_{b,a})\\ &\le \binom{k}{2}\frac{10\delta^2}{\beta}n+\left(\frac{1}{k-2}-\frac{1}{2}\zeta\right)\sum_{a<b}(r^-_{a,b}+r^-_{b,a}) \qquad \text{by Lemma 8.13}\\ &= \binom{k}{2}\frac{10\delta^2}{\beta}n+\left(\frac{1}{k-2}-\frac{1}{2}\zeta\right)\sum_{a\neq b}r_{a,b} \qquad \text{since } r^-_{a,b}\le r_{a,b}\\ &\le \binom{k}{2}\frac{10\delta^2}{\beta}n+\left(\frac{1}{k-2}-\frac{1}{2}\zeta\right)\sum_{a\neq b}\sum_{b'\notin\{a,b\}}\ell_{a,b'} \qquad \text{by (32)}\\ &\le \binom{k}{2}\frac{10\delta^2}{\beta}n+\left(\frac{1}{k-2}-\frac{1}{2}\zeta\right)(k-2)\sum_{a\neq b}\ell_{a,b}\\ &\qquad \text{since each } \ell_{a,b} \text{ appears } k-2 \text{ times in the double summation}\\ &= \binom{k}{2}\frac{10\delta^2}{\beta}n+(1-\frac{k-2}{2}\zeta)\sum_{a\neq b}\ell_{a,b}.\end{aligned}$$

Rearranging yields

$$\sum_{a\neq b}\ell_{a,b}\le\binom{k}{2}\frac{20\delta^2}{(k-2)\beta\zeta}n.$$

Since $\sum_{a\neq b}\ell_{a,b}\ge\delta n$ (from (31)), this is a contradiction for $\delta<\frac{\zeta\beta}{20k}$ from (28). □

8.3 Proof of Lemma 6.3

Recall the definitions of $\lambda_k(c),\xi_k(c),\mu_k(c),\tau_k(c)$ from Section 6, and the lemma:
Lemma 6.3 *For any $c>c_k$ w.h.p. we have that for every a,b, the subgraph induced by $K_{a,b}$ is connected and:*

(a) $|K_i|=\lambda_k(c)n+o(n)$;

(b) the 2-core of $K_{a,b}$ has $\xi_k(c)n+o(n)$ vertices in K_a and $\xi_k(c)n+o(n)$ vertices in K_b;

(c) the 2-core of $K_{a,b}$ has $\mu_k(c)n+o(n)$ edges;

(d) the 2-core of $K_{a,b}$ has $\tau_k(c)n+o(n)$ degree 2 vertices in K_a and $\tau_k(c)n+o(n)$ degree 2 vertices in K_b.

Proof: WLOG we will focus on $a = 1, b = 2$. Let H denote the 2-core of the giant component of $K_{1,2}$.

Part (a) is essentially a repetition of Lemma 5.2(b). (Note that the proof of Lemma 5.2(b) showed that each K_i has size $\lambda_k(c)n + o(n)$.)

For parts (b,c): Consider any constant $\delta > 0$, and choose I large enough that $\rho_I < \rho + \delta$. The proof of Lemma 8.7 says that w.h.p. $K_{1,2} \subseteq K_{1,2}^I$ is contained in the giant component of $(U(\rho_I n + h_I(n)), W(\rho_I n + h_I(n)))$ for some specific $h_I(n) = o(n)$. Lemma 8.11 implies that $K_{1,2}$ contains the giant component of $(U((\rho - \delta)n), W((\rho - \delta)n))$. So H is sandwiched between the 2-cores of those two giant components.

Standard techniques (eg. [55, 64, 45, 33, 40]) show that the giant component of $G_{\rho n, \rho n, p=c/n}$ has a 2-core with $\xi_k(c)n + o(n)$ vertices on each side and $\mu_k(c)n + o(n)$ edges. This also follows from Exercise 2.4.8 of [35]. We omit the details. Recalling that the subgraph induced by $(U(x), W(x))$ is distributed like $G_{x,x,p=c/n}$, and applying continuity to those same arguments, we see that the giant components of $(U(\rho_I n + h_I(n)), W(\rho_I n + h_I(n)))$ and $(U((\rho - \delta)n), W((\rho - \delta)n))$ have 2-cores with $\xi_k(c)n \pm \delta' n$ vertices on each side and $\mu_k(c)n \pm \delta' n$ edges, for some $\delta' = \delta'(\delta)$ which tends to zero with δ. Since H is sandwiched between those two giant components, it also has $\xi_k(c)n \pm \delta' n$ vertices on each side and $\mu_k(c)n \pm \delta' n$ edges. Since this holds for any $\delta > 0$ and hence any $\delta' > 0$, parts (b,c) follow.

Part (d) follows immediately from the assertion that the distribution of the degree sequence of the 2-core of a random bipartite graph is asymptotic to a truncated Poisson distribution, truncated at $d \geq 2$. This is well-known for k-cores of non-bipartite graphs, i.e. of $G_{n,M}$. The bipartite case follows from the same straightforward arguments, but we did not find a proof anywhere; so we sketch it here:

We first expose the set of vertices, $\mathcal{V}$, and number of edges, E, of H and assume that they satisfy (b,c). The first observation is that every simple bipartite graph with those vertices and that number of edges and with minimum degree at least 2 is equally likely to be H in $P_{n,p}$. To see this, consider two choices H_1, H_2. Take any graph G_1 that is planted on colouring σ such that, after applying Kempe-strip to (G_1, σ), the 2-core of the giant component in $K_{1,2}$ is H_1. Form G_2 by replacing H_1 in G_1 with H_2. Note that this does not change any steps of Kempe-Strip and so if we apply Kempe-Strip to (G_2, σ), the 2-core of the giant component in $K_{1,2}$ is H_2. Since G_1, G_2 have the same number of edges, $(G_1, \sigma), (G_2, \sigma)$ are equally likely to be selected in $P_{n,p}$. So H_1, H_2 are equally likely to be H.

Let Ω denote the set of bipartite simple graphs with E edges on vertex set $\mathcal{V}$. Let v_1, v_2 denote the number of vertices in $\mathcal{V}$ on each side of the bipartition. Let Φ denote the set of possible degree sequences of graphs in Ω; i.e. the set of all pairs $\mathcal{D}_1, \mathcal{D}_2$ of sequences of integers $d_1, ..., d_{v_1}; d'_1, ..., d'_{v_2}$ where each sequence sums to E, and where each $d_i, d'_i \geq 2$. Consider the set of possible left-side degree seqeunces; i.e. the set of degree sequences $d_1, ..., d_{v_1}$ summing to E with each $d_i \geq 2$. If we weight each such degree sequence by $\prod \frac{1}{d_i!}$ then we obtain the *truncated multinomial distribution.* Corollary 2 of [19] says that the distribution of the degrees is asymptotic to a truncated Poisson distribution with mean $E/v_1 = y_k(c) = c\lambda_k(c)$ (since E, v_1 satisfy parts (b,c)); the same is true of the right-side degree sequences. In particular, noting that

$$\tau_k(c) = \xi_k(c) \times \mathbf{Pr}[Po_{\geq 2}(y_k(c)) = 2],$$

Corollary 2 of [19] implies that there is some $g(n) = o(n)$ (defined in terms of the implicit $o(n)$ terms in (b,c)) such that if we define Φ^+ to be the number of pairs $(\mathcal{D}_1, \mathcal{D}_2)$ where each sequence contains the number two $\tau_k(c)n \pm g(n)$ times, then

$$\sum_{(\mathcal{D}_1, \mathcal{D}_2) \in \Phi^+} \prod \frac{1}{d_i!} \prod \frac{1}{d'_j!} = (1 - o(1)) \sum_{(\mathcal{D}_1, \mathcal{D}_2) \in \Phi} \prod \frac{1}{d_i!} \prod \frac{1}{d'_j!}. \tag{34}$$

We define $s(\mathcal{D}_1, \mathcal{D}_2)$ to be the number of simple graphs on $\mathcal{V}$ with degree sequence $(\mathcal{D}_1, \mathcal{D}_2)$. To bound $s(\mathcal{D}_1, \mathcal{D}_2)$, consider a random configuration with that degree sequence; i.e. assign d_i, d'_j vertex-copies to the ith vertex on the left and jth vertex on the right and consider a random matching from the left vertex-copies to the right vertex-copies. Each simple bipartite graph on this degree sequence arises from $\prod d_i! \prod d'_j!$ matchings. So

$$s(\mathcal{D}_1, \mathcal{D}_2) \leq \frac{E!}{\prod d_i! \prod d'_j!}. \tag{35}$$

Furthermore, applying the Method of Moments to the number of double edges in the configuration (see Section 6.1 of [41]) shows that for at least $(1-o(1))|\Phi|$ pairs $\mathcal{D}_1, \mathcal{D}_2$, the probability that the random matching forms a simple bipartite graph is at least $\epsilon > 0$. (This is true, eg. for all pairs of degree sequences that are both well-behaved in the sense of [57].) So for all such pairs,

$$s(\mathcal{D}_1, \mathcal{D}_2) \geq \frac{\epsilon E!}{\prod d_i! \prod d'_j!}.$$

Summing over all such pairs in Φ^+ and applying (35) then (34), we have

$$\sum_{(\mathcal{D}_1,\mathcal{D}_2)\in\Phi^+} s(\mathcal{D}_1, \mathcal{D}_2) \geq (1-o(1)) \sum_{(\mathcal{D}_1,\mathcal{D}_2)\in\Phi^+} \frac{\epsilon E!}{\prod d_i! \prod d'_j!} \geq (1-o(1))\epsilon \sum_{(\mathcal{D}_1,\mathcal{D}_2)\in\Phi} \frac{E!}{\prod d_i! \prod d'_j!}.$$

Furthermore (34) and (35) also yield

$$\sum_{(\mathcal{D}_1,\mathcal{D}_2)\in\Phi\setminus\Phi^+} s(\mathcal{D}_1, \mathcal{D}_2) \leq \sum_{(\mathcal{D}_1,\mathcal{D}_2)\in\Phi\setminus\Phi^+} \frac{E!}{\prod d_i! \prod d'_j!} = o(1) \times \sum_{(\mathcal{D}_1,\mathcal{D}_2)\in\Phi} \frac{E!}{\prod d_i! \prod d'_j!}.$$

So the probability that a uniform member of Ω has a degree sequence in Φ^+ is $1-o(1)$. This proves (d). $\square$

9 Vertices outside the Kempe-core are unfrozen

In this section, we prove Lemma 5.4.
Lemma 5.4 *For $k \geq 3$, in $P_{n,p=c/n}$:*

(a) If $c > c_k$ then w.h.p. at most $o(n)$ vertices outside of the Kempe-core are $\log n$-frozen.

(b) If $c < c_k$ then there exists $Q = Q(c,k)$ such that: w.h.p. no vertex is $Q \log n$-frozen.

Remark: In fact, the proof shows that we can replace $\log n$ in part (a) by any $\omega(n) \to \infty$ with n.

The intuition is as follows: Recall that STRIP is a parallel version of Kempe-Strip, where we repeatedly remove all Kempe chains of size less than $\log^2 n$. Our bounds on the size of the small components of a random bipartite graph show that w.h.p. all removed Kempe chains have size $O(\log n)$.

Consider any one of the Kempe-chains, Φ, that is removed. There is a sequence of Kempe-chains, removed during previous iterations, which eventually led to Φ being small and hence deleteable. A bit of thought shows that we should typically be able to switch the colours in each of the Kempe-chains, in order of deletion, in a manner that permits us to eventually switch the colours in Φ. Since each chain has size $O(\log n)$, this is a sequence of small changes which leads to every vertex in Φ changing colours, and hence the vertices of Φ are unfrozen.

The only thing that can go wrong in this scenario is if some of the Kempe-chains in this sequence intersect each other, or are adjacent to each other, such that flipping one Kempe-chain interferes with another. If this happens, then some of these chains must be arranged in a cyclic pattern. We will show that these patterns affect very few vertices, and that we can deal with those vertices another way.

We will simplify our analysis by modifying the procedure so that the Kempe chain sizes are bounded by $O(1)$ rather than $\log n$. Specifically, for any constant T, we define:

***T*-STRIP**
Input: a graph G and a k-colouring $\sigma = S_1, ..., S_k$ of G.
While there are any Kempe chains of size at most T
Remove the vertices of *every* such Kempe chain from G.

Since this is a valid way of beginning to carry out the procedure Kempe-Strip (from Section 5), no vertices of the Kempe-core are removed by T-STRIP. Furthermore, if T is large then very few Kempe chains have size greater than T, so we can show that what remains after i iterations of T-STRIP is very close to what remains after i iterations of STRIP.

We use the same notation as in our description of STRIP, STRIP1 in Section 8, adding "(T)"; eg. $V_a^i(T)$ is the set of vertices with colour a remaining after i iterations of T-STRIP. It is straightforward to check that the proofs of Observations 8.4 and 8.5 also apply in this setting, thus yielding:

Observation 9.1. *Given $V_a^i(T), V_b^i(T)$ and $x_a = |K_{a,b}^i(T) \cap V_a^i(T)|, x_b = |K_{a,b}^i(T) \cap V_b^i(T)|$, the vertices of $K_{a,b}^i(T) \cap V_a^i(T)$ and $K_{a,b}^i(T) \cap V_b^i(T)$ can be treated as uniformly random subsets of $V_a^i(T), V_b^i(T)$ of sizes x_a, x_b, respectively. Furthermore, if $(a, b) \neq (a', b')$ then the random subsets selected for $K_{a,b}(T)$ are independent of the subsets chosen for $K_{a',b'}(T)$.*

Observation 9.2. *For each i, the set of components of size greater than T in $(U^i(T), W^i(T))$ is identical to the set of components of size greater than T in $(V_1^i(T), V_2^i(T))$.*

It is well-known that the number of vertices on non-giant components of size greater than T in the random bipartite graph $G_{n_1,n_2,p=c/n}$ is at most ϵn where $\epsilon \to 0$ while $T \to \infty$. This allows us to adapt the proof of Lemma 8.7 to obtain:

Lemma 9.3. *For any constants I, ϵ, there exists T such that w.h.p. we have for each $1 \leq i \leq I$, $1 \leq a < b \leq k$:*

(a) $\beta_i n \leq |K^i(T)_{a,b} \cap V_a^i(T)|, |K_{a,b}^i(T) \cap V_b^i(T)| \leq (\beta_i + \epsilon)n$*;*

(b) $\rho_i n \leq |U^i(T)|, |W^i(T)| \leq (\rho_i + \epsilon)n$*;*

(c) $\nu_i n \leq |V_a^i(T)| \leq (\nu_i + \epsilon)n$*.*

Proof Outline: Straightforward bounds on the expected number of components of size between $T + 1$ and $\log^2 n$ in $G_{n_1,n_2,p=c/n}$ (such bounds appear in the proof of Theorem 6 of [42]) allow us to modify Lemma 8.1(a.i) to prove that for any $\delta > 0$ we can choose T such that at most $(\alpha_1 + \delta)n, (\alpha_2 + \delta)n$ vertices on each side of $G_{n_1,n_2,p=c/n}$ lie in components of size greater than T. This and Observations 9.1, 9.2 are enough for the proof of Lemma 8.7 to apply in this setting. That proof yields the upper bounds of this lemma, so long as we choose δ small enough that the increase in the recursive bounds accumulates to less than ϵ. The lower bounds are straightforward - clearly this procedure removes fewer vertices than STRIP does, and it is not hard to show that the difference is $\Theta(n)$, since at any stage there are $\Theta(n)$ vertices in components of size between T and $\log n$; i.e. vertices that would be removed by STRIP but are not removed by T-STRIP. $\square$

Next we show that the vertices removed by T-STRIP can have their colours changed unless they lie very close to a cycle of size $O(1)$.

Consider a Kempe chain C that is removed during our stripping procedure; i.e. C is a component of some $(V_a^i(T), V_b^i(T))$ with size less than $\log^2 n$. The *colours of* C are a, b - the two colours appearing on C. To *swap the colours of* C means to change the colour of every vertex in C from a to b or from b to a. To be clear: Suppose that a vertex $v \in V_a^i(T)$ has no neighbours in $V_b^i(T)$ nor in $V_{b'}^i(T)$ where a, b, b' are three different colours. Thus v forms a component of size 1 in $(V_a^i(T), V_b^i(T))$ and a component of size 1 in $(V_a^i(T), V_{b'}^i(T))$. These are considered to be two different components, one with colours a, b and the other with colours a, b'; swapping the colours in eg. the first of these components means changing the colour of v from a to b.

We now define the partially directed graph Γ on the components that are removed during T-STRIP. The edges of Γ join components which could interfere with each other when we swap their colours. An edge of Γ is directed from C_1 to C_2 if the removal of C_1 helped reduce the size of C_2 to be below $\log^2 n$. More formally:

Definition 9.4. *The vertices $V(\Gamma)$ are the components (i.e. Kempe chains) that are removed during T-STRIP. Two such components in $V(\Gamma)$ are joined by an edge of Γ if they share a vertex or if they are joined by an edge in the original graph, G.*

Furthermore, an edge of Γ is directed from component C_1 to C_2 if:

(i) C_1, C_2 have a colour in common,

(ii) C_1 is removed during an earlier iteration of T-STRIP than C_2, and

(iii) C_1, C_2 are joined by an edge in G whose endpoints are the colours of C_2.

For any component $C \in V(\Gamma)$, we define $\Upsilon(C)$ to be the subgraph of Γ induced by the set of components that can reach C using the directed edges *of Γ.*

A cycle *in Γ is a set of edges that form a cycle after removing any directions. We say that $\Upsilon(C)$ is a* Γ-tree *if (i) it has no cycles and (ii) no $C_1, C_2 \in \Upsilon(C)$ are joined by two edges in G.*

Remark: In what follows, the vertices of Γ are sometimes referred to as Kempe chains (as they are removed by T-STRIP) and are sometimes referred to as components (as they are components of one of the bipartite graphs).

Lemma 9.5. *If C is deleted by T-STRIP and $\Upsilon(C)$ is a Γ-tree then the vertices of C are not T-frozen.*

Proof We will prove that we can swap the colours on a sequence of some Kempe chains in $\Upsilon(C)$, culminating with a swap of the colours on C. Since each of these Kempe chains has size at most T, this proves the lemma. We prove this by induction on the iteration of T-STRIP in which C was removed. If it was removed in iteration 1, then we can swap the colours of C without first swapping any other colours.

WLOG suppose that C is a component of $(V_1^i(T), V_2^i(T))$ for some $i > 1$. Let $B_1, ..., B_q$ be the components that point to C in Γ. By induction, it is possible to swap the colours on a sequence of some components in each $\Upsilon(B_j)$ culminating in a swap of the colours of B_j. Furthermore, since $\Upsilon(C)$ is a Γ-tree, no component in any $\Upsilon(B_j)$ shares a vertex or is joined by an edge with any component in some other $\Upsilon(B_{j'})$, else the edge joining those two components in Γ would create a cycle. So the swaps in each $\Upsilon(B_j)$ do not alter the colours of any vertices in or adjacent to any components in any other $\Upsilon(B_{j'})$ and thus we can carry out every sequence of swaps. We argue that we can now swap the colours on $\Upsilon(C)$.

Every neighbour u of C outside of $(V_1^i(T), V_2^i(T))$ that had colour 1 or 2 lies in some B_j and so had its colour swapped. Note that the other colour in B_j cannot be 1 or 2, as otherwise B_j would have had to include u's neighbour in C; so the colour of u is now neither 1 nor 2. C does not have any other neighbours in B_j since $\Upsilon(C)$ is a Γ-tree. No neighbour of C outside of $B_1, ..., B_q$ lay in any components of $\Upsilon(B_1), ..., \Upsilon(B_q)$, since $\Upsilon(C)$ is a Γ-tree; so no such neighbour had its colour changed. So after our sequence of swaps, C has no neighbours outside of $(V_1^i(T), V_2^i(T))$ with colour 1 or 2. Since C is a component of $(V_1^i(T), V_2^i(T))$, we can swap the colours of C. $\square$

The fact that there are very few $O(1)$-length cycles in our underlying random graph implies that Lemma 9.5 applies to almost all vertices removed by T-STRIP:

Lemma 9.6. *(a) For any I, T, the expected number of vertices deleted during the first I rounds of T-STRIP which lie on Kempe chains C for which $\Upsilon(C)$ is not a Γ-tree is $O(1)$.*

(b) W.h.p. no two cycles of length at most $2IT$ in G intersect or are joined by a path of length at most $3IT$.

Proof *Part (a):* If $\Upsilon(C)$ is not a Γ-tree, then G contains two paths from some component C' to C. Those two paths and a path in C' must form a cycle. Each path has length less than I and each component in the path has size at most T. So that cycle corresponds to a cycle of length less than $2IT$ in the original graph G. Furthermore each vertex of C is within distance IT from that cycle. Since $2IT = O(1)$, a straightforward first-moment calculation shows that the expected number of vertices in the planted random graph $P_{n,p=c/n}$

shows that the expected number of vertices within distance IT of a cycle of length at most $2IT$ is $O(1)$. Indeed, letting ℓ be the length of the cycle and t be the length of the path (i.e. the number of edges), and overcounting by possibly choosing two vertices of the same colour to be adjacent, the expected number is at most

$$\sum_{1\leq \ell\leq 2IT, 0\leq t\leq IT} (\ell+t)n^{\ell+t}\left(\frac{c}{n}\right)^{\ell+t} = O(1),$$

since $I, T = O(1)$.

Part (b): We use a very similar first moment bound: ℓ_1, ℓ_2 are the lenghts of the two cycles and t is the length of the path. The expected number of such subgraphs is at most

$$\sum_{1\leq \ell_1,\ell_2\leq 2IT, 0\leq t\leq IT} n^{\ell_1+\ell_2+t-1}\left(\frac{c}{n}\right)^{\ell_1+\ell_2+t} = O(1/n).$$

□

This immediately yields Lemma 5.4 above the freezing threshold as follows:

Proof of Lemma 5.4(a) Consider any $\delta > 0$. Lemmas 9.3 and 8.6 imply that we can choose constants I, T sufficiently large so that w.h.p. I rounds of T-STRIP leave at most $(\beta+\delta)n$ remaining vertices of each colour. Lemmas 9.5 and 9.6 imply that w.h.p. at most $o(n)$ of the deleted vertices are T-frozen. W.h.p. the Kempe-core has $\beta n + o(n)$ vertices of each colour (Lemma 5.2(b) and (22)) and no vertices of the Kempe-core are removed by T-STRIP, so fewer than $2k\delta n$ vertices outside of the Kempe-core are T-frozen. Since this holds for every $\delta > 0$, and $T < \log n$, at most $o(n)$ vertices outside the Kempe-core are $\log n$-frozen. □

The same proof can show that, below the freezing threshold, at most $o(n)$ vertices are $\log n$-frozen. To show the stronger statement that **no** vertices are $Q\log n$-frozen, we have to be careful about how to handle the components C for which $\Upsilon(C)$ is not a Γ-tree.

Lemma 8.6 shows that for $c < c_k$, $\lim_{i\to\infty}\rho_i = 0$. We choose I such that $\rho_I < \frac{1}{4c}$. We let G_{I+1} denote the set of vertices remaining after I rounds of T-strip.

We take $\epsilon < \frac{1}{4c}$ and so Lemma 9.3(b) implies that $|U^I(T)|, |W^I(T)| < \frac{n}{2c}$. As described in the proof of Lemma 9.3, the argument from Lemma 8.7 applies in this setting; this implies that the bipartite graph $(U^I(T), W^I(T)$ is contained in $(U(n/2c), W(n/2c))$, and the latter bipartite graph is distributed like the random bipartite graph $G_{n/2c,n/2c,p=c/n}$. Thus w.h.p. every component of $(|U^I(T)|, |W^I(T)|)$ has size at most $Q\log n$, where Q comes from Lemma 8.1. Observation 9.2 thus yields:

$$\text{Every Kempe chain in } G_{I+1} \text{ has size at most } Q\log n. \tag{36}$$

Definition 9.7. *(a) $\mathcal{B}$ is the set of vertices that lie in any Kempe chain C removed during rounds $1, ..., I$ for which $\Upsilon(C)$ is not a Γ-tree. Abusing notation, we sometimes say that the Kempe chain C is in $\mathcal{B}$.*

(b) $\mathcal{B}^+$ is the set of vertices v which are joined by at least two edges to one Kempe chain which was deleted before v.

(c) A blocker *of a Kempe chain W in G_{I+1} is either*

(i) a neighbour of W that is in $\mathcal{B}$ and whose colour is a colour of W; or

(ii) a path of length at most $3IT$ through the vertices in $G\backslash G_{I+1}$ whose endpoints are both adjacent to W.

Note that $\mathcal{B}^+$ may intersect $\mathcal{B}$. Specifically, if $v \in \mathcal{B}^+$ is in a Kempe chain C that is removed during iterations $1, ..., I$ and if the Kempe chain removed prior to v which causes v to be in $\mathcal{B}^+$ is in $\Upsilon(C)$, then the two edges joining v to that Kempe chain will cause $\Upsilon(C)$ to not be a Γ-tree and so $v \in \mathcal{B}$. Note also that if $v \in \mathcal{B}^+$ is in G_{I+1} then any path joining the two neighbours of v in that Kempe chain will be a blocker of every Kempe chain in G_{I+1} containing v.

Note that Lemma 9.6 implies that $\mathbf{Exp}(|\mathcal{B}|) = O(1)$. In fact, the calculation used in that proof also implies $\mathbf{Exp}(|\mathcal{B}^+|) = O(1)$; indeed, every member of $\mathcal{B}^+$ lies in a cycle of length at most $T+1$ and it is well-known that the expected number of vertices lying in such cycles is $O(1)$.

Lemma 9.8. *Suppose W is a Kempe chain in G_{I+1} that has no blockers. Then no vertex in W is $Q\log n$-frozen.*

Proof Suppose WLOG that W is a Kempe chain with colours 1, 2; i.e. a component of $(V_1^I(T), V_2^I(T))$. By (36) $|W| \le Q\log n$. Let $u_1, ..., u_\ell$ be the neighbours of W in $G\backslash G_{I+1}$ that have colour 1 or 2. Each u_i lies in a Kempe chain C_i which was deleted during the first I rounds and which points to W in Γ; since W has no blocker, each $\Upsilon(C_i)$ is a Γ-tree. So we can swap the colours on a sequence of chains in $\Upsilon(C_i)$ and then swap the colours on C_i (as in the proof of Lemma 9.5). We must argue that that the sequence of swaps in $\Upsilon(C_i)$ does not conflict with the sequence in $\Upsilon(C_j)$ for any $1 \le i < j \le \ell$. That would require a Kempe chain in $\Upsilon(C_i)$ to either intersect or be joined by an edge to a Kempe chain in $\Upsilon(C_j)$. Since each Kempe chain has size at most T and they were both removed during the first I iterations, this would yield a path in G of length at most $2TI$ from u_i to u_j; but this would be a blocker of W. So we can change the colour of each u_i by swapping a sequence of chains of size at most T.

No C_i has colours $(1, 2)$ else the adjacent member of W would also be in C_i. And W is not adjacent to two vertices of any C_i as those vertices are joined by a path in G of length less than T which would form a blocker of W. And W has no neighbours other than u_i in any Kempe chain of $\Upsilon(C_i)$, else this would form a blocker path of length at most TI. Thus no neighbours of W had their colours changed to 1 or 2. So after these switches, W has no neighbours in G of colour 1 or 2. So we can swap the colours on W. Since $|W| \le Q\log n$ by (36), and all other chains in the sequence of flips have size at most $T = O(1)$, the vertices of W are not $Q\log n$-frozen. □

To prove that no vertices in G_{I+1} are frozen, we will show that for every vertex $v \in G_{I+1}$, at least one of the Kempe-chains containing v is not blocked.

Lemma 9.9. *(a) No Kempe-chain in G_{I+1} has two blockers;*

(b) No vertex $v \in G_{I+1}$ is in two Kempe-chains in G_{I+1} that have different blockers.

(c) No two Kempe-chains in G_{I+1} are adjacent to the same component in $\mathcal{B}$ and are each adjacent to a different endpoint of a path of length at most $3IT$ in $G\backslash G_{I+1}$.

(d) There is no path P of length at most $3IT$ in $G\backslash G_{I+1}$ and vertices v_1, u_1, v_2, u_2 such that: (i) $v_1 \ne u_1, v_2 \ne u_2$; (ii) each of v_1, u_1, v_2, u_2 is adjacent to an endpoint of P; and (iii) there are two different Kempe chains in G_{I+1}, one containing v_1 and u_1 and the other containing v_2 and u_2.

Remark: Note that part (b) proscribes *different* blockers. As explained above, if $v \in \mathcal{B}^+$ then a path through the deleted Kempe chain which is joined by two edges to v will be a blocker of every Kempe chain in G_{I+1} containing v. So in that case v will be in multiple Kempe-chains that all have *the same* blocker.

Proof Consider any two colours, WLOG the colours 1,2. Recall that the subgraph of G_{I+1} induced by colours 1,2 is a subgraph of (U^{I+1}, W^{I+1}) (see eg. Observation 9.2). As described above (36), (U^{I+1}, W^{I+1}) is a subgraph of $(U(n/2c), W(n/2c))$ which is distributed like the random bipartite graph $G_{n/2c,n/2c,p=c/n}$. A simple first moment argument shows that for any two vertices u, v of colours 1 or 2 in G^{I+1}, the probability that they are in the same component of G^{I+1} is $O(1/n)$. Indeed, letting ℓ denote the length of a path joining u, v, the probability that they are in the same component of $(U(n/2c), W(n/2c))$ is

$$\sum_{\ell\ge1}\left(\frac{n}{2c}\right)^{\ell-1}\left(\frac{c}{n}\right)^{\ell} = \frac{c}{n}\times\sum_{\ell\ge1}\left(\frac{1}{2}\right)^{\ell-1} = \frac{2c}{n}. \tag{37}$$

To prove parts (a,b), we will show that w.h.p. there is no vertex v and two different blockers ϕ_1, ϕ_2 each blocking a Kempe-chain in G_{I+1} containing v. We will do so by bounding the expected number of such v, ϕ_1, ϕ_2.

Case 1: ϕ_1, ϕ_2 are both vertices in $\mathcal{B}$. The computations from the proof of Lemma 9.6(a) show that the contribution of the number of choices for the vertices ϕ_1, ϕ_2, multiplied by the probability that they are both in $\mathcal{B}$ is $O(1)^2 = O(1)$. The number of choices for v is at most n. There are $(k-1)^2 = O(1)$ choices for the Kempe chains containing v that are adjacent to ϕ_1, ϕ_2 (we may choose the same Kempe chain to be adjacent to both). Choose u_1, u_2, the vertices in those Kempe chains that are adjacent to ϕ_1, ϕ_2. If $u_i \neq v$ then we get at most n choices for u_i, and we multiply by the probability that u_1, v are in the same Kempe chain - $\frac{2c}{n}$ by (37). The probability that u_i is adjacent to ϕ_i is $\frac{c}{n}$ for each i. Note that the edge sets we consider here are all disjoint: The edges of the Kempe chains containing v are in G_{I+1}; the edges of the short paths to short cycles that must exist if $\phi_1, \phi_2 \in \mathcal{B}$ are in $G \backslash G_{I+1}$ (see the proof of Lemma 9.6(a)); the edges $u_i\phi_i$ run between G_{I+1} and $G \backslash G_{I+1}$. Putting this together, the expected number of such v, ϕ_1, ϕ_2 is at most:

$$O(1) \times n \times \left(n \times \frac{2c}{n}\right)^{1,2 \text{ or } 3} \times O\left(\frac{1}{n}\right)^2 = O\left(\frac{1}{n}\right).$$

Case 2: ϕ_1 is a vertex in $\mathcal{B}$ and ϕ_2 is a path of length at most $3IT$. The endpoints of ϕ_2 are x, y and the length of ϕ_2 is ℓ; we allow $\ell = 0$ in which case $x = y$. If $\ell = 0$ there are fewer than n choices for x. If $\ell > 0$, there are fewer than $\binom{n}{2}$ choices for x, y, and calculation very much like the one in the proof of Lemma 9.6(a) yields that the probability that $\phi_1 \in \mathcal{B}$ and x, y are joined by a path of length at most $3IT$ is $O(1/n^2)$. The neighbours of x, y in the Kempe chain containing v are u_1, u_2, possibly $u_1 = u_2$ and either or both of them could be v. If v, u_1, u_2 are three different vertices, then the probability that they are all in the same Kempe chain of G_{I+1} is $O(1/n^2)$ by a calculation very much like (37); we omit the details. If they are two different vertices then that probability is $O(1/n)$ by (37). As in Case 1, the probability that u_1 is adjacent to x and u_2 is adjacent to y is $O(1/n)^2$. We use the calculations from Case 1 to bound the contribution of the choices for v, ϕ_1 and the probability that ϕ_1 is adjacent to the other Kempe chain containing v. Putting all this together again yields an expection of $O(1/n)$.

Case 3: ϕ_1, ϕ_2 are both paths of length at most $3IT$. Calculations just like those in Case 2, again yield an expectation of $O(1/n)$.

Parts (c,d) follow from very similar calculations, each time obtaining an expectation of $O(1/n)$. We omit the straightforward but tedious details. □

Corollary 9.10. *No vertex in $G_{I+1} \backslash \mathcal{B}^+$ is $Q \log n$-frozen.*

Proof Suppose WLOG $v \in V_1^{I+1}$. Since $k \geq 3$, v lies in at least two Kempe chains of G_{I+1}. By Lemma 9.9(b) either those two Kempe chains have the same blocker, or at least one of them has no blockers. In the latter case, Lemma 9.8 implies that v is not $Q \log n$-frozen. So we turn our attention to the former case.

If two Kempe chains in G_{I+1} containing v have the same blocker, then that blocker cannot be a vertex in $\mathcal{B}$ as the colour of that vertex can only be in one of the chains. So the blocker must be a path P whose endpoints w_1, w_2 are adjacent to vertices in both Kempe chains. If v is not adjacent to both w_1, w_2 then P will violate Lemma 9.9(d), possibly with $v_1 = v_2 = v$ and possibly with $u_1 = u_2$ if the two Kempe chains have two vertices in common. If v is adjacent to both w_1, w_2 then w_1, w_2 cannot be in the same deleted Kempe chain since $v \notin \mathcal{B}^+$. Therefore w_1, w_2 are in two different Kempe chains C_1, C_2 that were removed during iterations $1, ..., I$.

Neither w_1 nor w_2 have the same colour as v, so at least one of the two Kempe chains containing v does not have the colour of w_1 and the colour of w_2. Noting that that Kempe chain has no blocker other than P (otherwise we would violate Lemma 9.9(b)), the argument from the proof of Lemma 9.8 shows that the vertices of that Kempe chain are not $Q \log n$-frozen. □

Next we turn to $\mathcal{B}$:

Lemma 9.11. *No vertex in $\mathcal{B} \backslash \mathcal{B}^+$ is frozen.*

Proof Suppose WLOG v has colour 1 and is in a Kempe-chain $C \in \mathcal{B}$ with colours $1, 2$ which was removed during iteration $i \leq I$. As we described in the proof of Lemma 9.6, there must be a cycle in G of length at most $2TI$ through the components of $\Upsilon(C)$.

If v lies in another Kempe chain C' that is deleted during iteration i then we will argue that $C' \notin \mathcal{B}$. If $\Upsilon(C)$ and $\Upsilon(C')$ both have different cycles then those two cycles are joined by a path through v of length at most $2iT \leq 2IT$; this violates Lemma 9.6(b).

Suppose $\Upsilon(C)$ and $\Upsilon(C')$ both contain the same cycle. If v lies in that cycle then it must contains an edge joining v to a component $D \in \Upsilon(C)$ and an edge joining v to a component $D' \in \Upsilon(C')$; note that the endpoints of those two edges have the colours of C, C' and so they are two different edges. Since $v \notin \mathcal{B}^+$ we must have $D \neq D'$. Since the cycle is in $\Upsilon(C)$ and $\Upsilon(C')$, we must have $D \in \Upsilon(C')$ and $D' \in U(C)$. Since no edge of Γ is directed in both directions, this leads to a violation of Lemma 9.6(b); i.e. two short cycles close to v. Thus v does not lie in the cycle and so there are two paths in G from v to that cycle - one using an edge pointing to C in Γ and one using an edge pointing to C' in Γ. That cycle, along with the two paths to it from v, must violate Lemma 9.6(b). So $C' \notin \mathcal{B}$ and thus Lemma 9.5 implies that v is not T-frozen.

Of course, v cannot lie in two Kempe chains that are removed in different iterations, so we can assume:

$$v \text{ does not lie in any other Kempe chains that are removed during the first } I \text{ iterations.} \tag{38}$$

We will swap the colour of v to 3. Let $u_1, ..., u_\ell$ denote the neighbours of v in G that have colour 3.

Step 1: If u_j is in a Kempe-chain $W_j \notin \mathcal{B}$ removed during the first I iterations then we swap the colours of some chains in $\Upsilon(W_j)$ and then swap the colours of W_j as in the proof of Lemma 9.5.

Step 2: If $u_j \in G_{I+1}$ then let W_j denote the Kempe-chain in G_{I+1} with colours $2, 3$ containing u_j. Lemma 9.9(b) implies that W_j has no blocker, since v is a blocker for the Kempe-chain in G_{I+1} with colours $1, 3$ containing u_j and v cannot be a blocker for W_j because v has colour 1. We swap the colours of a sequence of chains in $G \backslash G_{I+1}$ and then swap the colours of W_j as in the proof of Lemma 9.8.

Step 3: Each remaining vertex u_j is in a Kempe chain $W_j \in \mathcal{B}$. We will argue that (i) W_j has colours $(1, 3)$ and so $C \in \Upsilon(W_j)$, and (ii) no neighbour of W_j with colour 1 or 3, other than v, is in $\mathcal{B}$. This implies that for each W_j we can uncolour some of the chains in $\Upsilon(W_j)$ as in the proof of Lemma 9.5 so that v is the only neighbour of W_j that still has colour 1 or 3. We then simultaneously swap the colours of v and every W_j. Since each W_j has size at most T, the number of vertices in this final swap is at most $T \deg_G(v) < Q \log n$ as the maximum degree in G is easily computed to by w.h.p. $o(\log n)$.

Consider any $W_j \in \mathcal{B}$ from Step 3. Thus $\Upsilon(W_j)$ contains a cycle. The argument from the proof of (38) shows that this must be the cycle in $\Upsilon(C)$ as otherwise we violate Lemma 9.6(b); the only difference is that v is joined by an edge to W_j whereas above we had $v \in C'$. Furthermore, the path from $\Upsilon(W_j)$ to that cycle must pass through C. This proves that $C \in \Upsilon(W_j)$ as claimed above. Finally, if any other neighbour of W_j which was deleted before W_j is in $\mathcal{B}$, then this would result in either a second cycle within distance TI of W_j, or two paths from W_j to C; either way, this violates Lemma 9.6(b).

Claim 1: All of these swaps do not interfere with each other.

Proof: If two sequences in Step 1 or 3 interfere with each other then this will create either two cycles of size at most $2TI$ that are both within distance TI of v, or one cycle of size $2TI$ that is joined to v by two different paths of length at most $2TI$; either way this violates Lemma 9.6(b).

If a sequence of swaps to change the colour of some W_j from Step 2 interferes with a sequence from Step 1 or 3 then these sequences form a path in $G \backslash G_{I+1}$ of length at most $2TI$ from v to a neighbour of W_j with colour 2 or 3; but this contradicts the fact, proven above, that W_j has no blocker.

Finally, if the sequences of swaps that change $W_j, W_{j'}$ in Step 2 interfere with each other, then they form a path of length at most $2TI$ in $G \backslash G_{I+1}$ from a neighbour of W_j to a neighbour of W_{j+1}. Since $W_j, W_{j'}$ are both adjacent to v, this violates Lemma 9.9(c).

Claim 2: None of these swaps change the colour of v, until the final swap in Step 3.

Proof: This also follows from Lemmas 9.6(b) and 9.9. We omit the argument since it is not neccessary - if a swap changes the colour of v then v is not $Q \log n$-frozen.

Claim 3: none of these swaps change the colour of a neighbour of v to 3.

Proof: If v has a neighbour, other than u_j in one of the Kempe chains swapped in Steps 1 or 3 in order to change u_j, then the edges from v to that neighbour and to u_j create a cycle of length less than TI. That cycle along with the cycle in $\Upsilon(C)$ violate Lemma 9.6(b). If v has a neighbour, other than u_j in one of the Kempe chains swapped in Step 2 in order to change u_j, then this forms a path of length at most TI from v to another neighbour of W_j which contradicts the fact that W_j has no blocker.

These Claims ensure that we can carry out all three steps, eventually changing the colour of v in Step 3. (36) implies that each Kempe chain we switch has size at most $Q \log n$. □

Having proven Lemmas 9.8, 9.11 and Corollary 9.10 it only remains to show:

Lemma 9.12. *No vertex in $\mathcal{B}^+$ is frozen.*

Proof Suppose WLOG that $v \in \mathcal{B}^+$ has colour 1, and that v has two neighbours in a Kempe chain removed before v with colours $2, 3$. If exactly one of those neighbours, y, has colour 3 then we will change the colours of v and y as follows: Let $u_1, ..., u_\ell$ be the neighbours of v with colour 3, other than y; let $w_1, ..., w_{\ell'}$ be the neighbours of y with colour 1, other than u. We can make a series of Kempe chain swaps, similar to those in Steps 1, 2, 3 in the proof of Lemma 9.11, to change each u_j to 1, change each w_j to 3, and change v to 3 and y to 1. Again, each chain has size less than $Q \log n$.

If neither of the two neighbours in that Kempe chain have colour 3, then we let $u_1, ..., u_\ell$ be all neighbours of v with colour 3. Again, we can make a series of Kempe chain swaps to change each u_j to 1 and change v to 3.

We omit the repetitive details, other than to remark that the short cycle through v and the Kempe chain containing y plays the role of the cycle in $\Upsilon(C)$ in the proof of Lemma 9.11. □

This completes our proof of Lemma 5.4.

Proof of Lemma 5.4(b): This follows immediately from Lemmas 9.8, 9.11, 9.12 and Corollary 9.10. □

References

[1] E. Abbe, A. Montanari. *On the concentration of the number of solutions of random satisfiability formulas.* arXiv:1006.3786v1

[2] D. Achlioptas and A. Coja-Oghlan. *Algorithmic Barriers from Phase Transitions.* Proceedings of FOCS (2008), 793 - 802. Longer version available at arXiv:0803.2122

[3] D. Achlioptas, A. Coja-Oghlan and F. Ricci-Tersenghi. *On the solution-space geometry of random constraint satisfaction problems.* Random Structures and Algorithms **38** (2011), 251 - 268.

[4] D. Achlioptas and E. Friedgut. *A sharp threshold for k-colorability.* Random Struct. Algorithms **14** (1999), 63 - 70.

[5] D. Achlioptas and M. Molloy. *The analysis of a list- coloring algorithm on a random graph.* Proceedings of FOCS (1997), 204 - 212.

[6] D. Achlioptas and M. Molloy. *The solution space geometry of random linear equations.Random Structures & Algorithms* **46** (2015), 197 - 231. arXiv:1107.5550 .

[7] D. Achlioptas and C. Moore. *Almost all graphs of degree 4 are 3-colorable.* J. Comp. & Sys. Sci. **67** (2003), 441 - 471.

[8] D. Achlioptas and A. Naor. *The two possible values of the chromatic number of a random graph.* Annals of Mathematics, **162** (2005), 1333 - 1349.

[9] D. Achlioptas and Y. Peres. *The threshold for random k-SAT is $2^k \log 2 - O(k)$.* J.AMS **17** (2004), 947 - 973.

[10] D. Achlioptas and F. Ricci-Tersenghi. *On the solution-space geometry of random constraint satisfaction problems.* Proceedings of STOC (2006), 130 - 139.

[11] V. Bapst, A. Coja-Oghlan and C. Efthymiou. *Planting colourings silently.* Manuscript (2014).

[12] V. Bapst, A. Coja-Oghlan, S. Hetterich, F. Rassmann and D. Vilenchik. *The condensation phase transition in random graph coloring.* Proc. 18th RANDOM (2014), 449 - 464.

[13] A. Braunstein, L. Dall'Asta, G. Semerjian and L. Zdeborova. *The large deviations of the whitening process in random constraint satisfaction problems.* arXiv:1602.01700 (2016).

[14] A. Braunstein, M. Mezard and R. Zecchina. *Survey propagation: an algorithm for satisfiability.* Random Structures and Algorithms **27** (2005), 201 - 226.

[15] B. Bollobás. *A probabilistic proof of an asymptotic formula for the number of labelled regular graphs.* Europ. J. Combinatorics **1** 311-316 (1980).

[16] B. Bollobás, *Random Graphs.* 2nd edition. Cambridge University Press, 2001.

[17] B. Bollobás, S. Janson and O. Riordan, *The phase transition in inhomogeneous random graphs.* Rand. Struc. & Alg. **31** (2007), 3122.

[18] B. Bollobás and O. Riordan, *Asymptotic normality of the size of the giant component via a random walk.* J.Comb.Th. (B) **102** (2012), 53 - 61.

[19] J. Cain and N. Wormald, *Encores on cores.* Electron. J. Combin., 13 (2006), no. 1, Research Paper 81.

[20] P. Cheeseman, B. Kanefsky and W. Taylor. *Where the really hard problems are.* Proceedings of IJCAI (1991), 331 - 337.

[21] A. Coja-Oghlan. *A better algorithm for random k-SAT.* SIAM Journal on Computing **39** (2010), 2823 - 2864.

[22] A. Coja-Oghlan. *On belief propagation guided decimation for random k-SAT.* Proc. 22nd SODA (2011), 957 - 966.

[23] A. Coja-Oghlan. *Upper-bounding the k-colorability threshold by counting covers.* Elec. J. Comb. **20** (2013), P32.

[24] A. Coja-Oghlan and C. Efthymiou. *On independent sets in random graphs.* Proc. 22nd SODA (2011), 136 - 144.

[25] A. Coja-Oghlan, C. Efthymiou, S. Hetterich. *On the chromatic number of random regular graphs.* JCT(B) **116** (2015), 367 - 439.

[26] A. Coja-Oghlan and K. Panagiotou. *The asymptotic k-SAT threshold.* Advances in Mathematics **288** (2015), 985 - 1068.

[27] A. Coja-Oghlan, W. Perkins and K. Skubch. *Limits of discrete distributions and Gibbs measures on random graphs.* arXiv: arXiv:1512.06798v1.

[28] A. Coja-Oghlan and D. Vilenchik. *Chasing the k-colorability threshold.* Proc. 54th FOCS (2013), 380 - 389.

[29] A. Coja-Oghlan and L. Zdeborov. *The condensation transition in random hypergraph 2-coloring.* Proceedings of SODA (2012).

[30] J. Ding, A. Sly, and N. Sun, *Proof of the satisfiability conjecture for large k.* arXiv:1411.0650, (2014).

[31] P. Erdős and A. Rényi. *On random graphs I.* Publ. Math. Debrecen **6** (1959), 290 - 297.

[32] P. Erdős and A. Rényi. *On the evolution of random graphs.* Magayar Tud. Akad. Mat. Kutato Int. Kozl. **5** (1960), 17 - 61.

[33] D. Fernholz and V. Ramachandran. *Cores and Connectivity in Sparse Random Graphs.* The University of Texas at Austin, Department of Computer Sciences, technical report TR-04-13 (2004).

[34] U. Feige, A. Flaxman, and D. Vilenchik. *On the diameter of the set of satisfying assignments in random satisfiable k-CNF formulas.* SIAM J. Disc.Math. **25** (2011), 736 - 749. (2011)

[35] A. Frieze and M. Karonski. *Introduction to Random Graphs.* Cambridge University Press (2016).

[36] P. Gao and M. Molloy, *Inside the clustering window for random linear equations.* Submitted.

[37] A. Gerschenfeld and A. Montanari. *Reconstruction for models on random graphs.* Proceedings of FOCS 2007.

[38] G.R.Grimmett and C.J.H.McDiarmid. *Onc olouring random graphs.* Math. Proc. Cambridge Philos. Soc., **77** (1975), 313324.

[39] M. Ibrahimi, Y. Kanoria, M. Kraning and A. Montanari. *The set of solutions of random XORSAT formulae.* Proceedings of SODA 2012. Longer version available at arXiv:1107.5377

[40] S. Janson and M. Luczak. *A simple solution to the k-core problem.* Random Structures Algorithms **30** (2007) 50 - 62 (2007).

[41] S. Janson, T. Łuczak and A. Ruciński. Random Graphs. Wiley, New York (2000).

[42] T. Johansson. *The giant component of the random bipartite graph.* M.Sc. thesis, Chalmers University of Technology (2012).

[43] R. Karp. *The transitive closure of a random digraph.* Random Structures and Algorithms **1** (1990), 73 - 93.

[44] A. B. Kempe. *On the geographical problem of the four colors.* Amer. J. Math., **2** (1879), 193 - 200.

[45] J.H.Kim. *Poisson cloning model for random graphs.* arXiv:0805.4133v1

[46] M. Krivelevich, B. Sudakov, and D. Vilenchik. *On the random satisfiable process.* Combinatorics, Probability and Computing **18** (2009), 775 - 801.

[47] F. Krzakala and J. Kurchan. *Constraint optimization and landscapes.* Eur. Phys. J. B **64** (2008), 563 - 565.

[48] F. Krzakala, A. Montanari, F. Ricci-Tersenghi, G. Semerjian and L. Zdeborova. *Gibbs States and the Set of Solutions of Random Constraint Satisfaction Problems.* Proc. Natl. Acad. Sci., (2007).

[49] S. Kudekar and N. Macris. *Decay of correlations for sparse graph error correcting codes.* SIAM J. Disc. Math. **25** (2011), 956 - 988.

[50] E. Maneva, E. Mossel and M. J. Wainwright. *A new look at Survey Propagation and its generalizations.* JACM **54** (2007), 2 - 41.

[51] M. Mezard and A. Montanari. *Information, Physics and Computation.* Oxford University Press, (2009).

[52] M. Mezard, R. Zecchina *The random K-satisfiability problem: from an analytic solution to an efficient algorithm.* Phys. Rev. E **66**, (2002).

[53] D. Mitchell, B. Selman and H. Levesque. *Hard and Easy Distributions of SAT Problems.* Proceedings of AAAI 1992, 459 - 465.

[54] M. Molloy. *Sets that are connected in two random graphs.* Preprint available at www.cs.toronto.edu/~molloy/papers.html

[55] M. Molloy. *Cores in random hypergraphs and boolean formulas.* Random Structures and Algorithms **27**, 124 - 135 (2005).

[56] M. Molloy. The freezing threshold for k-colourings of a random graph, *Proc. STOC*, 921 - 930, (2012).

[57] M. Molloy and B. Reed. *A critical point for random graphs with a given degree sequence.* Random Structures and Algorithms **6** 161 - 180 (1995).

[58] M. Molloy and B. Reed. *Graph Colouring and the Probabilistic Method.* Springer-Verlag, Berlin (2001).

[59] M. Molloy and R. Restrepo, Frozen variables in random boolean constraint satisfaction problems, *Proc. SODA*, 1306–1318, (2013).

[60] M. Molloy and M. Salavatipour. *The resolution complexity of random constraint satisfaction problems.* SIAM J. Comp. **37**, 895 - 922 (2007).

[61] A. Montanari, R. Restrepo and P. Tetali. *Reconstruction and clustering in random constraint satisfaction problems.* SIAM J. Disc. Math. **25** (2011), 771 - 808.

[62] E. Mossel and Y. Peres. *Information flow on trees.* Ann. Appl. Probab. **13** (2003),817 844.

[63] R. Mulet, A. Pagani, M. Weigt and R. Zecchina. *Coloring random graphs.* Phys. Rev. Lett. **89**, 268701 (2002).

[64] B. Pittel, J. Spencer and N. Wormald, *Sudden emergence of a giant k-core in a random graph.* J.Comb.Th.(B) **67** (1996), 111 - 151.

[65] B. Pittel and N. Wormald, *Counting connected graphs inside-out.* J. Comb.Th. B **93** (2005), 127 - 172.

[66] G. Semerjian, *On the freezing of variables in random constraint satisfaction problems.* J.Stat.Physics **130** (2008), 251 - 293.

[67] A. Sly. *Reconstruction of random colourings.* Commun. Math. Phys. **288** (2009), 943 961.

[68] V.E. Stepanov, *Some features of the structure of a random graph near a critical point.* (Russian) Teor. Veroyatnost. i Primenen. **32** (1987), 633657.

[69] U. Voll, *A novel giant-subgraph phase-transition in sparse random k-partite graphs.* Disc. Appl. Math. **153** (2005), 153 - 181.

[70] D. West, Introduction to Graph Theory. 2nd edition. Prentice Hall (2001).

[71] L. Zdeborová and F. Krzakala. *Phase transitions in the colouring of random graphs.* Phys. Rev. E 76, 031131 (2007)

[72] L. Zdeborová and F. Krzakala. *Quiet planting in the locked constraint satisfaction problems.* SIAM J. Discrete Math. **25** (2011) 750 - 770.

[73] L. Zdeborová. *Statistical physics of hard optimization problems.* Acta Physica Slovaca **59** (2009), 169 - 303.