On the Power of Choice for k-Colorability of Random Graphs

Varsha Dani
Ronin Institute, USA
Dept. of Computer Science, Rochester Institute of Technology, USA

Diksha Gupta
School of Computing, National University of Singapore, Singapore

Thomas P. Hayes
Dept. of Computer Science, University of New Mexico, USA

Abstract

In an r-choice Achlioptas process, random edges are generated r at a time, and an online strategy is used to select one of them for inclusion in a graph. We investigate the problem of whether such a selection strategy can shift the k-colorability transition; that is, the number of edges at which the graph goes from being k-colorable to non-k-colorable.

We show that, for $k \geq 9$, two choices suffice to delay the k-colorability threshold, and that for every $k \geq 2$, six choices suffice.

2012 ACM Subject Classification Mathematics of computing → Stochastic processes; Theory of computation → Generating random combinatorial structures

Keywords and phrases Random graphs, Achlioptas Processes, Phase Transition, Graph Colorability

Digital Object Identifier 10.4230/LIPIcs.APPROX/RANDOM.2021.59

Category RANDOM

Acknowledgements The authors would like to thank Will Perkins for suggesting the problem of shifting the threshold for k-coloring, and Will Perkins and Cris Moore for helpful conversations.

1 Introduction

In studying the evolution of a random graph, a field launched by the seminal paper of Erdős and Rényi [7], one starts from an empty graph, and adds edges one by one, generating each one independently and uniformly at random. In this context, a common object of study is the size of the graph at which some property of interest changes. For instance, if we are interested in k-colorability, there will eventually be some edge whose addition changes the graph from being k-colorable to non-k-colorable.

The k-colorability transition threshold conjecture states that there is a particular threshold $d(k)$ such that, almost surely, the k-colorability transition occurs when G has average degree approximately $d(k)$; more precisely, when the average degree lies between $(1-\varepsilon)d(k)$ and $(1+\varepsilon)d(k)$, for any fixed $\varepsilon > 0$. Substantial progress has been made on pinning down this transition threshold, especially by Achlioptas and Naor [2] and by Coja-Oghlan and Vilenchik [5], culminating in a rather precise formula for the asymptotics of $d(k)$ for large k. However, for fixed $k \geq 3$, the conjecture remains open.

An interesting twist on the evolution of the random graph was proposed by Achlioptas in 2001: Suppose that *two* random edges are sampled at each step in the construction of G, and an online algorithm selects one of them, which is then added to G. A more general version of this process proposes r random edges in each step, from which the algorithm selects one. After m edges have been chosen in this way, how different can the resulting graph be from the usual Erdős-Rényi random graph $G(n, m)$?

Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2021).
Editors: Mary Wootters and Laura Sanità; Article No. 59; pp. 59:1–59:17
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Earlier work on the "power of choice" to affect the outcome of random processes has investigated questions like load-balancing in balls and bins models, scheduling, routing and more; for more, see the excellent survey by Richa, Mitzenmacher and Sitaraman [10]. More specifically, Achlioptas processes have been studied in the context of formation of the giant component in a random graph [3, 4, 13, 11], and the satisfiability threshold for random boolean formulas [12, 6, 9]. In each of these cases, the upshot has been that fairly simple heuristics are capable of shifting the thresholds to a significant extent. However, the heuristics and their analyses remain fairly problem-specific.

The main contribution of the present work is a proof that, for every $k \geq 2$, there exist fairly simple choice strategies that significantly delay the k-coloring threshold, given a constant number of choices for each edge. Our proof leverages existing upper and lower bounds on the k-colorability threshold, and works even if the k-Colorability Threshold Conjecture turns out to be false. More precisely, we establish the following result.

▶ **Theorem 1.** *For every $k \geq 2$, there exist $2 \leq r \leq 6$, an explicit edge selection strategy for the r-choice Achlioptas process, and a real number d such that, if G is the graph produced by running our strategy for $dn/2$ steps, and H is an Erdős-Rényi random graph with the same number of edges, then G is almost surely k-colorable and H is almost surely not k-colorable. In particular, $r = 2$ choices suffice for $k = 2$ and $k \geq 9$, $r = 3$ suffices for all $k \neq 3$, and $r = 6$ suffices for all k.*

If, rather than delaying, one wants to *hasten* the k-colorability threshold, this can be done very easily by "densifying" the graph, an idea used in [6] to hasten the k-SAT threshold for random boolean formulas. Unlike our main result, this technique easily extends to any monotone graph property that has a sharp threshold in the Erdős-Rényi model. More precisely,

▶ **Observation 2.** *Let P be any graph property that is monotone in the sense that $P(G)$ implies $P(G')$ for every subgraph G' of G. Then, if the threshold conjecture is true for P, we can lower the threshold using r choices, whenever $r \geq 2$. Moreover, even without the threshold conjecture for P, if there exist real numbers $0 < \alpha_1 < \alpha_2$ such that P almost surely holds for $G(n, \alpha_1 n)$, and almost surely fails to hold for $G(n, \alpha_2 n)$, then there exists $r = r(\alpha_1, \alpha_2)$ and $d = d(\alpha_1, \alpha_2)$, and an explicit edge selection strategy for the r-choice Achlioptas process, such that, if G is the graph produced by running our strategy for $dn/2$ steps, and H is an Erdős-Rényi random graph with the same number of edges, then H almost surely has property P, and G almost surely does not. In the case when P is k-colorability, $r = 2$ choices suffices to lower the k-coloring threshold when $k = 2$ or $k \geq 12$, $r = 3$ suffices when $k \geq 6$, $r = 4$ suffices when $k \neq 4$, and $r = 5$ suffices for all k.*

The interested reader may refer to Appendix A for additional details.

1.1 Strategy for Delaying the k-Colorability Transition

Our basic strategy for delaying the k-colorability transition is to try to create a large bipartite subgraph. This can be achieved very simply by, *ab initio*, partitioning the vertex set into two equal parts, and then by choosing, whenever possible, a *crossing edge*, that is, one whose endpoints lie in both sides of the partition. As we shall see, this extremely simple heuristic suffices to establish Theorem 1 when $k \geq 6$, and with slight modifications, for $k \leq 5$ as well.

For intuition about why this approach works, think about what happens in the limit as r becomes very large. Since the probability of being offered all non-crossing edges in a particular step is less than 2^{-r}, by choosing crossing edges whenever possible, our graph becomes "more and more bipartite" as r increases. Indeed, when $r = 3 \log n$, G will almost surely become a complete bipartite graph before it is forced to include any non-crossing edges! Obviously, this is a huge delay to any of the k-colorability thresholds, which all take place after linearly many edges.

For more intuition, consider the case when k is very large, but $r \geq 2$ is constant. We expect about a 2^{-r} fraction of the edges to be non-crossing, and hence the average degree of the graph induced by one side of G will be about 2^{-r} times the average degree of G. Since, asymptotically for large k, we know that the k-colorability occurs somewhere around $d \approx 2k \ln(k)$, (See Theorem 3 for a more precise statement.) which is a nearly-linear function, this tells us that each side of G should need almost 2^r times fewer colors than $G(n, m)$. Hence, if we color the two sides with disjoint sets of colors, so that the crossing edges cannot cause any monochromatic edges, we would expect to need almost 2^{r-1} times fewer colors to color our graph than a random graph with the same average degree.

The above approach works as stated for $k \geq 6$. For smaller values of k, it is necessary to improve the above strategy by adding an additional "filtering" step that checks to see whether the edge proposed by the basic strategy would create an obstacle to k-coloring; in this case, we make a different edge choice. This is the most technical part of the paper. particularly the case $k = 3$, for which the filtering algorithm is fairly complicated.

For $k = 4$ and $k = 5$, since we are splitting the colors among the two sides of G, at least one side gets only two colors. This is a bit of a special case because, unlike with more colors, two-coloring does not have a sharp phase transition at a particular average degree. Instead, the transition for $G(n, d/n)$ is spread out over the range $0 < d < 1$. However, as we shall see, for Achlioptas processes, it is possible to delay this threshold until the emergence of a giant component at $d = 1$ (and even beyond!).

For $k = 3$, we need a further modification to our plan as outlined above. With only three colors, one side of the graph would only get one color, and would need to remain empty of edges! Since this is clearly impossible, we modify our plan of prescribing disjoint sets of colors for the two sides of the graph. Instead, we allow one of the three colors to be used on both sides. As will be seen, this complicates both the edge selection process and its analysis, and increases the number of choices we need, to $r = 6$.

We point out an interesting qualitative difference between the problem of delaying the k-coloring threshold and that of delaying the k-SAT threshold. Earlier work on delaying the k-SAT threshold, in particular by Perkins [9] and by Dani et al. [6], took advantage of the fact that, with enough choices, the 2-SAT threshold can be shifted past the k-SAT threshold. The analogous statement for k-coloring would require us to keep our graph bipartite past the formation of a giant component. Although Bohman and Frieze [3] showed that it is possible to delay the formation of a giant component, it obviously cannot be delayed past $d = 2$, and indeed, as shown by Bohman, Frieze and Wormald [4, Theorem 1(d)], not past $d = 1.93$. After a linear-size giant component has formed, each step of our Achlioptas process has a constant probability that all r offered edges will fall within the giant component, and moreover all violate bipartiteness. Thus, there is no hope of keeping the graph 2-colorable past the 3-colorability threshold, for any constant (or indeed sub-logarithmic) number of choices. This "fragility" of the property of 2-colorability may provide some intuition for the increased difficulty of our attempts to shift the k-colorability threshold for small values of k.

1.2 Organization of the Paper

The remainder of the paper is divided into numbered sections. For the most part, each section introduces one or two new ideas that are needed for a particular range of the number of colors, k. Many of the sections depend on concepts introduced in earlier sections, so it is easiest to read them in order.

In Section 2 (Preliminaries), we introduce various notation and terminology, as well as stating the key results from past work that we will need for our work. In Section 3 we formally state the PreferCrossing strategy, and show how it can be used directly to raise the k-colorability threshold for $k \geq 6$. In Section 4, we handle the case $k = 2$ by showing that odd cycles (indeed all cycles) can be delayed until a giant component forms, and that this idea can be combined with previous work on delaying the birth of the giant component. In Section 5, we handle the cases $k = 4$ and $k = 5$. These are treated separately from the large k cases because now one of the two sides will be colored using only two colors, which requires the cycle-avoidance technique developed in Section 4. In Section 6, we handle the hardest case: $k = 3$, which involves a significant extension to the technique for avoiding cycles introduced in Section 4. In Section 7, we show how an improved bound on the 3-coloring transition threshold, due to Achlioptas and Moore [1], can be used to reduce the number of choices we need for $k = 9$ from 3 to 2.

Finally, Appendix A presents a proof of Observation 2, about hastening the transition for (almost) any monotone graph property.

2 Preliminaries

Let V be a fixed vertex set, of size n. In the rest of the paper, unless otherwise specified, whenever we use aymptotic notations such as big-O and little-O, these refer to limits as $n \to \infty$, while all the other key parameters, namely, average degree $\bar{d}$, number of choices, r, and number of colors, k, are held constant. When we state that something happens "almost surely," we mean that the corresponding event has probability $1 - o(1)$.

When we talk about the Erdős-Rényi random graph, $G(n, m)$, we assume that m independent random edges are sampled from $\binom{V}{2}$, with replacement. Edges are undirected and self-loops are not allowed.

In an r-choice Achlioptas process, at each step, r independent random edges are sampled from $\binom{V}{2}$, with replacement. An online algorithm, which we call a "strategy" is used to select one of these edges for inclusion in the edge set of the graph, which is initially empty. We allow duplicate edges both in the set of proposed edges, as well as the graph itself. However, observe that, when the total number of edges is linear in n, and $r = O(1)$, the expected number of duplicate edges seen during the entire process is $O(1)$. Consequently, in this range of parameters, it should be easy to see that very similar results hold even when duplicate edges are not allowed.

Key Results from Prior Work

The following result is due to Achlioptas and Naor [2, See Lemma 3 and Proposition 4]

▶ **Theorem 3** (Achlioptas and Naor). *Suppose k is a positive integer, and $d < 2(k-1)\ln(k-1)$. Then, almost surely, $G(n, dn/2)$ is k-colorable. If, instead, $d > (2k-1)\ln(k)$, then, almost surely, $G(n, dn/2)$ is not k-colorable.*

For notational convenience, we introduce a shorthand for the upper and lower bounds on the transition threshold from Theorem 3.

▶ **Definition 4.** *For k a positive integer, denote*

$$L_k = 2(k-1)\ln(k-1) \quad \textit{and} \quad U_k = (2k-1)\ln(k).$$

Subsequent work by Coja-Oghlan and Vilenchik [5] established an asymptotically sharper bound, pinning down the chromatic number for a set of degrees having asymptotic density one. However, their bounds are only stated asymptotically in k, and do not lead to improved bounds for fixed values of k.

For the case $k = 3$, Achlioptas and Moore [1] proved a tighter lower bound on the 3-colorability threshold by analysing the success probability of a naive 3-coloring algorithm using the differential equations method.

▶ **Theorem 5** (Acliopṭas and Moore). *Almost all graphs with average degree 4.03 are* 3*-colorable.*

Although Theorem 3 is sharp enough to derive most of our bounds, we will need Theorem 5 in order to shift the transition threshold for $k = 7$ using $r = 3$ choices, and $k = 9$ using only $r = 2$ choices. We note that future improvements to the bounds on the k-coloring transition thresholds for $G(n, m)$ might produce further improvements to our bounds.

For the cases whose analysis involve 2-coloring, we will make use of past work on accelerating or delaying the formation of the giant component. We start with a classical result of Erdős and Rényi:

▶ **Theorem 6.** *When $d < 1$, almost surely, all connected components of $G(n, dn/2)$ have size $O(\log n)$, but when $d > 1$, almost surely, $G(n, dn/2)$ has a "giant" component of size $\Theta(n)$.*

Bohman and Frieze [3] showed that, in an Achlioptas process, it is possible to delay this threshold, inspiring many related papers. The following result is due to Spencer and Wormald [13].

▶ **Theorem 7.** *There exists an edge selection strategy for the* 2*-choice Achlioptas process, in which, almost surely, the largest component size is still $O(\log n)$ after the inclusion of $dn/2$ edges, where $d = 1.6587$.*

The details of Spencer and Wormald's elegant algorithm will not be important in the present work. In Section 4 we will show how to modify their strategy to additionally delay G's first cycle until the giant component forms, but these modifications treat the original strategy as a black box. We note that, in the same paper, Spencer and Wormald presented another strategy for hastening the arrival of giant component, causing it to appear at average degree $d = 0.6671$.

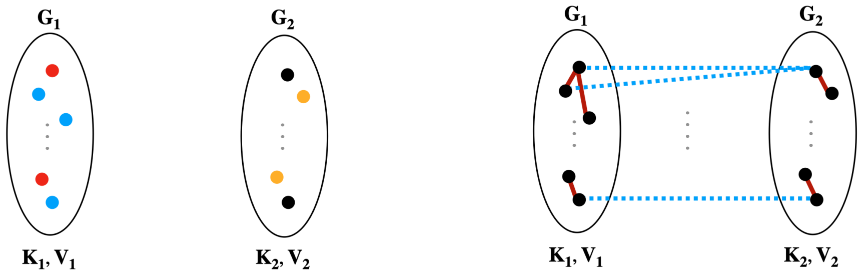
3 Main Idea, Many Colors

Our general approach to delaying the k-colorability threshold is to partition both the vertex and color sets into two parts, and then to assign a disjoint set of colors to each side of the graph. The intuition for this was already discussed in Section 1.1. We now formalize some of these ideas.

Let V be the set of vertices and K the set of colors. Then $|V| = n$ and $|K| = k$. We will partition V into disjoint subsets V_1 and V_2, called "sides," each of size $n/2$. (Since we are interested in the asymptotic behaviour in n we do not need to worry about its parity.)

We also partition K into disjoint sets K_1 and K_2. When we color the graph, we will use colors in K_i to color side V_i. Most of the time we will partition the set of colors so that $|K_1| = \lfloor k/2 \rfloor$ and $|K_2| = \lceil k/2 \rceil$, although we will have some occasions to deviate from this.

We will use an Achlioptas process to build a graph G with m edges on V. G_1 and G_2 will denote the subgraphs of G induced by V_1 and V_2. By abuse of notation, we will also refer to the graphs obtained partway through the Achlioptas process as G, G_1 and G_2.



(a) (b)

Figure 1 Illustration for: (a) Disjoint color sets K_1 and K_2, and vertex sets V_1 and V_2 assigned to G_1 and G_2, respectively; and (b) Types of edges using solid lines for non-crossing edges and dashed lines for crossing edges.

Based on the partition $V = V_1 \bigsqcup V_2$, we classify the possible edges into two types:

- **a crossing edge:** An (undirected) edge $\{u, v\}$ with $u \in V_1$ and $v \in V_2$.
- **a non-crossing edge on side i, or an edge in G_i :** An (undirected) edge $\{u, v\}$ where both $u, v \in V_i$.

Note that since we are using disjoint sets of colors for V_1 and V_2, a crossing edge is never violated by a coloring.
Each edge offered to us in the Achlioptas process is sampled uniformly at random from all $\binom{n}{2}$ pairs of vertices. The probability of a single offered edge being a crossing edge is

$$\frac{(n/2)(n/2)}{n(n-1)/2} = \frac{1}{2} + \frac{1}{2(n-1)} = 1/2 + o(1) \approx 1/2$$

while for $i = 1, 2$, the probability of a single offered edge being a non-crossing edge on side i is

$$\frac{\frac{1}{2}(n/2)((n/2)-1)}{n(n-1)/2} = \frac{1}{4} - \frac{1}{4(n-1)} = 1/4 - o(1) \approx 1/4$$

Let r denote the number of edges offered to the algorithm at each step of the Achlioptas process. As a reminder, each edge is sampled independently and uniformly from $\binom{V}{2}$. We use the following strategy to select an edge at every step, unless stated otherwise:

> ▶ **Strategy 1. PreferCrossing**
> *Select the first crossing edge, if any. Otherwise, select the first edge.*

Note that in the event that no crossing is available, the selected edge is equally likely to be on either side, and is a uniformly random edge conditioned on being on the side it is.

Let m be the total number of edges inserted into G, so the average degree of G is $\bar{d} = 2m/n$. For $i \in \{1, 2\}$, let $\bar{d}_i$ denote the average degree of the graph G_i.

We use the PreferCrossing strategy to choose the edge to be inserted into G at each step. A non-crossing edge is inserted only if all r candidate edges are non-crossing, so the probability of inserting a non-crossing edge is at most $1/2^r$. Also, in this case we insert the first edge, which is equally likely to be on either side. So the probability of inserting an edge into G_i is $1/2^{r+1}$.

It follows that in expectation, there are $m/2^{r+1}$ edges in each G_i and the rest are crossing edges. Using this we can calculate the expected average degrees on the two sides as follows:

$$\mathbb{E}[\bar{d}_i] < \frac{2m/2^{r+1}}{n/2} = \left(\frac{2m}{n}\right)\frac{1}{2^r} = \frac{\bar{d}}{2^r}$$

By the Law of Large Numbers, it follows that, almost surely,

$$\bar{d}_i < (1 + o(1))\frac{\bar{d}}{2^r}. \tag{1}$$

Now, since whichever of the three classes of edge (crossing, non-crossing on side 1, non-crossing on side 2) the PreferCrossing strategy selects, the edge is uniformly random within that class, it follows that, conditioned on $\bar{d}_1$ and $\bar{d}_2$, G_1 and G_2 are uniformly random graphs with that number of edges. Therefore, assuming each $\bar{d}_i$ is below a known lower bound on the k_i-colorability transition, it will follow that each G_i is almost surely k_i-colorable, and hence G is $(k_1 + k_2)$-colorable. If, additionally, $\bar{d}$ is greater than a known upper bound on the k-colorability threshold, and $k = k_1 + k_2$, we will have shifted the k-colorability transition threshold.

Theorem 3 tells us that for $\kappa \geq 3$, the κ-colorability transition threshold (if it exists) lies between L_κ and U_κ (see Definition 4.) Additionally, we will sometimes also use the improved lower bound $L'_3 = 4.03$ from Theorem 5

Since the expression for L_κ is monotone, the graphs G_i are k_i-colorable (and hence G is k-colorable) until $\bar{d}_1 = \bar{d}_2 = \min\{L_{k_1}, L_{k_2}\}$. It therefore makes sense to split the colors as evenly as possible. We will set $k_1 = \lfloor k/2 \rfloor$, $k_2 = \lceil k/2 \rceil$. Then G_1 and G_2 are k_1- and k_2-colorable respectively until $\bar{d}_1 = \bar{d}_2 = L_{\lfloor k/2 \rfloor}$

Now, we know from Eq. (1) that

$$\bar{d} \geq 2^r \bar{d}_1 \geq 2^r L_{\lfloor k/2 \rfloor}$$

and we will have delayed the k-colorability transition if this exceeds U_k

Since L_k and U_k are both asymptotically equal to $2k \ln k$, this shows that for sufficiently large k two choices suffice to raise the k-colorability threshold. Indeed, using Mathematica to solve the inequalities

$$2^r L_{\lfloor k/2 \rfloor} \geq U_k$$

for $r = 2, 3$ and 4, we see that

- two choices suffice for even $k \geq 10$ and odd $k \geq 13$
- three choice suffice for even $k \geq 6$ and odd $k \geq 9$ and
- four choices suffice for $k = 7$.

Moreover, if we use the improved lower bound $L'_3 = 4.03$, instead of L_3 for the case of $k = 7$, then we see that

$$8L'_3 = 8 \times 4.03 = 32.24 > 25.3 = U_7$$

so that three choices suffice $k = 7$. This establishes Theorem 1 for $k \geq 6$, except for the cases $k = 9$ and $k = 11$.

For $k = 9, 11$ we have established that three choices suffice, but we want to show that in fact we only need two. We will tackle the case $k = 11$ here and leave $k = 9$ for Section 7.

When $k = 11$, we allocate five colors to side 1, and six colors to side 2. The five-colorability of G_1 is only guaranteed until $\bar{d}_1 = 8 \ln 4 \approx 11.09$. With $r = 2$ choices, at this point $\bar{d}$ is about 44.36, smaller than $20 \ln 10 = 46.05$, so that although G is 11-colorable, so is

$G(n, m = 44.36n/2)$, so we have not shifted the threshold. In order to increase $\bar{d}$ past $U_{11} = 21 \ln 11 = 50.356$, we note that $\bar{d}_2$ is *also* about 11.09, since we are equally like to add a non-crossing edge to side 2 as to side 1. But $\bar{d}_2$ is allowed to go to $10 \ln 5 \approx 16.09$ before we can no longer guarantee the 6-colorability of G_2. This means we have a fair bit of slack to favor G_2 when adding non-crossing edges. Suppose we put a $\varphi < 1/2$ fraction of the non-crossing edges into G_1 and a $(1 - \varphi)$ fraction of them into G_2. What should φ be to ensure the best outcome? Note that we need $\varphi \geq 2^{-r}$, since if all the non-crossing choices are on side 1, then we cannot add an edge in side 2. However, subject to this constraint, we are adding $m\varphi/2^r$ edges to G_1 and $m(1-\varphi)/2^r$ edges to G_2 in expectation. But this means that $\mathbb{E}[\bar{d}_1] = \bar{d}\varphi/2^{r-1}$ and $\mathbb{E}[\bar{d}_2] = \bar{d}(1-\varphi)/2^{r-1}$. Since these random variables stay close to their expectations, it follows that $\bar{d}_1$ and $\bar{d}_2$ are in the ratio $\varphi/(1-\varphi)$. Now, its is best if we can arrange it so that both G_1 and G_2 lose their guarantee of colorability at the same time (so that there is no slack). But this means

$$\frac{L_5}{L_6} = \frac{\varphi}{1-\varphi}$$

But this means we should set

$$\varphi = \frac{L_5}{L_5 + L_6} = \frac{11.09}{11.09 + 16.09} \approx \frac{2}{5}$$

Since $2/5 > 1/4$, it *is* possible to achieve a $2/5 - 3/5$ split of the non-crossing edges, when there are two choices.

Finally, what does this make the average degree of the graph G at the time when 11-colorability can no longer be guaranteed?? Since

$$\bar{d} \approx \frac{2^{r-1}\bar{d}_1}{\varphi} \approx \frac{2^{r-1}\bar{d}_2}{1-\varphi}$$

when $r = 2$ and $\varphi = 2/5$ we get

$$\bar{d} \approx \frac{2L_5}{2/5} = 55.45 > 50.356 = U_{11}.$$

Thus two choices suffice to raise the 11-colorability threshold.

To write down an *explicit* edge selection strategy, note that if when we are not forced to take an edge on a particular side, we toss a biased coin that selects side 1 with probability γ, then the overall probability of adding an edge to side 1 conditioned on adding a non-crossing edge is $1/4 + \gamma/2$. Since we want this to be $2/5$ we should set $\gamma = 3/10$. Here is the strategy we use.

> ▶ **Strategy 2. BiasedPreferCrossing for $k = 11$**
> *Given two edges, select the first crossing edge, if any.*
> *Otherwise if both non crossing edges are on the same side, select the first one*
> *Otherwise there is one edge offered on each side. Select the one on side 1 with probability 0.3, and the one on side 2 with probability 0.7.*

4 Emergence of Giant component and Emergence of Cycles

The case $k = 2$ differs from larger k in one very important way: namely, the k-Colorability Threshold Conjecture is false when $k = 2$; for $G(n, p)$ where $p = d/n$, rather than a sharp transition from colorable to non-colorable at a critical value of d, instead this transition is spread across the whole range $0 < d < 1$.

To see this, we observe that the expected number of triangles is $\binom{n}{3}p^3 \approx d^3/6$, which is a positive constant for all $0 < d < 1$. It is not much harder to prove that the probability that at least one triangle exists is also $\Theta(1)$ whenever $d = \Theta(1)$, and hence the probability that $G(n,p)$ is not 2-colorable is bounded away from zero.

On the other hand, it is also not hard to prove that, as long as $p < (1-\epsilon)/n$, $G(n,p)$ is a forest with probability bounded below by a constant, and hence the probability that $G(n,p)$ is 2-colorable is also bounded away from zero. In other words, the transition from $G(n,d/n)$ being almost surely 2-colorable to being almost surely not 2-colorable is not sharp, but is rather spread over the entire interval $0 < d < 1$.

Even though there isn't a sharp threshold for 2-colorability in $G(n,p)$, we will prove in this section that, given $r = 2$ choices, we can both create a sharp threshold, and shift it.

Two-colorability is of course, equivalent to the absence of odd cycles, and it turns out that the presence of odd cycles–indeed, of any cycles–is intimately linked with the emergence of the giant component.

Consider a 2-choice Achlioptas process, using the following, very simple, edge selection rule:

> ▶ **Strategy 3. SimpleAvoidCycles**
> *Select the first edge, unless it would create a cycle, in which case, select the second edge.*

SimpleAvoidCycles manages to avoid the emergence of cycles until the average degree is 1, the threshold for the emergence of the giant component. On the other hand, once a giant component forms, it very quickly grows to size $\omega(\sqrt{n})$, at which point it is almost certain that a pair of edges will be offered within $o(n)$ steps, both of which lie within the giant component. Therefore it is not possible to avoid cycles for more than a few steps after the formation of a giant component. Thus, with two choices, this very simple heuristic results in a sharp threshold for the emergence of cycles (and similarly for odd cycles, a.k.a. non-2-colorability).

4.1 Analysis of SimpleAvoidCycles

As before, let $m = dn/2$, where $d < 1$. Consider the graph $G' = G(n,m')$, where $m' = m + \log n$.

For our purposes, $G(n,m')$ means the graph obtained from sampling m' independent edges uniformly from $\binom{n}{2}$ (with replacement).

▶ **Lemma 8.** *The number of edges of G' contained in one or more cycles is $o(\log n)$, almost surely.*

Proof. This is a standard result, so we present an abbreviated proof. The expected number of cycles of length k in the $G(n,p)$ model is

$$\binom{n}{k}\frac{k!}{2k}p^k < \frac{(np)^k}{2k}.$$

Since each k-cycle contains k edges, it follows that the expected number of edges in k-cycles is less than $(np)^k/2$. If we set $np = 1-\epsilon$ and sum over all $k \geq 3$, we get

$$\sum_{k=3}^{n}\frac{(1-\epsilon)^k}{2} < \sum_{k=3}^{\infty}\frac{(1-\epsilon)^k}{2} = \frac{(1-\epsilon)^3}{2\epsilon} = O(1).$$

We omit the details of the comparison between the $G(n, m)$ model and the $G(n, p)$ model, which are standard. Since the expected number of edges in cycles is $O(1)$, whereas $\log(n)$ tends to infinity, by Markov's inequality it is almost certain that the actual number of edges in cycles is $o(\log n)$. ◀

▶ **Lemma 9.** *The probability that any of the edges $e_{m+1}, \ldots e_{m'}$ are contained in a cycle of G' is $O(\log(n)/n)$.*

Proof. Since, by Lemma 8, the expected number of edges in cycles is $O(1)$, and since the m' edges of G' are identically distributed, it follows that each edge e_j has probability $O(1/m')$ to be part of a cycle. Hence, by linearity of expectation and Markov's inequality, the probability that any of the edges $e_{m+1}, \ldots, e_{m'}$ is part of a cycle is $O((m' - m)/m') = O(\log(n)/n)$. ◀

▶ **Theorem 10.** *For $d < 1$, SimpleAvoidCycles outputs a cycle-free graph, almost surely.*

Proof. We couple the m choices made by SimpleAvoidCycles with the edges chosen in $G(n, m')$. For each $1 \le i \le m$, let e_i be the first edge offered to SimpleAvoidCycles. For each j'th edge rejected by SimpleAvoidCycles, we let e_{m+j} be the second edge offered to SimpleAvoidCycles. When j is greater than the number of edges rejected by SimpleAvoidCycles, we let e_{m+j} be a uniformly random edge, chosen independently from all others.

Our first observation is that the sequence of edges $e_1, \ldots, e_{m'}$ is uniformly random in $\binom{n}{2}^{m'}$. This is because each e_j is uniformly random, conditioned on $e_1, \ldots, e_{j-1}$.

Now, suppose the output of SimpleAvoidCycles contains a cycle. This means that at least one of the "second edges" chosen by SimpleAvoidCycles is contained in a cycle in the output of SimpleAvoidCycles. This implies that either SimpleAvoidCycles rejected more than $m' - m$ first edges, in which case $e_1, \ldots, e_m$ contains more than $m' - m$ cycles, and hence so does G'. This is unlikely by Lemma 8. Or SimpleAvoidCycles rejected fewer than $m' - m$ edges, but one of the second edges formed a cycle in its output, which is a subgraph of G'. But Lemma 9 bounds the probability of this event. Applying the union bound to these two events, we get the desired upper bound on the probability that the output of SimpleAvoidCycles contains a cycle. ◀

4.2 Avoiding Cycles Longer

Next we will show how to keep G a forest as long as the average degree is less than 1.6587, the threshold from Theorem 7. More generically, we will show how, if any strategy for a 2-choice Achlioptas process can delay the giant component until average degree d, we can tweak it to additionally keep G a forest up to the same average degree threshold. We will refer to this strategy as DelayGiant. To be more precise, we will assume that, for every $d' < d$, DelayGiant run for $d'n/2$ steps almost surely outputs a graph whose components all have $O(n^{1/4})$ vertices.

First, we argue that, without loss of generality, DelayGiant can be assumed to have the following two properties:

1. If exactly one of the two offered edges make a cycle, DelayGiant selects it.
2. In this case, the subsequent behavior of DelayGiant is independent of the second, unselected edge.

The first property is obvious, since if an edge forms a cycle, adding it to G does not increase any of the component sizes; therefore it dominates any edge that doesn't form a cycle. The second property is less obvious, but the idea is that any strategy can be made "forgetful" by making it resample any state information it might be maintaining, from its conditional distribution, conditioned on the edges it has accepted so far. It follows from the Law of Total Probability that this does not change the distribution of the output. An algorithm that is forgetful in this sense, and satisfies property 1, necessarily satisfies property 2 as well. The motivation for property 2 is that it will allow us to apply the Principle of Deferred Decisions to the edges chosen by our strategy in steps when it deviates from DelayGiant's choices.

Now our strategy for delaying the appearance of the first cycle in G can be described in one sentence:

▶ **Strategy 4. AvoidCycles**
Select the edge chosen by the DelayGiant algorithm, unless it would form a cycle, in which case, select the other edge.

▶ **Theorem 11.** *For $d < 1.6587$, with high probability, the 2-choice Achlioptas process run for $m = dn/2$ steps using strategy AvoidCycles outputs a cycle-free graph.*

Consider a run of the DelayGiant algorithm. Let $\{(e_1, e'_1), (e_2, e'_2) \dots (e_m, e'_m)\}$ be the edges that are offered to the algorithm during this run. Let G_i be the graph produced by DelayGiant after the first i steps, *i.e.* G_i has i edges, one out of each pair $(e_j, e'_j), 1 \le j \le i$. Let

$$S := \{i | \text{ neither of the edges } e_i, e'_i \text{ forms a cycle when added to } G_{i-1}\}$$

Let DelayGiant′ be an algorithm that emulates DelayGiant on the steps in S, but adds no edge on the $m - |S|$ steps when DelayGiant would add a cycle-forming edge. Let G'_i be the intermediate graph produced by DelayGiant′ after i steps. Note that for all i, G'_i is a spanning forest of G_i.

By assumption, almost surely, all the components of G_m have size $o(n^{1/4})$. Hence also, for all $1 \le i \le m$, the components of G_i, and therefore also G'_i have size $O(n^{1/4})$. Now, consider an arbitrary forest all of whose components are of size at most t. We make two observations:

▶ **Observation 12.** *Let G be a graph, all of whose components are of size at most t. Then the probability that adding one random edge to G creates a cycle is at most $\frac{t-1}{n-1}$.*

▶ **Observation 13.** *Let G be a graph, all of whose components are of size at most t. Add any ℓ edges to G. Then, the largest component of the resulting graph has size at most ℓt*

Applying Observation 12 inductively to each G'_i, with $t = O(n^{1/4})$, we see that the expected number of steps on which DelayGiant′ adds no edge, $\mathbb{E}[m - |S|]$, is at most $m\left(\frac{t-1}{n-1}\right)$, which is $O(n^{1/4})$.

When DelayGiant′ has run for m steps, the resulting graph G'_m is a forest with $|S|$ edges, whose components are size $O(n^{1/4})$. Let DelayGiant″ be the algorithm that runs DelayGiant′ and then expands G'_m to a graph with m edges by adding $m - |S| = O(n^{1/4})$ uniformly random edges. Applying Observation 13, the components of this graph have size at most $O(n^{1/2})$. Since each of the $O(n^{1/4})$ random edges to be added has at most $O(n^{-1/2})$ chance of forming a cycle, by Markov's inequality, the probability that this graph contains a cycle is at most $O(n^{-1/4})$. Thus, the graph produced by DelayGiant″ is almost surely a forest.

The proof of Theorem 11 will be complete once we establish the following Lemma, relating AvoidCycles to DelayGiant″.

▶ **Lemma 14.** *AvoidCycles is better at avoiding cycles than DelayGiant$''$*, i.e., *for every m,*

$$\mathbb{P}(\textit{AvoidCycles is cycle-free after } m \textit{ edges}) \geq \mathbb{P}(\textit{DelayGiant}'' \textit{ is cycle-free after } m \textit{ edges}).$$

Proof. It will suffice to couple the choices made by the two algorithms in such a way that each edge chosen by DelayGiant$''$ is either the same as the one chosen by AvoidCycles, or forms a cycle. Consider the edge chosen by AvoidCycles at a particular timestep $i \in [m] \setminus S$. Also, let A_i denote the set of all possible edges that would form a cycle if added to G_{i-1}, and let $B_i = \binom{n}{2} \setminus A_i$. We apply the principle of deferred decisions to the edges (e_i, e_i'). Conditioned on G_{i-1} and the event that $i \notin S$, the distribution of (e_i, e_i') is uniform in $(A_i \cup B_i)^2 \setminus B_i^2$. This means that the edge selected by AvoidCycles in step i has a conditional distribution which is uniform in A_i with probability $\frac{|A_i|}{|A_i|+2|B_i|}$ and uniform in B_i with probability $\frac{2|B_i|}{|A_i|+2|B_i|}$.

Let us compare this distribution with that of a uniformly random edge. A uniformly random edge is uniform in A_i with probability $\frac{|A_i|}{|A_i|+|B_i|}$, and uniform in B_i with probability $\frac{|B_i|}{|A_i|+|B_i|}$. Now, observing that

$$\frac{a}{a+2b} < \frac{a}{a+b}$$

whenever $a, b > 0$, we see that the edge selected by AvoidCycles can be coupled with the uniformly random edge so that either the two edges are either equal, or the edge selected by AvoidCycles is in B_i and the uniformly random edge is in A_i. Since an edge in A_i would have formed a cycle even at step i, it definitely forms a cycle when added to the final result of DelayGiant.

Moreover, conditioned on the edges $e_1, \ldots, e_m$, the deferred edges e_{m+j} are fully independent, since Property 2 tells us that DelayGiant does not take the identities of previously rejected edges into account when making its decisions. Thus, the sequence of edges $e_{m+1}, \ldots, e_{m'}$ is less likely to make a cycle than a sequence of $m' - m$ uniformly random edges. It follows that there is a coupling between the output of AvoidCycles and DelayGiant$''$ such that the graphs produced are always identical except when DelayGiant$''$ contains at least one cycle. ◀

5 Four or Five Colors

When we get down to fewer than six colors, the basic PreferCrossing strategy runs into some difficulties, since at least one of the sides has fewer than three colors. This is problematic because even at low edge densities, $G(n, m)$ has a constant chance of having an odd cycle and therefore cannot be two-colored. This means that the subgraph G_i of G on the side with only two colors will stop being two-colorable even before it has a linear number of edges. Fortunately, as we saw in the previous section, given a choice of two edges to choose from, we can can avoid the appearance of cycles and keep the graph two-colorable until it reaches an average degree of about 1.6587.

When $k = 4$, we partition V into two sides as usual, and assign two of the four colors to each side. We prefer crossing edges as usual, and select a crossing edge whenever we are offered one. If there at least three edges to choose from, and we are not offered any crossing edges, then at least two of the offered non-crossing edges are on the same side, and we have some room to be selective about the edge we are adding, and avoid cycles in the graph. Note that either side is equally likely to have two or more edges, and conditioned on the side, the edge choices are uniformly random from that side.

Here is an explicit description of the edge-selection strategy used:

▶ **Strategy 5. PreferCrossing with Two-sided Cycle Avoidance (PCTCA)**
Choose $r = 3$ edges independently and uniformly at random
if *there are any crossing edges* **then**
| *Select the first crossing edge.*
else
| *Let G_i be the side with more candidate edges*
| *Select the edge chosen by AvoidCycles on G_i.*
end

Using the above edge selection strategy, we can show that

▶ **Theorem 15.** *Three choices suffice to increase the 4-colorability threshold.*

Proof. Let m be the total number of edges inserted into G, so the average degree of G is $\bar{d} = 2m/n$. For $i \in \{1, 2\}$, let $\bar{d}_i$ denote the average degree of the graph G_i.

The probability of inserting a crossing edge into G is $7/8$. When there are no crossing edges, the chance that a particular side has two edge choices is $1/16$. We choose one of the two or more offered edges using the AvoidCycles strategy so that for $i \in \{1, 2\}$ the expected number of edges inserted into G_i is $m/16$. Thus $\mathbb{E}[\bar{d}_i] = \frac{2m/16}{n/2} = \frac{\bar{d}}{8}$, and as usual,

$$\bar{d}_i \leq (1 + o(1))\bar{d}/8$$

Since we are using the AvoidCycles strategy to insert edges into G_1 and G_2, by Theorem 11 we can push $\bar{d}_1$ to 1.6587 before G_i stops being two-colorable. At that point,

$$\bar{d} = 8 \times 1.6587 = 13.2696 > 9.704 = 7 \ln 4 = U_4$$

so that G is 4-colorable at a density where $G(n, m)$ isn't, and we have shifted the threshold. ◀

When $k = 5$ we assign two colors to G_1 and three colors to G_2. Again, we choose crossing edges whenever we can; if there are $r = 3$ choices we can do this about $7/8$ of the time.

What happens when we can't choose a crossing edge? Half the time, there will be two edges offered on side 1 and we can use AvoidCycles to choose one of them. If we choose an edge on side 2 the other half the time, the we will have $\bar{d}_1 = \bar{d}_2 = \bar{d}/8$ and as we know from the four-colorability analysis above, we can push this up to $\bar{d}_1 = 1.6587$ and $\bar{d} = 13.2696$ before the 2-coloring on G_1 breaks down. But $13.2696 < 14.485 = 9 \ln 5 = U_5$ so we haven't shifted the 5-colorability threshold. Of course, at this point, $\bar{d}_2$ is also only 1.6587, and has a lot of slack before it reaches $L'_3 = 4.03$, or even $L_3 = 2.77$.

So we want to use a biased strategy that favors choosing edges from side 2 when no crossing edges are available. We could figure out the optimal bias that makes both sides reach their limits at the same time, as we did in the $k = 11$ case. Instead we opt for the following simple explicit strategy.

▶ **Strategy 6. PreferCrossing with One-sided Cycle Avoidance (PCOCA)**
Choose $r = 3$ edges independently and uniformly at random
if *there are any crossing edges* **then**
  Select the first crossing edge.
else
  if *the first two edges are both in V_1^2* **then**
    Select one of them according to AvoidCycles, run on G_1
  else
    (In this case at least one edge is in V_2^2)
    Select the first edge in V_2^2.
  end
end

▶ **Theorem 16.** *Three choices suffice to increase the 5-colorability threshold.*

Proof. Let m be the total number of edges inserted into G so the average degree of G is $\bar{d} = 2m/n$. Similarly, for $i \in \{1, 2\}$, let $\bar{d}_i$ denote the average degree of the graph G_i.

The probability of choosing a crossing edge is $7/8$. The probability of choosing an edge on side 1 is $(1/4)(1/4)(1/2) = 1/32$, and the probability of choosing an edge on side 2 is $3/32$. Then $\mathbb{E}[\bar{d}_1] = \bar{d}/16$ and $\mathbb{E}[\bar{d}_2] = 3\bar{d}/16$

If we set $m = 8n$ then $\bar{d} = 16 > U_5$ is a density at which $G(n, m)$ is not 5-colorable. On the other hand if $\bar{d} = 16$ in G constructed using PCOCA, then $\bar{d}_1 = 1 < 1.6586$ and $\bar{d}_2 = 3 < 4.03 = L'_3$, so that G_1 is two-colorable, G_2 is 3-colorable and hence G is 5-colorable. ◀

6 Three Colors

For $k = 3$, we face a new challenge to our approach, namely: there is no longer any hope of using disjoint color sets to color the two sides of our graph. Instead, we try to make the color sets as disjoint as possible. Specifically, we try to color G using red and yellow for the first side, and blue and yellow for the second side. Although the crossing edges may cause problems now, at least the only bad color assignment for a crossing edge is (yellow, yellow). We call this kind of 3-coloring a $(Y, *)$-coloring, since the non-yellow colors are determined by their side.

Note that this specific type of coloring can be found in linear time, since it is a special case of Constrained Graph 3-Coloring, which is reducible to 2-SAT (see [8, Problem 5.6]). Here is our strategy:

▶ **Strategy 7. PreferCrossingButCheck (PCBC)**
Choose $r = 6$ edges independently and uniformly at random.
Let e be the edge chosen by the PreferCrossing heuristic.
Check whether $G \cup \{e\}$ remains (Y,)-colorable. If it is, select e. Otherwise, select the first edge other than e.*

We note that with an appropriate data structure, all m of the colorability checks can be performed in combined expected time $O(n)$. However, since our goal is just to show that the colorability transition can be shifted, we leave the details as an exercise.

We claim that, when $r = 6$, the output of PCBC is almost surely $(Y, *)$-colorable. To see this, observe that, in order for a greedy approach to coloring to fail, the graph must have a cycle of length $2k+1$ with edges (in order) $(e_1, e_2, \ldots, e_{2k+1})$, where the k even edges e_{2i} are all non-crossing. This is analogous to the fact that a graph fails to be 2-colorable if and only if it has an odd cycle. However, note that in the case of 2-coloring, the criterion is "if and only if," whereas here there is only an implication; the cycle is only guaranteed to cause a problem if we start by coloring the wrong vertex yellow. We call a cycle of this type a "bad odd cycle."

▶ **Proposition 17.** *Let $d > 0$. Let G be the output of an Achlioptas process with r choices, running the PreferCrossing heuristic, for $dn/2$ steps. Also suppose that $d^2 < 2^r$. Then the expected number of edges contained in bad odd cycles of G is $O(1)$.*

Proof. Note that, for every vertex v, the expected degree is d, but the expected number of non-crossing edges incident with v is $d2^{-r}$. With a little work we can see that the expected number of walks of length $2k$ starting at a particular node, in which all the even steps are along non-crossing edges is at most $d^k(d2^{-r})^k$. In order to complete such a walk to a cycle of length $2k+1$, we need a particular edge to be present, which is an event of probability at most $d/(n/2)$. Since there are n possible starting points for our walk, this gives the following bound on the number of edges contained in a bad odd cycle:

$$n \sum_{k \geq 1} \frac{d}{n/2} (d^2 2^{-r})^k (2k+1) = 2d \sum \sum_{k \geq 1} (2k+1)(d^2 2^{-r})^k,$$

which since $d^2 < 2^r$, is a convergent sum. ◀

Thus, in expectation, PCBC deviates from the choices made by PreferCrossing on only $O(1)$ steps. Denote this number of steps by $m' - m$. On the steps when it deviates, it takes the first alternative edge. Since PreferCrossing makes its edge choice based only on which edges are crossing or not, this alternative edge must be uniformly random, conditioned on whether it is a crossing edge or not. It follows that PCBC succeeds at least as often as a variant PCBC$'$ that, instead of taking each rejected edge from PreferCrossing, instead adds one uniformly random crossing edge and one uniformly random non-crossing edge.

PCBC$'$, in turn, will almost surely perform at least as well as another variant, PCBC$''$, which, instead of adding one uniformly random crossing edge, and one uniformly random non-crossing edge, instead adds $C2^r$ edges chosen by an Achlioptas process running the PreferCrossings strategy, where $C \to \infty$. But now, observe that PCBC$''$ is just PreferCrossings run for $m'' = m + o(n)$ steps, with all of its bad odd cycles from the first m steps broken up. Since PreferCrossings run for m'' steps still has, in expectation, $O(1)$ edges involved in bad odd cycles, and these edges are uniformly randomly distributed among the m' steps, the probability that any of them occur in the last $m'' - m$ steps is $O((m'' - m)/m'') = o(1)$. Hence the output of PCBC$''$ almost surely has no bad odd cycles, and is therefore $(Y, *)$-colorable. Since by our earlier remarks, PCBC almost surely performs at least as well as PCBC$''$, this establishes the result.

7 Two choices for 9 colors

In Section 3 as part of a unified analysis for $k \geq 6$ we showed that three choices were enough to raise the 9-colorability threshold. In this section we will show that in fact just two choices suffice. Surprisingly, this result involves a more uneven split of the colors, with three colors reserved for V_1 and six for V_2. This helps partly because, for $k = 3$, the improved lower bound $L'_3 = 4.03$ of Theorem 5 is significantly better than the bound of Theorem 3.

The main idea is to use a biased PreferCrossing strategy which favors the six color side when a non-crossing edge is forced. We have two choices, so we will be putting in a non-crossing edge only a fourth of the time. Conditioned in that, we want to make the colorabilty on the two sides break at roughly the same time. As we saw before, this means that we should add edges to side 1 (with three colors) with probability φ, where

$$\frac{\varphi}{1-\varphi} = \frac{L'_3}{L_6} = \frac{4.03}{16.094} \approx \frac{1}{4}$$

from which we get that φ should be approximately 1/5... and that is a problem. If the probability of selecting an edge from side 1 conditioned on a non-crossing edge is 1/5, then the overall probability is 1/20, but this is not achievable with two choices, since there is a 1/16 chance that both edges are on side 1!

So where does that leave us? It turns out that we can still tweak this to make it work. From the beginning, we have made the *a priori* division of the vertex set into two equal sized disjoint subsets because that maximizes our ability to put in crossing edges. But having found ourselves in a situation where we want to put in fewer edges into side one than is possible, the obvious solution seems to be to make side one smaller. So let's start over, and partition V into disjoint sets V_1 and V_2, where $|V_1| = \alpha n$ and $|V_2| = (1-\alpha)n$. It turns out that $\alpha = 0.47$ works well. With this parameter setting, we choose crossing edges whenever possible, and failing that, edges in G_2, with edges in G_1 as a last resort. This leads to average degrees $\bar{d}_1 = 0.1038\bar{d}$ and $\bar{d}_2 = 0.3830\bar{d}$. Since $0.1038U_9 \le L_3$ and $0.3830U_9 \le L_6$, this shows that we have shifted the 9-coloring threshold with $r = 2$ choices.

References

1. Dimitris Achlioptas and Cristopher Moore. Almost all graphs with average degree 4 are 3-colorable. *Journal of Computer and System Sciences*, 67(2):441–471, 2003.
2. Dimitris Achlioptas and Assaf Naor. The two possible values of the chromatic number of a random graph. *Annals of mathematics*, 162(3):1335–1351, 2005.
3. Tom Bohman and Alan Frieze. Avoiding a giant component. *Random Structures & Algorithms*, 19(1):75–85, 2001.
4. Tom Bohman, Alan Frieze, and Nicholas C Wormald. Avoidance of a giant component in half the edge set of a random graph. *Random Structures & Algorithms*, 25(4):432–449, 2004.
5. Amin Coja-Oghlan and Dan Vilenchik. Chasing the k-colorability threshold. In *2013 IEEE 54th Annual Symposium on Foundations of Computer Science*, pages 380–389. IEEE, 2013.
6. Varsha Dani, Josep Diaz, Thomas Hayes, and Cristopher Moore. The power of choice for random satisfiability. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 484–496. Springer, 2013.
7. Paul Erdős and Alfréd Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5(1):17–60, 1960.
8. Cristopher Moore and Stephan Mertens. *The nature of computation*. OUP Oxford, 2011.
9. Will Perkins. Random k-SAT and the power of two choices. *Random Structures & Algorithms*, 47(1):163–173, 2015.
10. Andrea W Richa, M Mitzenmacher, and R Sitaraman. The power of two random choices: A survey of techniques and results. *Combinatorial Optimization*, 9:255–304, 2001.
11. Oliver Riordan and Lutz Warnke. Achlioptas process phase transitions are continuous. *The Annals of Applied Probability*, 22(4):1450–1464, 2012.
12. Alistair Sinclair and Dan Vilenchik. Delaying satisfiability for random 2-SAT. *Random Structures & Algorithms*, 43(2):251–263, 2013.

13 Joel Spencer and Nicholas Wormald. Birth control for giants. *Combinatorica*, 27(5):587–628, 2007.

Appendix A: Hastening the threshold

Here we present a sketch of the proof of Observation 2. Since the choice strategy and the proof technique are exactly the same as in [6], we omit most of the details.

Proof Sketch for Observation 2. The choice strategy is to favor some vertex set S, where $|S| = \gamma n$. For instance, let $S = \{1, 2, \ldots \gamma n\}$. By always choosing a random edge in $\binom{S}{2}$ when one is available, we find that the induced graph on S is uniformly random, but denser than G as a whole, having average degree asymptotically equal to $(1-(1-\gamma^2)^r)/\gamma$ times the average degree of G. Choosing γ to maximize this expression, we obtain the desired choice strategy. For instance, setting $\gamma = 1/\sqrt{r}$, we can see that $(1-(1-\gamma^2)^r)/\gamma = \Theta(\sqrt{r})$, which tends to infinity. This shows that the favored subgraph can be made arbitrarily more dense than G, thus bridging the gap between any upper and lower bounds on the threshold. ◀