To Sharon and to my parents

Contents

1 Introduction **1**

1.1 Goals . . . 2

1.2 Notation and definitions . . . 3

1.3 The model . . . 5

1.3.1 Some properties of these graphs . . . 6

1.4 Results . . . 6

1.5 Related work . . . 8

2 One shot coloring propagation algorithm **9**

2.1 A proof of Theorem 2.1 for constant d . . . 10

2.2 A proof of Theorem 2.1 for $d = c(\log\log n)$, where $c < \frac{2}{\log(4/3)}$ is a constant. . . . 12

2.2.1 Sphere caps cover . . . 13

2.2.2 Geometric proof . . . 14

2.2.3 Algorithm 1 does not color $G_{n,k,d}$ when $d(n)$ is too large . . . 16

3 2-Neighborhood Algorithm **16**

3.1 Proof plan for Theorem 3.1 . . . 17

3.2 Proof of Lemma 3.2 . . . 18

3.3 Proof of Lemma 3.3 . . . 19

3.4 Proof of Lemma 3.4 . . . 19

3.5 The existence of a threshold . . . 20

3.6 Summary . . . 22

4 Iterative coloring propagation algorithm **22**

4.1 Notation . . . 23

4.2 Analysis of Algorithm 3 . . . 23

5 Tools **24**

5.1 Bounds on the intersection of two sphere caps . . . 24

5.2 Geometric Routing . . . 28

6 3-Neighborhood Algorithm **29**
6.1 Algorithm description 30
6.2 Proof of Theorem 6.1 31

7 Uniquely colorable graphs **33**
7.1 Definitions 33
7.2 Proof of Theorem 7.1 33

A Sphere caps cover appendix **36**

B Algorithm 2 Appendix **37**

C Algorithm 3 Appendix **42**

D Intersection of sphere caps appendix **43**
D.1 Proof of Lemma 5.4 44

E Uniquely colorable graphs Appendix **45**

F A note on the dimension range **45**

Finding planted k-coloring in vector k-colorable graphs

By: Roee David

Advisor: Prof. Uriel Feige

Department of Computer Science and Applied Mathematics

Weizmann Institute of Science

Abstract

We consider the following model for generating random 3-colorable graphs. The vertex set V corresponds to a set of n independent random vectors in $\mathbb{S}^d$, where $\mathbb{S}^d$ denotes the unit sphere in d dimensions. Each $v \in V$ is randomly and independently assigned a color from $\{1, 2, 3\}$. These colors are referred as the planted coloring. Place an edge between $u, v \in V$ if and only if the inner angle between u, v is larger than $120°$ and u, v have different colors. The motivation for considering this distribution is that it appears to fool the coloring algorithms with currently best performance guarantees, such as those based on semidefinite programming.

We design and analyze algorithms for 3-coloring such graphs with high probability (over the choice of graph). The input to the algorithms is the adjacency matrix of such a graph (without being given the actual geometric embedding of the vertices in $\mathbb{S}^d$). The task of 3-coloring the graphs becomes more difficult as the dimension d grows (though only up to a point, because when the dimension is very large the graphs have no edges and hence can trivially be 3-colored). Our algorithms work up to dimension $d < 4.93260 \log n$, which corresponds to average degree of roughly $n^{0.29}$.

Acknowledgements

I thank my advisor, Uriel Feige, for being so patient with me, for giving me inspiration and for his willingness to help at any time.

I thank Sharon for her deep love.

I thank my family and my friends for being there for me.

Lastly, I thank the Weizmann Institute and its staff for creating a great place to study in.

1 Introduction

Given a graph G and a set Col of k-colors, a vertex coloring of G assigns to every vertex a color from Col. The coloring is legal if and only if no two adjacent vertices are assigned the same color. A graph is k-colorable if it admits a legal coloring. The smallest k such that G is k-colorable is called the *chromatic number* of G, and is denoted by $\chi(G)$. Computing the chromatic number is well known to be $NP-hard$, see [10]. Therefore we do not expect to find a polynomial time algorithm for that problem. Likewise, given a k-colorable graph, finding a legal k-coloring is $NP-hard$ for every $k \geq 3$. We call this last problem the k-coloring problem.

Given that the k-coloring problem is NP-hard, one often considers a relaxed version of the problem. Given a k-colorable graph and a value $\hat{k} > k$, one needs to find in polynomial time a legal $\hat{k}$-coloring. Ideally, one would like $\hat{k}$ to be as close as possible to k, and the ratio $\frac{\hat{k}}{k}$ is referred to as the approximation ratio of the algorithm. One of the first such approximation algorithms colors 3-colorable graphs with at most $O(\sqrt{n})$ colors, where n is the number of vertices in the graph [15].

The currently best approximation algorithms use Semi Definite Programming (SDP). For example, the algorithm of [9] colors a 3-colorable graph with $O\left(n^{1/4}\right)$ colors. In these works, SDP is used in order to solve with high precision and in polynomial time optimization systems called Vector Programming (see for example [14], chapter 26). The following vector program is a relaxation of the k-coloring problem.

$$\begin{array}{lll} \min & k' & \\ \text{s.t.} & \langle v_i, v_j \rangle \leq -\frac{1}{k'-1} & \forall (i,j) \in E \\ & \langle v_i, v_i \rangle = 1 & \forall i \in V \\ & v_i \in \mathbb{R}^n & \end{array} \tag{1.1}$$

This system embeds every vertex of the graph in the unit hyper sphere in following way: there exists a threshold angle such that every pair of neighbors have an inner angle that is larger than the threshold angle. Notice that as k' gets smaller we get a larger threshold angle. We can treat (1.1) as a relaxation of the coloring problem since for k-colorable graphs we have a solution with $k' \leq k$. In the case where $k = 3$ we can see it in the following way: consider an equilateral triangle with all its endpoints in the unit circle. Treating each of its vertices as a vector one can get the following vectors: $[a, -a, 0]$, $[0, a, -a]$, $[-a, 0, a]$, where $a = \sqrt{\frac{1}{2}}$. One can see that the inner angle between each pair of these unit vectors is $-\frac{1}{2} = -\frac{1}{k'-1}$. So we can embed the 3 color classes to these 3 vectors with all the constraints being satisfied. We call it the naive solution. Generalizing it for general k is easy, see Karger et al. [9]. Hence for k-colorable graphs, we have that $k' \leq k$ in the optimal solution for (1.1).

Given an optimal solution of (1.1) we try to use the embedding we have to produce a coloring with a limited number of colors, or in different terminology, we try to find an algorithm that "rounds" the embedding to a coloring. We note that by (1.1) constraints every pair of neighboring vertices are far away from each other on the sphere. Using this observation, one of the rounding algorithms in [9] does the following procedure iteratively: It randomly picks a sphere cap, considers the induced graph of the vertices it contains, picks up isolated vertices from that graph and colors them with the same color (a new color is used in each iteration).

Now let's look on a different problem. Let $\chi_v(G)$ denote the optimum value of k' in the vector program (1.1) for G. We call this value the vector chromatic number of G. One may ask how large might the ratio $\frac{\chi(G)}{\chi_v(G)}$ be? Sadly there are graph families satisfying $\frac{\chi(G)}{\chi_v(G)} = \frac{n}{poly(\log(n))}$, see [5]. Using the algorithm from [9] on degree bounded graphs (each vertex degree is bounded by Δ) with $\chi_v(G) = k$ results in a $\Delta^{1-\frac{2}{k}}$ coloring. Can we hope for better? It turns out that not by much because for any ϵ there exist graphs families with $\chi_v(G) = k$ but with $\chi(G) \geq \Delta^{1-\frac{2}{k}-\epsilon}$, see [5].

1.1 Goals

The examples in [5] imply that if one is given an optimal solution for (1.1) showing that $\chi_v(G) \leq 3$, this by itself does not suffice in order to compute a 3-coloring (and not even a k-coloring for values of k nearly as high as $\Delta^{\frac{1}{3}}$), simply for the reason that G might not be 3-colorable. Hence a rounding technique that only uses the fact that a graph is vector 3-colorable has no hope of producing approximation ratios significantly better then those of Karger et al. [9]. However, this argument leaves something to be desired, because there might be rounding techniques that use additional information beyond the solution to (1.1). In particular, it may use the assumption that the graph is 3-colorable (even though it does not actually know a 3-coloring). The above examples do not exclude the possibility that given graphs that are actually 3-colorable, SDP (perhaps augmented with additional algorithmic techniques) can lead to significantly better approximations than those shown in Karger et al. [9]. The following are questions that motivate our work.

1. Can we find graph families that are 3-colorable (or more generally, k-colorable) for which current approximation algorithms require a number of colors much larger than 3? (Recall that the graphs of [5] are not 3-colorable.)

2. As the notion of "current approximation algorithms" is rather vague, we may rephrase the above question so that it refers only to a well defined specific class of algorithms. To employ such an approach, we would need to propose a definition for the class of algorithms that we wish to refer to.

3. Given a family of 3-colorable graphs that may serve as a positive answer for either of the above two questions (namely, that known algorithms do not perform well on it), can we design new polynomial time algorithms that do color these graphs with few colors.

4. Given that we design new algorithms that are tailored for a specific family of 3-colorable graphs, do the ideas that underlie these algorithms extend to coloring of 3-colorable graphs in general.

We consider a family of graphs that appears to be a good starting point for addressing these questions, because on this family the algorithm of Karger et al. [9] might use many colors. This family is parametrized by a notion of dimension. We provide polynomial time algorithms for coloring graphs from this family for a certain range of parameters.

1.2 Notation and definitions

In this work we shall consider high dimensional objects such as spheres and caps. Given an object , obj, its surface area is denoted by $S(obj)$ and its volume is denoted by $V(obj)$. Given a graph $G = (V, E)$, where V is the vertex set and E is the edge set, $N(v)$ denotes the set of the neighbors of $v \in V$. Given a graph and a set of k-colors, a $k-$coloring of G is an assignment of a color, from the k possible colors, to any vertex of the graph in such a way that no two vertices with an edge between them have the same color. The graph chromatic number $\chi(G)$ is the smallest k such that the graph has a k coloring. The set of colors is denoted $Col(G)$ and the color of a vertex $v \in V$ is $Col(v)$. If a random variable X takes values from the set S uniformly we'll denote it $X \in_R S$ and if D is some distribution we denote $d \sim D$ for d chosen by D. Given a vector v we note that $-v$ is the antipodal vector to v . Let u, v be two unit vectors then $d(u, v) := \arccos(u^t v)$ i.e $d(u, v)$ is u, v angular distance. When we state upper and lower bounds on some function f of an object we denote them by $\hat{f}, \check{f}$. We denote $a(n) \approx b(n)$ when $\lim_{n\to\infty}\left(\frac{a(n)}{b(n)}\right) = 1$. The initials SDP stand for "semi definite programming". Unless otherwise stated, $\log(x)$ is the natural base logarithm.

Definition 1.1. ($d-$Sphere). The d dimensional sphere with radius r is defined as follows: $\mathbb{S}^d_r = \{x \in \mathbb{R}^d : \|x\| = r\}$, using Euclidean norm. So S^2_1 is a unit circle. When the radius is 1 we omit it.

Definition 1.2. ($d-$Ball). The d dimensional ball with radius r is defined as follows: $\mathbb{B}^d_r = \{x \in \mathbb{R}^d : \|x\| \leq r\}$, using Euclidean norm.

Definition 1.3. (The $d-$Sphere Surface Area). There is a known formula for the sphere surface area: $S\left(\mathbb{S}^d_r\right) = \frac{2\pi^{\frac{d}{2}}}{\Gamma\left(\frac{d}{2}\right)} r^{d-1}$. where Γ is the gamma function . The Gamma function is an extension of the factorial function to real numbers. $\Gamma(n) = (n-1)!$.

Throughout, given a sphere $\mathbb{S}^d_r = \{x \in \mathbb{R}^d : \|x\| = r\}$, the term *sphere volume* will refer to the volume of the enclosed ball $\mathbb{B}^d_r$.

Definition 1.4. (The $d-$Sphere Volume). There is a known formula for the sphere Volume: $V\left(\mathbb{S}^d_r\right) = \frac{\pi^{\frac{d}{2}}}{\Gamma\left(\frac{d}{2}+1\right)} r^d$.

Definition 1.5. (Sphere Caps). Let $a \in$[0,1], $r \in [0, \pi]$, $l > 0$, and $\vec{x} \in \mathbb{S}^d$. An a-cap centered at $\vec{x}$ is defined to be the set $C^d_{a,l}(\vec{x}) = \left\{\vec{u} \in \mathbb{S}^d_l : \langle \vec{u}, \vec{x} \rangle \geq a\right\}$. A cap of angular radius r is defined to be the set $CR^d_l(\vec{x}, r) = \left\{\vec{u} \in \mathbb{S}^d_l : d(\vec{u}, \vec{x}) \leq r\right\}$. When d and r are known from the context we omit them and usually $r = 1$.

Definition 1.6. (Sphere Caps Radius $R(C_a)$). The radius of a sphere cap is the distance on the sphere form the center of the cap to the boundary of the cap. It's easy to see $R(C_a) = \arccos(a)$.

Definition 1.7. (Sphere Cap Measure $\mu(C_a)$). The sphere Cap measure is the relative surface area of the sphere cap to the surface area of the entire sphere. $\mu(C_a) = \frac{S(C_a)}{S(\mathbb{S}^d_r)}$. For example in S^n the measure of C_0 is $\frac{1}{2}$.

Fact 1.8. *(Bounds on the sphere Cap measure). We use the same bounds as in Feige and Schechtman [7]:* $\frac{c}{\sqrt{d}}(1-a^2)^{\frac{d-1}{2}} \leq \mu(C_a) \leq \frac{1}{2}(1-a^2)^{\frac{d-1}{2}}$, *where c is some constant independent of d.*

Fact 1.9. *(Bounds on surface area of a sphere cap in dimension d).*
$S(C_a^d) \geq V\left(\mathbb{S}^{d-1}_{\sqrt{1-a^2}}\right)$.
$S(C_a^d) \leq (1-a)\,S\left(\mathbb{S}^{d-1}_{\sqrt{1-a^1}}\right) + V\left(\mathbb{S}^{d-1}_{\sqrt{1-a^2}}\right)$.
This fact is easily illustrated in Figure 1

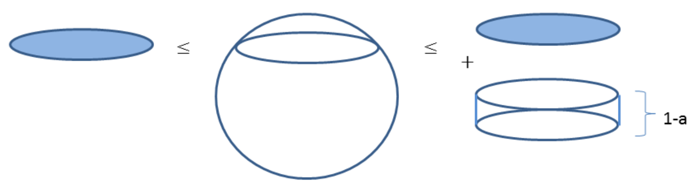


Figure 1: Illustration of Fact 1.9.

1.3 The model

In this section we define families of random geometric graphs, $G_{n,k,d}$,$G'_{n,k,d}$. Let $k \geq 3$ be a constant, let $d = d(n)$ be a function of the number of vertices, and let $\mathbb{S}^d = \{v \in R^d \mid \|v\| = 1\}$. The graph distribution $G'_{n,k,d}$ is defined as follows: take n random vectors $[\vec{v}_1 ... \vec{v}_n]$ each of which is drawn uniformly and independently from S^d (there are some techniques to do this, for example Knuth [11]), each one corresponds to a vertex $[v_1...v_n]$. Place an edge between two vertices if $\langle \vec{v}_i, \vec{v}_j \rangle \leq -\frac{1}{k-1}$. Now we modify G' to obtain G. For each vertex randomly choose a color $[1..k]$ (we call these the original colors of the graph) with some probability vector $\vec{P}$ and every edge between vertices of the same color is removed. If it is not stated otherwise then $\vec{P}$ is the uniform probability vector. $G \in G_{n,k,d}$ denotes that G is created that way.

The modification process can be applied also on arbitrary graphs. In the literature the original coloring is sometimes called the planted coloring.

1.3.1 Some properties of these graphs

First of all $G'_{n,k,d}$ is a distribution of a vector k-colorable graphs and actually it's the same that achieved the main results in [5]. The first observation holds because we can use $[\vec{v}_1 ... \vec{v}_n]$ to form a solution for (1.1), we call it the embedding solution. Note that $G_{n,k,d}$ is a distribution of a k-colorable graphs so (1.1) can be solved also with the naive solution. Suppose we try to use the SDP (1.1) in order to color the graph. Then if the SDP provides us the embedding solution then the rounding technique of Karger et al. [9] will use far more than k colors.

Now we give some technical properties of these graphs:

1. Since the density of vertices in the sphere is uniform, then every vertex has $\frac{k-1}{k}(n-1)\mu\left(C_{\frac{1}{k(n)-1}}\right)$ neighbors in expectation. Using some large deviation bounds one can show that these graphs tend to be almost regular.
2. As the dimension of these graphs is larger the graph has less edges (this follows by direct calculation of the expression $\frac{k-1}{k}(n-1)\mu\left(C_{\frac{1}{k(n)-1}}\right)$ with the provided lower/upper bounds).
3. Given $u \in N(v)$ of some vertex v then $Col(u) \in_R Col(G) \setminus Col(v)$.

1.4 Results

We provide polynomial time algorithms for coloring $G \in_R G_{n,3,d}$ with high probability over the choice of G. All the results that are mentioned here can be generalized for any constant k. The input for the algorithms is a graph adjacency matrix.

When d gets large the average degree of vertices in the graph gets small. If $d \geq 6.95212 \log n$ then our graphs have isolated vertices with high probability (see Section F). An interesting question that arises is what is the largest (in terms of asymptotic behavior) function of $d(n)$ such that our graphs admit polynomial time coloring algorithms. We show the following theorem:

Theorem 1.10. *Let $d = (c_{GRD} - \epsilon)\log n$, where ϵ is some small positive constant and $c_{GRD} \approx 4.93260$. Let $G \in G_{n,3,d}$. Algorithm 4 (see Section 6) legally colors G with high probability.*

Though we believe that ϵ in Theorem 1.10 can be chosen to be arbitrarily small, we prove the theorem for the case that ϵ is a fixed small constant, say $\epsilon = 1/10$. When

$d = (c_{GRD} - \epsilon)\log n$ then the expected degree of vertices in $G \in G_{n,3,d}$ is approximately $n^{0.29}$.

We also analyze natural algorithms for 3-coloring of random graphs, and show that for our distribution $G_{n,3,d}$ they are not as effective as Algorithm 4. Namely, the range of dimensions d for which they succeed is more limited.

Algorithm 1 samples a small set of vertices, colors them and thereafter deduces a coloring for the rest of the vertices. The algorithm is similar to the one in Arora et al. [3].

Theorem 1.11. *Let $d = c\log\log n$ for $c < 6.95$ and let $G \in G_{n,3,d}$. Algorithm 1 legally colors G with high probability.*

The analysis uses the fact that sphere caps have a low VC-dimension. Note that when $d = \Theta(\log\log n)$ then the expected degree of vertices in $G \in G_{n,3,d}$ is $c\frac{n}{(\log n)^{c'}}$, where $c, c' > 0$ are some constants.

Algorithm 2 is based on finding a threshold t such that any pair of vertices with more than t common neighbors have the same original color. It is similar to an algorithm that was presented in Blum and Spencer [4].

Theorem 1.12. *Let $d = c_0\frac{\log n}{\log\log n}$, where c_0 is an arbitrary constant smaller than 1, and let $G \in G_{n,3,d}$. Algorithm 2 legally colors G with high probability.*

Algorithm 3 is based on coloring three carefully chosen vertices by different colors (rather than choosing these vertices carefully, the algorithm tries all triplets), and iteratively propagating colors to additional vertices, as long as there are uncolored vertices whose neighbors are colored by exactly two different colors.

Theorem 1.13. *Let $d = c_0\log n$, where c_0 is an arbitrary constant smaller than* approximately 1.4426950*, and let $G \in G_{n,3,d}$. Algorithm 3 legally colors G with high probability.*

When $d = c_0\log n$, where c_0 is approximately 1.4426950, then the expected degree of vertices in $G \in G_{n,3,d}$ is approximately $n^{0.79}$.

The analysis of Algorithm 3 is the basis for the proofs in Section 7 where we show the following theorem:

Theorem 1.14. *Let $d = c_0\log n$, where c_0 is an arbitrary constant smaller than* approximately 1.4426950*, and let $G \in G_{n,3,d}$. With high probability the only legal 3-coloring of G is the planted coloring.*

1.5 Related work

Let $G \in G_{n,p}$ be a random graph with n vertices where an edge between each pair of vertices is placed independently from other pairs with probability p. Planting a 3-coloring is the following random process: Given a graph G, each vertex is assigned randomly, independently and uniformly with a color ($\{1, 2, 3\}$). Every edge that is shared by two vertices with the same color is removed. The distribution of graphs from the plating process applied on $G \in G_{n,p}$ is denoted by $G_{n,p,3}$. The expected degree of each vertex in $G \in G_{n,p,3}$ is $\frac{2}{3}np$. For this discussion we denote the expected degree of a vertex in a random graph by Deg.

Arora et al. [3] have shown a polynomial time algorithms that colors 3-colorable dense graphs (every vertex in a dense graph has $\Omega(n)$ neighbors) with high probability. This algorithm also colors $G \in G_{n,p,3}$ with high probability when p is a constant. One can show that if p is a function of n that tends to zero as n grows then this algorithm will not color G. In our model this algorithm performs better, we show that this algorithm colors $G \in G_{n,3,d}$ with average degree of $c\frac{n}{(\log n)^{c'}}$, where $c, c' > 0$ are some constants. For more details see Algorithm 1.

Blum and Spencer [4] present algorithms that find (with high probability) a coloring for $G \in G_{n,p,3}$. One of these algorithms colors $G \in G_{n,p,3}$ with high probability, when $p \geq n^{\epsilon - 1/2}$ for any fixed $\epsilon > 0$. This algorithm is based on counting paths of length 2 between any two vertices. In our model this algorithm (Algorithm 2) performs worse, $\forall \epsilon > 0$ this algorithm fails to color $G \in G_{n,3,d}$ with average degree smaller than $n^{1-\epsilon}$. Another algorithm that was presented in Blum and Spencer [4] is a generalization of the former algorithm and it is based on counting paths of different lengths. This algorithm colors $G \in G_{n,p,3}$ with high probability, when $p \geq n^{\epsilon - 1}$ for any fixed $\epsilon > 0$. We did not try to analyze it due to the poor performance of the former algorithm.

In Alon and Kahale [1] an algorithm that finds (with high probability) a coloring for $G \in G_{n,p,3}$ is presented, where $p = cn^{-1}$ for some constant c. The analysis of this algorithm uses spectral techniques. They proved and used the following fact: almost surely the adjacency matrix of $G \in G_{n,p,3}$ has two eigenvalues of value roughly $-Deg/2$, while the rest of the eigenvalues have value at least $-O\left(\sqrt{Deg}\right)$. Moreover, the eigenvectors corresponding to these highly negative eigenvalues are correlated with the planted 3-coloring in a way that leads to recovering this 3-coloring. See Proposition 2.1 in Alon and Kahale [1]. However in our model almost surely the adjacency matrix A of $G \in G_{n,3,d}$ has additional eigenvalues (unrelated to the planted coloring) of value roughly

$-Deg/2$. Consider for example a hyperplane through the origin, and consider the n-dimensional vector corresponding to the cut induced by the hyperplane. Namely, its entries are indexed by the vertices of G, and its jth coordinate is either $+1$ or -1 depending on the side of the hyperplane in which the corresponding vertex lies. It is not difficult to show that roughly two thirds of the edges of G are cut by the hyperplane, which using Rayleigh quotient considerations implies the existence of an eigenvalue of value smaller than $-Deg/3$. More generally, one may consider the following vector family $\{\vec{y}_j | 1 \leq j \leq d\}$, where $\vec{y}_j(i) = \begin{cases} 1 & \vec{v}_i(j) \geq 0 \\ -1 & \vec{v}_i(j) < 0 \end{cases}$. If d is $O(\log n)$ then all these vectors are nearly orthogonal to each other. This (with some extra work) can be used in order to show that A has multiple eigenvalues of value $-\Omega(Deg)$. The leading constant in the Ω notation can be increased either by considering refinements of the vector family $\{\vec{y}_j | 1 \leq j \leq d\}$ (so that the value of a coordinate depends on the distance of the corresponding vertex from the hyperplane), or by considering refinements of the geometric graphs (placing an edge if the inner angel is greater than $135°$ rather than $120°$). Details omitted.

In the $G_{n,p,3}$ model Algorithm 3 performs worse than Algorithm 2. One can show that if $p < n^{-\frac{1}{3}}$ then Algorithm 3 does not color $G \in G_{n.p.3}$ with high probability (as opposed to Algorithm 2 which colors $G \in G_{n,p,3}$ if $p \geq n^{\epsilon-1/2}$). But in our model ($G_{n,3,d}$) Algorithm 3 performs much better than Algorithm 2, see the results section.

We remark that Algorithm 4 was designed specifically to work in the model $G_{n,3,d}$, and uses properties of this model in an essential way. Even relatively small changes in the model might cause the algorithm to fail. In contrast, Algorithm 3 is more robust and works for a wider range of models. A formal approach for discussing such robustness notions is via semi-random models (see Blum and Spencer [4] and Feige and Kilian [6]) but this aspect will not be discussed here.

2 One shot coloring propagation algorithm

In this section we describe an algorithm for coloring $G \in G_{n,k,d}$. The algorithm is similar to the one in Arora et al. [3] but the analysis is different. We provide two proofs. The first proof is of correctness when d is a constant. Basically this proof only uses the fact that these graphs are dense and the way that the original coloring was created. The second proof is of correctness when $d = \Theta(\log \log n)$. This proof uses some of the geometric structure that was used to build $G \in G_{n,k,d}$. One of the geometric structure properties is that sphere caps of the unit sphere in $\mathbb{R}^d$ have a low VC - dimension (definition will be given in Section 2.2.1).

Algorithm 1

Input: A graph G.
Output: A k -coloring of the graph or a failure message.

1. Sample a random set S of vertices. Each vertex is in S with probability p independently from other vertices.
2. For every legal coloring of S.
   (a) For every vertex not in S
      i. If $k-1$ different colors are assigned to its neighbors that are within S then assign to it the remaining color.
      ii. Otherwise color it randomly.
   (b) If a correct coloring was found return the coloring and terminate the algorithm.
3. If no correct coloring was found return a failure message.

The running time of the algorithm is $O(nk^{|S|})$. To keep the running time polynomial $|S|$ must be of size $O(\log n)$. To assure this we can set $p = \frac{\Theta(\log n)}{n}$.

In the next sections we show proofs of the next theorem for various values of d.

Theorem 2.1. *Let $G \in G_{n,3,d}$. Algorithm 1 legally colors G with high probability.*

2.1 A proof of Theorem 2.1 for constant d

The proof prerequisites are that the input graph is of type $G_{n,k,d}$ where d and k are both constants.

Lemma 2.2. *If $p = \frac{c_1 \log n}{n}$, where c_1 is a constant then for any arbitrary small γ_1 it holds that* $(1-\gamma_1) c_1 \log(n) < |S| < (1+\gamma_1) c_1 \log n$ with high probability

Proof. For each vertex in the graph v we define a random variable

$$x_v = \begin{cases} 1 & \text{with probability}\, p \\ 0 & \text{otherwise} \end{cases}$$

Let $X = \Sigma_{v \in V} x_v$. $E[X] = pn = c_1 (\log n)$. X is distributed like $|S|$. We can apply the Chernoff bound:

$$\Pr[X > (1+\gamma_1) c \log n] < e^{-\frac{c_1 \log n \gamma_1^2}{4}}$$

and

$$\Pr[X < (1-\gamma_1) c \log n] < e^{-\frac{c_1 \log n \gamma_1^2}{2}}$$

so by union bound the probability that the statement will not hold is smaller then

$$\epsilon_1 = e^{-\frac{c_1 (\log n) \gamma_1^2}{2}} + e^{-\frac{c_1 (\log n) \gamma_1^2}{4}}$$

It is easy to see that we can set γ_1 to be arbitrary small and ϵ_1 tends to zero as n grows. □

Denote by $N(v, c)$ the group of neighbors of v that were originally assigned with color c and let $q = \frac{\mu\left(C_{\frac{1}{k-1}}\right)}{k}$. Observe that q is a constant independent of n. This is because d and k are constants independent of n. Another lemma that we need is the following:

Lemma 2.3. *For any $v \in G, c \neq col(v)$ and for any arbitrary small γ_2 it holds that* $(1-\gamma_2) nq \leq |N(v,c)| \leq (1+\gamma_2) nq$ w.h.p.

Proof. For each vertex in $u \neq v$ we define a random variable

$$x_u = \begin{cases} 1 & \text{with probability } q \\ 0 & \text{otherwise} \end{cases}$$

Let $X = \Sigma_{u \in V/\{v\}} x_u$. $E[X] = qn$. X is distributed like $|N(v,c)|$. We can apply the Chernoff bound:

$$\Pr[X > (1+\gamma_2) qn] < e^{-\frac{qn\gamma_2^2}{4}}$$

and

$$\Pr[X < (1-\gamma_2) qn] < e^{-\frac{qn\gamma_2^2}{2}}$$

so by union bound the probability that the statement will not hold is smaller then

$$\epsilon_2 = n(k-1)\left(e^{-\frac{qn\gamma_2^2}{4}} + e^{-\frac{qn\gamma_2^2}{2}}\right)$$

Set γ_2 to be arbitrary small. Recall that q is a constant. Therefore ϵ_2 tends to zero as n grows. $\square$

We prove that for $G \in G_{n.k,d}$ as defined in Section 1.3 the algorithm returns a legal coloring with high probability. It suffices to prove that once the algorithm finds the original coloring of S (we can assume it happens because there is an enumeration of all legal coloring of S) then w.h.p the algorithm finds the original coloring of G.
The probability of failure is at most :

$$\Pr\left[\ \exists v \in V \setminus S \,\text{s.t at least one of the}\ Col\left(G\right) \setminus Col\left(v\right) \text{colors is missing in}\ S \cap N\left(v\right)\right]$$

By union bound

$$\Pr\left[Failure\right] \leq n\left(k-1\right)\Pr\left[\ \text{given}\ v \in V \setminus S\, \text{the color}\ r \in Col\left(G\right) \setminus Col\left(v\right) \text{is missing in}\ S \cap N\left(v\right)\right]$$

By Lemma 2.3

$$\Pr\left[Failure\right] \leq n\left(k-1\right)\left(1-q\left(1+\gamma_2\right)\right)^{|S|}$$

If d and k are constants then $1-q\left(1+\gamma_2\right)=b<1$ is constant. For any ϵ_3 setting $p=\frac{1}{n}\left(\log_{\frac{1}{b}}\left(n\left(k-1\right)\right)+\log_{\frac{1}{b}}\left(\frac{1}{\epsilon_3}\right)\right)$ gives us :

$$\Pr\left[Failure\right] \leq n\left(k-1\right)\left(1-q\left(1+\gamma_2\right)\right)^{\left(\log_{\frac{1}{b}}\left(n(k-1)\right)+\log_{\frac{1}{b}}\left(\frac{1}{\epsilon_3}\right)\right)(1-\gamma_1)} \leq \epsilon_3$$

Here we use Lemma 2.2 to bound from bellow the size of S. Note that γ_1 can be arbitrary small. So Algorithm 1 fails to color the graph with probability at most $\epsilon_1+\epsilon_2+\epsilon_3$. If d and k are constants then $\epsilon_1+\epsilon_2$ tends to zero as n grows and ϵ_3 can be arbitrary small.

2.2 A proof of Theorem 2.1 for $d=c\left(\log\log n\right)$, where $c<\frac{2}{\log(4/3)}$ is a constant.

We assume for simplicity that $k=3$ (the proof can be extended to other values of k). The proof prerequisites are that the input graph is of type $G_{n,3,d}$ where: $d=c\left(\log\log n\right)$, where c is some constant satisfying $c<\frac{2}{\log(4/3)}$. The proof is divided to two parts. In the first part we show a lemma about the number of random sphere caps needed to cover the sphere with high probability. In the second part we use this fact to show that the algorithm colors the graph with high probability. We also show that if $c > \frac{2}{\log(4/3)}$ then the algorithm fails with high probability.

2.2.1 Sphere caps cover

In this section we prove that it is sufficient to sample a small set of sphere caps in order to get a sphere caps cover. The following definitions and theorem are well known and we use them later on.

Definition 2.4. (Range Space). A range space is a pair (X, R), where X is a set of elements and R is a set of subsets of X.

Definition 2.5. (Vapnic and Chervonenkis (VC) Dimension). A range space (X, R) *shatters* a set $A \subseteq X$ if for every $a \subseteq A$ there exists $r \in R$ such that $a = A \cap r$ (we say that r separates a). The VC-dimension of (X, R) is the size of the largest set it can shatter. We denote it by $D_{V.C}((X, R))$.

Definition 2.6. (ϵ-net). Let (X, R) be a range space and let $N \subseteq A \subseteq X$. N is an ϵ-net for A if $\forall r \in R$ such that $r \cap A \geq \epsilon A$ it holds that $r \cap N \neq \emptyset$.

The following theorem is due to Haussler and Welzl [8]:

Theorem 2.7. *(ϵ-net theorem). Let (X, R) be a range space of VC-dimension D_{VC} and let $A \subseteq X$. If N is a random subset of A (each element of N is drawn uniformly and independently) of size larger than max $\left(\frac{4}{\epsilon}\log\left(\frac{2}{\gamma}\right), \frac{8D_{VC}}{\epsilon}\log\left(\frac{8D_{VC}}{\epsilon}\right)\right)$ then N is an ϵ-net for A with probability at least $1 - \gamma$.*

A proof of the last theorem can be found in Alon and Spencer [2].

Corollary 2.8. *Consider the range space $\left(\mathbb{S}^d, \left\{C_a(\vec{v})\,|v \in \mathbb{S}^d\right\}\right)$. If N is a ϵ-net for $\mathbb{S}^d$ and $\epsilon = \mu(C_a)$ then $\bigcup_{x \in N} C_a(x) = \mathbb{S}^d$.*

Proof. For each $\vec{v} \in \mathbb{S}^d$ there exists $x \in N$ such that $x \in C_a(\vec{v})$ (by the definition of the $\mu(C_a)$-net). But in general $x \in C_a(\vec{v}) \Leftrightarrow \vec{v} \in C_a(x)$. And because this holds for each $\vec{v} \in \mathbb{S}^d$ then $\bigcup_{x \in N} C_a(x)$ is a sphere cover. □

Lemma 2.9. *If $a > 0$ then $D_{VC}\left(\mathbb{S}^d, \left\{C_a(\vec{v})\,|v \in \mathbb{S}^d\right\}\right) \leq d$.*

Proof. It is a well known fact that the set of the half-spaces in $\mathbb{R}^d$ has VC-dimension of $d+1$, i.e $D_{VC}\left(\mathbb{R}^d, \left\{\left\{x \in \mathbb{R}^d | \langle x, v\rangle \geq b\right\} | v \in \mathbb{R}^d, b \in \mathbb{R}\right\}\right) = d+1$. Let A be set of n points in S^d (and hence also in $\mathbb{R}^d$). We show that if the set of sphere caps $\left(\mathbb{S}^d, \left\{C_a(\vec{v}) | v \in \mathbb{S}^d\right\}\right)$ shatters A then $\left(\mathbb{R}^d, \left\{\left\{x \in \mathbb{R}^d | \langle x, v\rangle \geq b\right\} | v \in \mathbb{R}^d, b \in \mathbb{R}\right\}\right)$ shatters $A \cup \left\{\vec{0}\right\}$.

If $z \subseteq A \subseteq A \cup \left\{\vec{0}\right\}$ then there exists a sphere cap $C_a(\vec{v})$ that separates z therefore in $\mathbb{R}^d$ we can separate z by $\left\{x \in \mathbb{R}^d | \langle x, \vec{v}\rangle \geq a\right\}$ (note that all the caps in $\left(\mathbb{S}^d, \left\{C_a(\vec{v}) | v \in \mathbb{S}^d\right\}\right)$ cannot contain $\vec{0}$ since $a > 0$).

If $z \subseteq A \cup \left\{\vec{0}\right\}$ and $\vec{0} \in z$ then the set $\left(A \cup \left\{\vec{0}\right\}\right)/z$ has a sphere cap $C_a(\vec{v})$ that separates it. Therefore in $\mathbb{R}^d$ we can separate z by $\left\{x \in \mathbb{R}^d | \langle x, -\vec{v}\rangle \geq -a - \epsilon\right\}$ when ϵ is sufficiently small.

This shows that

$$D_{VC}\left(\mathbb{S}^d, \left\{C_a(\vec{v}) | v \in \mathbb{S}^d\right\}\right) + 1 \leq D_{VC}\left(\mathbb{R}^d, \left\{\left\{x \in \mathbb{R}^d | \langle x, v\rangle \geq b\right\} | v \in \mathbb{R}^d, b \in \mathbb{R}\right\}\right)$$

and the claim follows. □

The following theorem is a corollary of Theorem 2.7, Corollary 2.8 and Lemma 2.9.

Theorem 2.10. *Given a sphere cap $\mathbb{S}^d$, a random set of sphere caps C_a (each sphere cap center point is drawn uniformly and independently) of size larger than*

$$max\left(\frac{4}{\mu(C_a)}\log\left(\frac{2}{\gamma}\right), \frac{8d}{\mu(C_a)}\log\left(\frac{8d}{\mu(C_a)}\right)\right)$$

covers the sphere with probability at least $1-\gamma$.

In appendix A we show a self contained alternative proof of a theorem similar to Theorem 2.10, but with weaker bounds.

2.2.2 Geometric proof

We show that Algorithm 1 outputs a correct coloring with high probability if $k = 3$ and $d(n) = c(\log\log n)$, where c is defined later. Set the size of S to be some large enough multiple of $\log n$. This is done by setting the appropriate value for p and and using Lemma 2.2. We omit these details for simplicity.

Denote the original color classes red, blue and green. With high probability at least $\frac{|S|}{6}$ vertices out of S are red (similar proof as in Lemma 2.2 and this will hold for for all the colors with high probability). Denote them S_{red}. Assume that the algorithm has the original coloring of S (we can assume it happens because there is an enumeration of all legal coloring of S). Every $v \in S_{red}$ implies that vertices in $N(v)$ know that they are not colored with red color. In order that Algorithm 1 will succeed the following property must hold: every vertex in V/S that is not colored red has a red colored neighbor vertex in S. If $v \in S_{red}$ then each vertex that is not colored red whose corresponding vector is in the sphere cap $C_{\frac{1}{2}}(-\vec{v})$ will satisfy this property. Let $cover_{red}$ be defined as follows:

$$\bigcup_{v \in S_{red}} C_{\frac{1}{2}}(-\vec{v})$$

If $cover_{red}$ covers the whole sphere this property holds. By Theorem 2.10 and Fact 1.8

$$s = max\left(\frac{4}{\left(\frac{3}{4}\right)^{c\log\log n/2}}\log\left(\frac{2}{\gamma}\right), \frac{8c\log\log n}{\left(\frac{3}{4}\right)^{c\log\log n/2}}\log\left(\frac{8c\log\log n}{\left(\frac{3}{4}\right)^{c\log\log n/2}}\right)\right)$$

random sphere caps suffice in order to cover the sphere with probability $1-\gamma$. Since

$$\left(\frac{4}{3}\right)^{c\log\log n/2} = (\log n)^{-c\log(4/3)/2}$$

then

$$s = O\left((\log n)^{c\log(4/3)/2}\log\log n\right)$$

The size of the sampled set should not exceed $O(\log n)$ (otherwise Algorithm 1 running time will not be a polynomial) therefore $c\log(4/3)/2 < 1 \Rightarrow c < \frac{2}{\log(4/3)}$. One can see that for these values of c we can set γ to be arbitrary small and $(O(\log n))$ random sphere caps form a cover with probability $1-\gamma$.

If $cover_{red}$, $cover_{blue}$,$cover_{green}$ are all covering the sphere then Algorithm 1 will succeed, so by union bound on all the color classes the probability that Algorithm 1 fails is at most 3γ.

The proof extends for any constant k. The only thing that changes is the constant c.

2.2.3 Algorithm 1 does not color $G_{n,k,d}$ when $d(n)$ is too large

We showed that if $d(n) < \frac{2}{\log\left(\frac{4}{3}\right)} (\log\log n)$ then Algorithm 1 colors the graph with high probability. Now we show that if $d(n) > \frac{2}{\log\left(\frac{4}{3}\right)} (\log\log n)$ then Algorithm 1 will fail to color the graph with high probability. Note that $|S| = O(\log(n))$ otherwise the algorithm's running time would not be polynomial. We show that if $d(n) > \frac{2}{\log\left(\frac{4}{3}\right)} (\log\log n)$ and $|S| = O(\log(n))$ then arbitrary $v \in V/S$ has no neighbors in S.

$$\Pr[\text{The algorithm fails}] = \Pr[\exists v \in V/S \text{ such that the algorithm cannot determine its color}]$$

$$\geq \Pr[v_i \in V/S \text{ with no neighbors in } S]$$

$$\geq (1 - \mu(a))^{|S|} \approx e^{-\mu(a)O(1)\log(n)}$$

Now we use 1.8 and further observe that for our choice of d and c one has $\log(n) = e^{\frac{d}{c}}$.

$$e^{-\mu(a)O(1)\log(n)} = e^{-\frac{1}{2}O(1)\left(1-\frac{1}{2}^2\right)^{\frac{d-1}{2}} e^{\frac{d}{c}}}$$

$$= e^{-\frac{1}{2O(1)} e^{\log\left(1-a^2\right)\frac{d-1}{2}+\frac{d}{c}}} \approx e^{-\frac{1}{2O(1)} e^{d\left(\log\left(1-a^2\right)\frac{1}{2}+\frac{1}{c}\right)}}$$

$d(n) = c'\log(n)$. If $\log\left(\frac{3}{4}\right)\frac{1}{2} + \frac{1}{c} < 0 \Rightarrow c > \frac{2}{\log\left(\frac{4}{3}\right)}$ then the whole term tend to 1, which means that Algorithm 1 fails with high probability.

3 2-Neighborhood Algorithm

In this section we describe an algorithm for coloring $G \in G_{n,k,d}$ with different parameters. The algorithm is similar to an algorithm in Blum and Spencer [4]. We show the following theorem

Theorem 3.1. *Let $d = c_0 \frac{\log n}{\log\log n}$, where c_0 is an arbitrary constant smaller than 1, and let $G \in G_{n,3,d}$. Algorithm 2 legally colors G with high probability.*

Algorithm 2

Input: A graph G.

Output: A k -coloring of the graph or a failure message.

1. **Grouping:** Let t be a parameter whose value will be determined later. For each vertex $v \in V$ let logical-cap$_v$ be the set of vertices from G with more than t common neighbors with v, i.e logical-cap$_v = \{u | N(u) \cap N(v) \geq t\}$.
2. **Merge step:** Initially each one of the logical caps is referred to as a group. Iteratively replace any two groups that intersect by a new group formed by their union.
3. If after the merging stage the number of disjoint groups is k then color the graph in the following way: every group has a unique color and a vertex's color is determined by the group that it belongs to. If the coloring is valid return the coloring, otherwise return a failure message.

3.1 Proof plan for Theorem 3.1

Let $[\bar{v}_1 ... \bar{v}_n]$ be the random vectors on the sphere used to build $G \in G_{n,k,d}$ and let δ_1, δ_2 be arbitrary small constants. If t satisfies

$$(1+\delta_2)n\frac{k-2}{k}\mu\left(C_{\frac{1}{k-1}}\right) < t < (1-\delta_1)n\frac{k-1}{k}\check{\mu}_1 \tag{3.1}$$

where $\check{\mu}_1 = \mu\left(CR\left(R\left(C_{\frac{1}{k-1}}\right) - \frac{1}{2}R\left(C_{1-\epsilon}\right)\right)\right)$ and $\epsilon = \frac{1}{d^2}$ then with high probability over of $G \in G_{n.k,d}$ the following two lemmas hold.

Lemma 3.2. *For every $v \in V$ if $\vec{u} \in C_{1-\epsilon}(\vec{v})$ and u has the same color as v in the original coloring then $u \in$ logical-cap$_v$.*

Lemma 3.3. *For every $v \in V$ no vertex of G with a different color from that of v is in logical-cap$_v$.*

Using Lemma 3.2 and Lemma 3.3 we prove:

Lemma 3.4. *The merge stage returns the original coloring.*

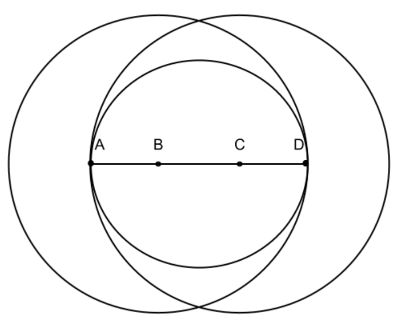
In Section 3.5 we show that there exists t satisfying Equation 3.1.

3.2 Proof of Lemma 3.2

Proof. Let μ_1 be the expected fraction of common neighbors of two vertices with angular distance less than $R(C_{1-\epsilon})$. We show that $\check{\mu}_1 \leq \mu_1$. Let $\vec{v_i}, \vec{v_j}$ be two vectors with angular distance less than $R(C_{1-\epsilon})$. When $d(\vec{v_i}, \vec{v_j})$ gets larger $\mu(\cap) := \mu\left(C_{\frac{1}{k-1}}(-\vec{v_i}) \cap C_{\frac{1}{k-1}}(-\vec{v_j})\right)$ gets smaller. Therefore assuming $\vec{v_j}$ is on the edge of $C_{1-\epsilon}(\vec{v_i})$ results in a lower bound on $\mu(\cap)$. $\mu(\cap)$ is larger than the measure of the largest sphere cap L that can be placed there. Note that $R(L) = R\left(C_{\frac{1}{k-1}}\right) - \frac{1}{2}R(C_{1-\epsilon})$, see Figure 2. Therefore $\frac{k-1}{k}\mu_1 \geq \frac{k-1}{k}\mu\left(CR\left(R\left(C_{\frac{1}{k-1}}\right) - \frac{1}{2}R(C_{1-\epsilon})\right)\right) = \frac{k-1}{k}\check{\mu}_1$.

Let us denote by $\Pr[\bar{E}]$ the probability that there exist two vertices in the graph such that: $\vec{v_j} \in C_{1-\epsilon}(\vec{v_i})$, both of them colored by the same original color and they have less than t common neighbors. t is smaller than $(1-\delta_1)n\frac{k-1}{k}\check{\mu}_1$ (recall that $\check{\mu}_1 \leq \mu_1$). By Union Bound $\Pr[\bar{E}] \leq n^2 \Pr[\bar{E}^{i,j}]$ where $\bar{E}^{i,j}$ is the event that: $\vec{v_j} \in C_{1-\epsilon}(\vec{v_i})$, both have the same original color and they have less than $(1-\delta_1)n\frac{k-1}{k}\check{\mu}_1$ common neighbors. By applying the Chernoff bound it follows that $\Pr[\bar{E}^{i,j}] \leq e^{-\Omega(\delta_1^2 \frac{k-1}{k}\check{\mu}_1 n)}$.

The probability that the lemma does not hold is at most $err_1 = \left(n^2 e^{-\Omega(\delta_1^2 \frac{k-1}{k}\check{\mu}_1 n)}\right)$ and in Lemma B.6 we prove that err_1 tends to zero as n grows. □



.

Figure 2: B,C are the centers of the two big spheres. $AC = BD = R(C_{\frac{1}{k-1}})$, $BC = R(C_{1-\epsilon})$ and $AD = 2R(L)$ and also $AD = AC + BD - BC$. Therefore $2R(L) = 2R\left(C_{\frac{1}{k-1}}\right) - R(C_{1-\epsilon})$.

3.3 Proof of Lemma 3.3

Proof. Let v_j,v_i be two vertices that are colored by different original colors. Let $\frac{k-2}{k}\mu_2$ be the expected fraction of vertices that are common neighbors of v_j,v_i. If $\vec{v_j} = \vec{v_i}$ then $\mu_2 = \mu\left(C_{\frac{1}{k-1}}\right)$, therefore $\mu_2 \leq \mu\left(C_{\frac{1}{k-1}}\right)$. Denote by $\Pr[\bar{E}]$ the probability that there are two vertices in the graph such that: they have different original colors and they have more then t common neighbors. t is larger than $(1+\delta_2)n\frac{k-2}{k}\mu\left(C_{\frac{1}{k-1}}\right)$. By union bound $\Pr[\bar{E}] \leq n^2 \Pr[\bar{E}^{i,j}]$ where $\bar{E}^{i,j}$ is the event that: v_j, v_i have different original colors and they have more than $(1+\delta_2)\mu_2 n$ common neighbors. By applying the Chernoff bound it follows that $\Pr[\bar{E}] \leq n^2 e^{-\Omega\left(\delta_2^2 \frac{k-2}{k}\mu\left(C_{\frac{1}{k-1}}\right)n\right)}$.

The probability that the lemma does not hold is at most $err_2 = n^2 e^{-\Omega\left(\delta_2^2 \frac{k-2}{k}\mu\left(C_{\frac{1}{k-1}}\right)n\right)}$ and similar proof as in Lemma B.6 can show that err_2 tends to zero as n grows. □

3.4 Proof of Lemma 3.4

This lemma shows that the merging process works with high probability. The proof prerequisites is that k is constant. With high probability $(1 - err_3)$ the number of vertices from each color is $\Theta(\frac{n}{k})$. The proof is similar to that of Lemma 2.2.

Consider the following random graph with $\Theta(\frac{n}{k})$ vertices: given a sphere S^d each vertex is a random point (points are drawn uniformly and independently from the sphere surface area) on it and $u \in N(v)$ iff $\vec{u} \in C_{1-\epsilon}(\vec{v})$. Followed by the previous lemmas it is left to prove that performing a merging process on this graph ends with one component i.e this graph is connected.

Assume towards a contradiction that after the merging process $t > 1$ groups remain. Let P be the closest (in terms of angular distance) pair of vertices v_1, v_2 from different groups. There exists a cap $CR\left(\frac{d(v_{i,1},v_{i,2})}{2}\right)$, B, centered between them (Area B in Figure 3) that contains no other vertex of the graph. The event that exist $CR\left(\frac{d(v_{i,1},v_{i,2})}{2}\right)$ cap that does not contain vertices is denoted by E.

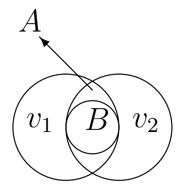


Figure 3: No vertex can be in A.

Recall that $\epsilon = \frac{1}{d^2}$ and note that $d(v_1, v_2) > R(C_{1-\epsilon}) = \arccos(1-\epsilon)$. Otherwise v_1, v_2 are merged together. For each vertex v place a sphere cap centered at $\vec{v}$ with radius r with the property $2r < \arccos(1-\epsilon)$. If these $\frac{n}{k}$ caps cover the sphere then there are no $CR\left(\frac{d(v_{i,1}, v_{i,2})}{2}\right)$ caps that do not contain vertices (otherwise the sphere caps will not cover the whole sphere).

Therefore by Lemma 3.5 E occurs with low probability. We conclude that $t = 1$.

Lemma 3.5. *$\frac{n}{k}$ random caps with radius $r < \frac{\arccos(1-\epsilon)}{2}$ cover the sphere with high probability.*

Proof. By Lemma 2.10

$$n' = max\left(\frac{4}{\mu(C_a)}\log\left(\frac{2}{\gamma}\right), \frac{8d}{\mu(C_a)}\log\left(\frac{8d}{\mu(C_a)}\right)\right)$$

random sphere caps are needed in order to cover the sphere with probability $1-\gamma$, where $a = \cos\left(\frac{1}{2}\arccos\left(1-\frac{1}{d^2}\right)\right)$. If for arbitrary γ it holds that $\lim_{n\to\infty}\frac{n'}{\frac{n}{k}} = 0$ then for large enough n the probability that the lemma does not apply is $err_4 = \gamma$. This follows by Lemma B.7. $\square$

3.5 The existence of a threshold

By Equation 3.1 it follows that if $\frac{\check{\mu}_1}{\mu\left(C_{\frac{1}{k-1}}\right)} > \frac{(1+\delta_1)}{(1-\delta_2)}\frac{k-2}{k-1}$ then there exists a threshold t. Note that δ_1, δ_2 are arbitrary small so if $\frac{\check{\mu}_1}{\mu\left(C_{\frac{1}{k-1}}\right)}$ tends to 1 as n grows then the claim follows.

Note that $\check{\mu}_1 = \mu\left(CR\left(R\left(C_{\frac{1}{k-1}}\right) - \frac{1}{2}R(C_{1-\epsilon})\right)\right)$ and $\epsilon = \frac{1}{d^2}$. So generally it is clear that setting ϵ to be any function that tends to zero fast enough as n grows will suffice to prove that $\frac{\check{\mu}_1}{\mu\left(C_{\frac{1}{k-1}}\right)}$ tends to 1 as n grows as well. The tradeoff is that this function cannot go too fast to zero otherwise Lemma 3.4 does not hold.

In order to evaluate $\frac{\check{\mu}_1}{\mu\left(C_{\frac{1}{k-1}}\right)}$ we could try to use Fact 1.8 directly but the ratio of the upper bound and the lower bound of the same sphere cap is $\sqrt{d}$. Therefore we use the following lemma.

Lemma 3.6. *If $a < b$ then*

$$S\left(C_a\right) - S\left(C_b\right) \leq (b-a)\, S\left(\mathbb{S}^{d-1}_{\sqrt{1-a^2}}\right) + V\left(\mathbb{S}^{d-1}_{\sqrt{1-a^2}}\right) - V\left(\mathbb{S}^{d-1}_{\sqrt{1-b^2}}\right)$$

Proof. $S\left(C_a\right) - S\left(C_b\right)$ can be upper bounded by a cylinder and an annulus. $(b-a)\, S\left(\mathbb{S}^{d-1}_{\sqrt{1-a^2}}\right)$ is the surface area of the cylinder and $V\left(\mathbb{S}^{d-1}_{\sqrt{1-a^2}}\right) - V\left(\mathbb{S}^{d-1}_{\sqrt{1-b^2}}\right)$ is the area of the annulus. See Figure 4. □

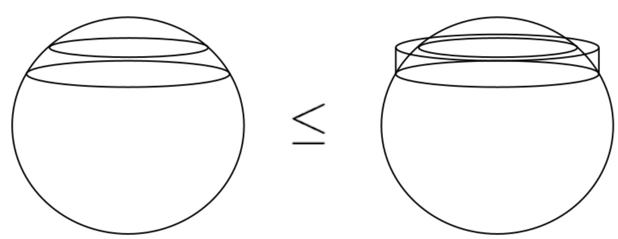


Figure 4: On the left we see a graphical representation of $S\left(C_a\right) - S\left(C_b\right)$ and on the right we see a graphical representation of $(b-a)\, S\left(\mathbb{S}^{d-1}_{\sqrt{1-a^2}}\right) + V\left(\mathbb{S}^{d-1}_{\sqrt{1-a^2}}\right) - V\left(\mathbb{S}^{d-1}_{\sqrt{1-b^2}}\right)$.

Note that by Fact 1.3 and Fact 1.4 $S\left(\mathbb{S}^{d-1}_{\sqrt{1-a^2}}\right)$, $V\left(\mathbb{S}^{d-1}_{\sqrt{1-a^2}}\right)$ and $V\left(\mathbb{S}^{d-1}_{\sqrt{1-b^2}}\right)$ have a closed formula.

The existence of a threshold proof itself is technical and it can be found in the appendix, Lemma B.1.

3.6 Summary

The probability that Algorithm 2 fails is at most $err_1+err_2+err_3+err_4$ and we showed this tends to zero as n grows.

4 Iterative coloring propagation algorithm

In this section we describe an algorithm for the coloring $G \in G_{n,3,d}$ with $d = c\log n$, where c is an arbitrary positive constant that satisfies $c \le c_0 = -\frac{1}{\log(\sin(30^\circ))} \approx 1.4426950$. The main idea is that given 3 vertices v_1, v_2, v_3 with 3 different original colors we can infer the color of any vertex u such that $\vec{u} \in C_{\frac{1}{2}}(-\vec{v}_1) \cap C_{\frac{1}{2}}(-\vec{v}_2) \cap C_{\frac{1}{2}}(-\vec{v}_3)$. This is because regardless the color of u it will have two neighbors with different original colors so it must have been assigned with the third color. As $\vec{v}_1, \vec{v}_2, \vec{v}_3$ are closer then the measure of $C_{\frac{1}{2}}(-\vec{v}_1) \cap C_{\frac{1}{2}}(-\vec{v}_2) \cap C_{\frac{1}{2}}(-\vec{v}_3)$ is larger (up to $\mu\left(C_{\frac{1}{2}}\right)$) and if $\vec{v}_1, \vec{v}_2, \vec{v}_3$ are too far away then $\mu\left(C_{\frac{1}{2}}(-\vec{v}_1) \cap C_{\frac{1}{2}}(-\vec{v}_2) \cap C_{\frac{1}{2}}(-\vec{v}_3)\right) = 0$. By enumerating all the possible triplets of vertices in the graph we will find three vertices that are roughly close and they have three different colors in the original coloring. Now we can color vertices that are in $C_{\frac{1}{2}}(-\vec{v}_1) \cap C_{\frac{1}{2}}(-\vec{v}_2) \cap C_{\frac{1}{2}}(-\vec{v}_3)$ and try to find new triplets (with conditions as before) among them in order to color new vertices.

Algorithm 3

Input: A graph G.
Output: A 3 -coloring of the graph or a failure message.

1. For every triplet of vertices v_1, v_2, v_3 do
   (a) Set $Col(v_1) = 1$,$Col(v_2) = 2$ and $Col(v_3) = 3$.
   (b) Set the rest of the graph's vertices as uncolored vertices.
   (c) While there exists a vertex v with more than 2 colored neighbors do
      i. If the colored neighbors of v have exactly 2 different colors, color v with the remaining color.
      ii. If the colored neighbors of v have 3 different colors abort the while loop.
   (d) If all the vertices in G are colored return this coloring.
2. Return a failure message.

4.1 Notation

Let $\mathbb{S}^d$ be a sphere of radius 1 in R^d, centered at the origin. Fix $\epsilon > 0$ to be arbitrarily small and fix $0 < \epsilon' \leq 2\epsilon$. Let $\mathcal{C}$ be the collection of caps C_x centered at $x \in S^d$, each of radius $30° - \epsilon$. We assume throughout that $\epsilon < \pi/12$.

4.2 Analysis of Algorithm 3

Definition 4.1. Given (d, r), we say that a graph $G \in G_{n,3,d}$ is (r)-*dense* if every cap of radius r contains at least one vertex from each color class.

Theorem 4.2. *For every d, ϵ, Algorithm 3 legally colors every $(d, 30° - \epsilon)$-dense graph.*

Before proving Theorem 4.2 we introduce the following definition.

Definition 4.3. Given (d, ϵ') the infinite graph H has as vertices all $x \in \mathbb{S}^d$, and (x, x') forms an edge if $d(x, x') \geq 180° - \epsilon'$.
Note that since $\epsilon' > 0$ then H is connected.

We now prove Theorem 4.2.

Proof. Algorithm 3 will pick three differently colored vertices in the cap C_x for some $x \in X$. Let $y \in X$ be a neighbor of x in H. Then all vertices in C_y will be colored correctly by Algorithm 3, as any vertex in C_x is at distance at least $180° - \epsilon' - 2(30° - \epsilon) \geq 120°$ from any vertex in C_y. Hence C_y also has three different colored vertices. The coloring propagates to all caps in $\mathcal{C}$ because H is connected and because the graph is $(d, 30° - \epsilon)$-dense. As $\mathcal{C}$ cover the sphere, the whole graph is legally colored (by the planted coloring). □

It remains to compute for which values of n and d is $G \in G_{n,3,d}$ likely to be dense. Observe that Theorem 4.2 works for any value of $\epsilon > 0$, and hence we may let ϵ tend to 0. This is equivalent to requiring that every cap of radius $30°$ has vertices from the three colors.

Lemma 4.4. *(Density lemma). Let γ be an arbitrary small constant. If $d < -\frac{1}{\log(\sin(30°))} \log n \approx 1.4426950 \log n$ then with probability $1 - 3\gamma$ it holds that $CR(\vec{x}, 30°)$ contains vectors that correspond to vertices from each original color class for every $\vec{x} \in S^d$.*

Proof. Assume that each color class has $\frac{n}{3} \pm o(n)$ colors. This can be proved using the Chernoff and the union bounds, we omit the details. If $\frac{n}{3} \pm o(n)$ random sphere caps $CR(30°)$ cover the sphere with high probability then by applying the union bound the following holds for each color class. Every $\vec{x} \in S^d$ is contained in some $C(\vec{v}, 30°)$ where v is in that color class and also $\vec{v} \in CR(\vec{x}, 30°)$. It is left to verify that indeed $\frac{n}{3}$ random sphere caps $C_{1-\epsilon}$ cover the sphere with high probability. By Lemma 2.10

$$n' = max\left(\frac{4}{\mu(CR(30°))}\log\left(\frac{2}{\gamma}\right), \frac{8d}{\mu(CR(30°))}\log\left(\frac{8d}{\mu(CR(30°))}\right)\right)$$

random sphere caps are needed in order to cover the sphere with probability $1-\gamma$. By Lemma C.1, for arbitrary γ it holds that $\lim_{n\to\infty} \frac{n'}{\frac{n}{3}} = 0$. □

5 Tools

5.1 Bounds on the intersection of two sphere caps

Given two sphere caps $C_b(\vec{v}_1), C_b(\vec{v}_2)$ at angular distance $\alpha < 90°$ between their centers we would like to estimate $\mu(\cap, \alpha, b) := \mu(C_b(\vec{v}_1) \cap C_b(\vec{v}_2))$. Denote $\mu(\cap, \alpha) := \mu\left(\cap, \alpha, \frac{1}{2}\right)$. Denote $H_b(\vec{v}_i) := \{\vec{x} \in \mathbb{R}^n \mid \langle \vec{v}_i, \vec{x}\rangle = b\}$, in other words $H_b(\vec{v}_i)$ is the hyperplane that defines $C_b(\vec{v}_i)$. Let $x = \frac{\vec{v}_1 + \vec{v}_2}{2}$. Let $\bar{x} = cx$ where c is some scalar satisfying $cx \in H_b(\vec{v}_1) \cap H_b(\vec{v}_2)$. For illustration see Figure 5.

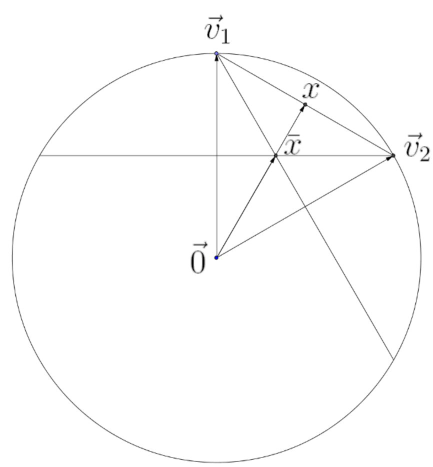


Figure 5: The vectors $x, \bar{x}, \vec{v}_1, \vec{v}_2$ in the special case where $a = b = \frac{1}{2}$ and $\alpha = 60°$.

To bound $\mu(\cap, \alpha, b)$ from above we use a sphere cap that encloses the intersection area, for illustration see Figure 6.

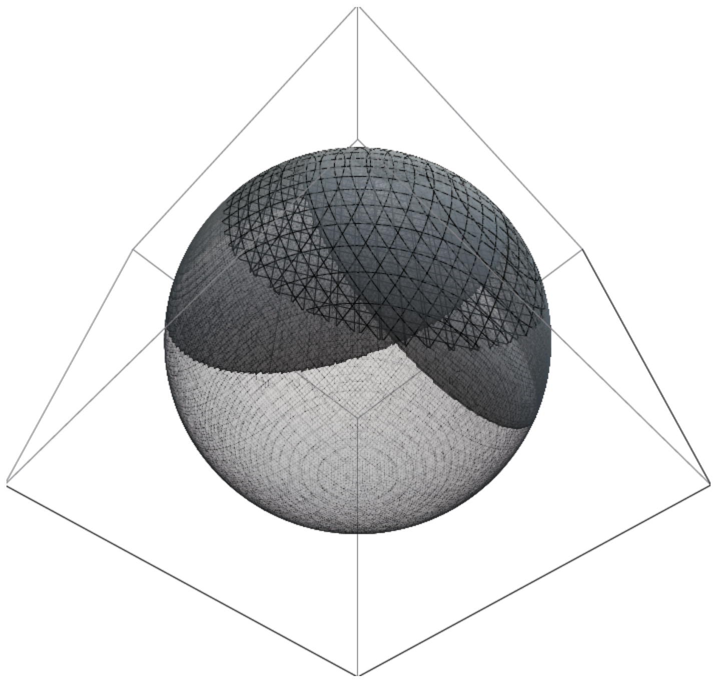


Figure 6: Bounding the intersection of sphere caps with a sphere cap.

Let z_1 be the following sphere cap

$$z_1 = C_{\|\bar{x}\|}\left(\frac{\bar{x}}{\|\bar{x}\|}\right)$$

z_1 encloses $C_b(\vec{v}_1) \cap C_b(\vec{v}_2)$. Therefore z_1 surface area is larger than $C_b(\vec{v}_1) \cap C_b(\vec{v}_2)$ surface area. Using standard trigonometry it follows that $\|\bar{x}\| = \frac{b}{\cos\left(\frac{\alpha}{2}\right)}$. Fact 1.8 implies the following lemma.

Lemma 5.1. $\mu(\cap, \alpha, b) \leq \mu(z_1) \leq O(1)\left(\sqrt{1 - \left(\frac{b}{\cos\left(\frac{\alpha}{2}\right)}\right)^2}\right)^d$

We define the complete sphere cap to be sphere cap of $\mathbb{B}_r^d$ (compared to sphere cap that was defined before as sphere cap of $\mathbb{S}_r^d$).

Definition 5.2. (Complete Sphere Caps). Let $a \in [0,1]$ and $\vec{x} \in \mathbb{S}^d$. A complete a-cap centered at $\vec{x}$ is defined to be the set $CC_{a,r}^d(\vec{x}) = \{\vec{u} \in \mathbb{B}_r^d : \langle \vec{u}, \vec{x} \rangle \geq a\}$. When d and r are known from the context we omit them and usually $r = 1$.

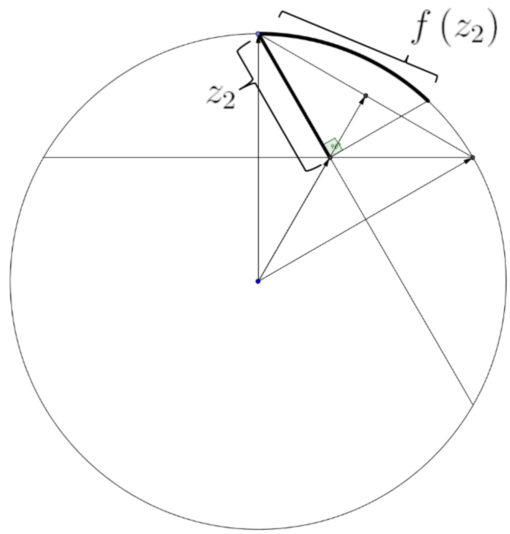


Figure 7: Illustration of Lemma 5.3

To bound $\mu(\cap, \alpha, b)$ from below we use the following region:

$$z_2 = H_b(\vec{v}_1) \cap CC_b(\vec{v}_2)$$

Lemma 5.3. *The volume of z_2 is smaller than the surface area of $C_b(\vec{v}_1) \cap C_b(\vec{v}_2)$.*

Proof. z_2 is a $d-1$ dimensional region which is contained in $H_b(\vec{v}_1)$ and therefore $\forall x \in z_2$ it holds that $\langle x, \vec{v}_1 \rangle = 0$. Let $f : z_2 \longrightarrow C_b(\vec{v}_1) \cap C_b(\vec{v}_2)$ be defined as follows $f(\vec{x}) = \vec{x} + c_x \vec{v}_1$, where $c_{\vec{x}}$ is a positive constant satisfying $\|f(\vec{x})\| = 1$. Since $\alpha < 90°$ than it holds that $f(z_2) \subset C_b(\vec{v}_1) \cap C_b(\vec{v}_2)$. Let $\vec{x}_1, \vec{x}_2 \in z_2$, it holds that $\|\vec{x}_1 - \vec{x}_2\| \leq \|f(\vec{x}_1) - f(\vec{x}_2)\|$ and the claim follows. For illustration see Figure 7 □

The proof of the following lemma uses Lemma 5.3 and it can be found in Section D.1.

Lemma 5.4. $\mu(C_b(\vec{v}_1) \cap C_b(\vec{v}_2)) \geq \frac{O(1)}{d} \left(\sqrt{1 - \left(\frac{b}{\cos\left(\frac{\alpha}{2}\right)} \right)^2} \right)^d$

Now we show bounds of intersection measure of two sphere caps of different size $C_{b_1}(\vec{v}_1), C_{b_2}(\vec{v}_2)$ at angular distance $\alpha < 90°$ between their centers. Denote $\mu(\cap, \alpha, b_1, b_2) := \mu(C_{b_1}(\vec{v}_1) \cap C_{b_2}(\vec{v}_2))$. Let D_2 be the plane spanned by $\vec{v}_1, \vec{v}_2$. In this work we deal only with the case that $\bar{x} = H_{b_1}(\vec{v}_1) \cap H_{b_2}(\vec{v}_2) \cap D_2$ contained in the triangle $\vec{0}, \vec{v}_1, \vec{v}_2$. Let $R = \|\bar{x}\|_2$ and note that R is a function of b_1,b_2 and α.

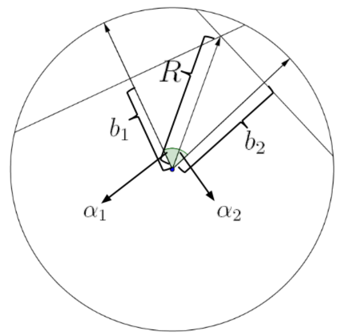


Figure 8: $C_{b_1}(\vec{v}_1)$ and $C_{b_1}(\vec{v}_1)$ on the plane D_2.

With similar arguments as in the proofs of Lemma 5.1 and in Lemma 5.3 the following lemma can be proved:

Lemma 5.5. $\frac{c''}{d^{1.5}}\left(\sqrt{1-R^2}\right)^d \leq \mu\left(\cap, \alpha, b_1, b_2\right) \leq c'\left(\sqrt{1-R^2}\right)^d$

It is left to calculate $R = \|\bar{x}\|_2$. We define α_1, α_2 to be the inner angle between $\vec{x}$ and $\vec{v}_1, \vec{v}_2$ and we need to solve the following equations:

1. $\cos\left(\alpha_1\right) = \frac{b_1}{R}$
2. $\cos(\alpha_2) = \frac{b_2}{R}$
3. $\alpha_1 + \alpha_2 = \alpha$

$$\Rightarrow R = \frac{b_1}{\cos\left(\alpha - \arccos\left(\frac{b_2}{R}\right)\right)}$$

$$\Rightarrow R = \sqrt{{b_2}^2 + {b_2}^2\left(\cot\alpha\right)^2 - 2b_1b_2\frac{\cos\alpha}{\left(\sin\alpha\right)^2} + {b_1}^2\frac{1}{\left(\sin\alpha\right)^2}}$$

The solution can be verified using standard trigonometry, for illustration see Figure 8.

5.2 Geometric Routing

We define another type of graphs distribution $\bar{G'}_{n,3,d}$ that is related to $G'_{n,3,d}$.

Definition 5.6. ($\bar{G'}_{n,3,d}$ Graph Distribution). The $\bar{G'}_{n,3,d}$ graph distribution is defined as follows: let N be the set of vertices. For each vertex $v \in N$ let $\vec{v}$ a vector drawn uniformly (and independently) from $\mathbb{S}^d$. $u \in N(v)$ if and only if $\vec{u} \in C_{\frac{1}{2}}(\vec{v})$ (rather than $\vec{u} \in C_{\frac{1}{2}}(-\vec{v})$ in $G'_{n,3,d}$).

It will be convenient to use the following definitions for the next section:

Definition 5.7. ($K'_{n,a,d}$ Graph Distribution). The $K'_{n,a,d}$ graph distribution is defined as follows: let N be the set of vertices. For each vertex $v \in N$ let $\vec{v}$ a vector drawn uniformly (and independently) from $\mathbb{S}^d$. $u \in N(v)$ if and only if $\vec{u} \in C_a(-\vec{v})$.

$K'_{n,a,d}$ is related to $G'_{n,a,d}$. $K_{n,a,d}, \bar{K}'_{n,a,d}$ are defined similarly to the way $K'_{n,a,d}$ is defined and they are related to $G_{n,a,d}, \bar{G}'_{n,a,d}$.
A nice property of $G \in \bar{G'}_{n,3,d}$ that is likely to hold only when d is small enough is the following:

Definition 5.8. (Geometric Routing). We say that a graph $G \in \bar{G'}_{n,3,d}$ supports Geometric Routing if any two vertices $u, v \in G$ are either neighbors or u has a neighbor w strictly closer (in terms of angular distance) to v. We also say that a graph $G \in G'_{n,3,d}$ supports Geometric Routing if the graph $\bar{G}$ supports geometric routing, where $\bar{G}$ has the same vertex set as G and $u \in N(v)$ if and only if $\vec{u} \in C_{\frac{1}{2}}(\vec{v})$.

This definition is useful because of the following lemma:

Lemma 5.9. *If $G \in \bar{G'}_{n,3,d}$ supports geometric routing then G is connected.*

Proof. The proof follows from Definition 5.8. If G supports geometric routing then for every two vertices u, v there is a path in G that is connecting them. The path starts in u. If u is a neighbor of v then the path ends at v. Otherwise the next vertex in the path is a neighbor of u which is strictly closer (in terms of angular distance) to v (by Definition 5.8 there exists such a vertex). The rest of the vertices on that path are defined in a similar manner. Since G is a finite graph then eventually one of the vertices of the path that was defined is a neighbor of v. □

Let c_{GRD} be a constant that satisfies the following: if $d < c_{GRD} \log(n)$ then $G \in \bar{G'}_{n,3,d}$ supports Geometric Routing with high probability.

Lemma 5.10. $c_{GRD} \geq \frac{2}{\log\left(\frac{3}{2}\right)} \approx 4.93260$ up to terms that tend to infinity as n grows.

Proof. Let $G \in \bar{G}'_{n,3,d}$, where $d = c \log n$. Let u, v be two vertices with angular distance α. If $\alpha > 60°$ then a vertex that satisfying $w \in N(u)$ and $d(v,w) < d(v,u)$ could be placed only at

$$int = C_{\frac{1}{2}}(\vec{u}) \cap CR(\vec{v}, d(u,v))$$

The measure of *int* is lower bounded by $\mu(\cap, 60°)$. The expectation of the number of vertices in this region is $E = (n-2)\mu(\cap, 60°) \approx n\frac{c''}{\sqrt{d}}\left(\sqrt{1 - \left(\frac{1}{2\cos(30°)}\right)^2}\right)^d = n\frac{c''}{\sqrt{d}}\left(\sqrt{\frac{2}{3}}\right)^d$. By using similar calculation as in Section F one can show that if $1 + \log\left(\frac{2}{3}\right)\frac{c}{2} > 0 \Rightarrow c < \frac{2}{\log\left(\frac{3}{2}\right)} \approx 4.93260$ then E is some polynomial of n. By applying the union bounds on each pair of vertices and by applying the Chernoff bounds one can show that G supports geometric routing with high probability and therefore $c_{GRD} \geq \frac{2}{\log\left(\frac{3}{2}\right)} \approx 4.93260$. □

In the analysis of Algorithm 4 (that we present next) we use the following lemma:

Lemma 5.11. *Let $G \in G'_{n,3,d}$. If $d < c_{GRD} \log n$ then G is connected with high probability.*

Proof. Let u, v be any two vertices of G. If $d(\vec{u}, \vec{v}) \leq 60°$ then $\exists w$ such that $u, v \in N(w)$. This follows since

$$\mu\left(C_{\frac{1}{2}}(-\vec{u}) \cap C_{\frac{1}{2}}(-\vec{v})\right) = \mu(\cap, d(\vec{u}, \vec{v})) \geq \mu(\cap, 60°)$$

and one can see that similar arguments as in Lemma 5.10 would apply this statement. Let G_1 be a graph that is a result of replacing the edge set of G with a new edge set: $u \in N(v) \Leftrightarrow d(\vec{u}, \vec{v}) \leq 60°$. It is easy to see that if G_1 is connected then G is also connected. Because G_1 has the same distribution as $H \in \bar{G}'_{n,3,d}$ the proof follows from Lemma 5.9. □

6 3-Neighborhood Algorithm

To simplify the presentation we assume that $k = 3$. In this section we describe Algorithm 4 for coloring graphs and show the following theorem:

Theorem 6.1. *Let $d = (c_{GRD} - \epsilon)\log n$, where ϵ is a small constant larger than 0, and let $G \in G_{n,3,d}$. Algorithm 4 legally colors G with high probability.*

Suppose that the random vectors on the sphere, denote them by $[\vec{v}_1 ... \vec{v}_n]$, that were used to build $G_{n,3,d}$ are given as part of the problem's input. For every vertex v_i of the graph define Logical-Cap$_{v_i}$ $:= \left\{ v_j | \left(\langle \vec{v}_i, \vec{v}_j \rangle \leq -\frac{1}{2} \right) \wedge (v_j \notin N(v_i)) \right\}$. This set contains only vertices that have the same original color as v_i. Apply the merge step on these sets (as defined in Algorithm 2). These groups behave like random spherical caps of radius 60° that contain vertices with the same original color. By applying similar arguments as in Lemma 3.4 one can see that the merge step colors $G_{n,3,d}$ with higher dimensions than in Algorithm 2. This is because in Algorithm 2 the merge step uses a small logical caps ($C_{1-\epsilon}$) and here these logical caps are much bigger ($C_{\frac{1}{2}}$).

If there exists an algorithm for restoring $[\vec{v}_1 ... \vec{v}_n]$ then we can color $G \in G_{n,3,d}$ with high probability. It seems that in order to restore $[\vec{v}_1 ... \vec{v}_n]$ we can use SDP as in Equation 1.1. Note, however, that because this SDP has the naive solution (as mentioned in the introduction this can be found in Karger et al. [9], Lemma 4.1) and also any convex combination between it and $[\vec{v}_1 ... \vec{v}_n]$ then it is unclear how to find $[\vec{v}_1 ... \vec{v}_n]$ using SDP and in any other method. So instead of finding $[\vec{v}_1 ... \vec{v}_n]$ we would like to determine the following: given a vertex v find *almost* all the vertices laying in $C_{\frac{1}{2}}(-\vec{v})$.

Definition 6.2. $Close_v^t = \{u \in V \mid N(u) \cap N(v) \geq t\}$

Intuitively if t is large enough and $u \in Close_v^t$ then $\vec{u}$ and $\vec{v}$ are close.

6.1 Algorithm description

Algorithm 4

Input: A graph G.
Output: A k -coloring of the graph.

1. **Building Logical caps:** For all $v \in G$:

   (a) $T_v^1 \leftarrow \{u \in V \mid N(u) \cap Close_v^{t_1} \geq t_2\}$

   (b) $T_v^2 \leftarrow \{v\} \cup (T_v^1 / N(v))$

2. **Merging:** For any two groups if $T_v^2 \cap T_u^2 \neq \emptyset$ then merge them into one group $T_v^2 \leftarrow T_v^2 \cup T_u^2$. Continue this process until no two groups share a common vertex.

Later we set t_1,t_2 and we prove that the algorithm returns a coloring with high probability. We show that T_v^1 contains almost all the original neighbors of v (vertices that were neighbors of v before the modification). Therefore T_v^2 contains almost all the original neighbors of v who are in the same color as v. The algorithm's running time is clearly polynomial.

6.2 Proof of Theorem 6.1

Recall the notions and definitions of sphere caps intersections (see Section 5.1). Let $d = c\log(n)$ where c is a constant satisfying $c < c_{GRD}$.

Lemma 6.3. *$\forall\epsilon' > 0\ \exists t_1$ such that with high probability for $G \in G_{n,3,d}$ it holds that $\forall u,v$ if $\vec{u} \in C_{\frac{1}{2}+\epsilon'}(\vec{v})$ then $u \in Close_v^{t_1}$ and if $\vec{u} \notin C_{\frac{1}{2}}(\vec{v})$ then $u \notin Close_v^{t_1}$.*

Proof. Let $f(\epsilon') = R\left(C_{\frac{1}{2}}\right) - R\left(C_{\frac{1}{2}+\epsilon'}\right)$. Let G' be the graph before the modification and let v, u be arbitrary vertices.
For $u \in C_{\frac{1}{2}+\epsilon'}(\vec{v})$ let $\hat{t_1'}(u,v) = |N(u) \cap N(v)|$ in G', it holds that

$$E\left[\hat{t_1'}\right] > (n-2)\,\check{\mu}\left(\cap, 60^\circ - f(\epsilon')\right)$$

In order to get a lower bound on $\hat{t_1'}$ in G we can assume that u, v are not in the same color in G and set $\hat{t_1} \approx \frac{1}{3}\hat{t_1'}$.
For $u \notin C_{\frac{1}{2}}(\vec{v})$ let $\check{t_1'}(u,v) = |N(u) \cap N(v)|$ in G', it holds that

$$E\left[\check{t_1'}\right] < (n-2)\,\hat{\mu}\left(\cap, 60^\circ\right)$$

In order to get an upper bound on $\check{t_1'}$ in G we can assume that u, v are in the same color in G and set $\check{t_1} \approx \frac{2}{3}\check{t_1'}$.
Recall that $d < c\log(n)$. With similar calculations as in Lemma 5.10 it follows that $\frac{1}{3}\hat{t_1'} = O(1)\,n^{\alpha}$ and $\frac{2}{3}\check{t_1'} = O(1)\,n^{\beta}$, where α, β are constants satisfying $0 < \beta < \alpha$. Therefore by applying the Chernoff and the Union bounds one can show that for arbitrary small ϵ' and large enough n it holds that $\hat{t_1} > \check{t_1}$ for every set of vertices in the graph. Therefore if $\check{t_1} < t_1 < \hat{t_1}$ then the proof follows. □

Lemma 6.4. *$\exists\epsilon'', t_1, t_2$ such that with high probability for $G \in G_{n,3,d}$ it holds that $\forall u,v$ if $\vec{u} \in C_{\frac{1}{2}+\epsilon''}(-\vec{v})$ then $u \in T_v^1$ and if $\vec{u} \notin C_{\frac{1}{2}}(-\vec{v})$ then $u \notin T_v^1$.*

Proof. Let v be an arbitrary vertex and set $\epsilon'' = 0.1$. By applying the previous lemma with $\epsilon' = 10^{-6}$ we can assume that there exists t_1 satisfying $\forall u, v\ \vec{u} \in C_{\frac{1}{2}+\epsilon'}(\vec{v}) \Rightarrow u \in Close_v^{t_1}$ and $\vec{u} \notin C_{\frac{1}{2}}(\vec{v}) \Rightarrow u \notin Close_v^{t_1}$. Therefore the quantity$E\left[\left|N(u) \cap Close_v^{t_1}\right|\right]$ can be lower/upper bounded as follows:

For $\vec{u} \in C_{\frac{1}{2}+\epsilon''}(-\vec{v})$ let $\hat{t_2}(u,v) = \left|N(u) \cap Close_v^{t_1}\right|$, it holds that

$$E\left[\hat{t_2}\right] \geq (n-2)\,\mu\left(\cap, d(-u,v), \frac{1}{2}, \frac{1}{2}+\epsilon'\right)$$

$$\geq (n-2)\,\mu\left(\cap, R\left(\frac{1}{2}+\epsilon''\right), \frac{1}{2}, \frac{1}{2}+\epsilon'\right)$$

$$\geq \frac{c''}{\sqrt{d}}(0.829156)^d$$

Where the last inequality is by Lemma 5.5. For $\vec{u} \notin C_{\frac{1}{2}}(-\vec{v})$ let $\check{t_2}(u,v) = \left|N(u) \cap Close_v^{t_1}\right|$, it holds that

$$E\left[\check{t_2}\right] = E\left[\left|N(u) \cap Close_v^{t_1}\right|\right] \leq (n-2)\,\mu\left(\cap, 60^\circ, \frac{1}{2}, \frac{1}{2}\right)$$

$$\leq c'(0.816497)^d$$

Where the last inequality is by Lemma 5.5. Recall that $d < c\log(n)$ and therefore $\frac{c''}{\sqrt{d}}(0.829156)^d, c'(0.816497)^d \in \Omega\left(n^{O(1)}\right)$. By applying the Chernoff and Union bounds one can show that for large enough n it holds that $\check{t_2} < \hat{t_2}$ for every set of vertices in the graph. Therefore if $\check{t_2} < t_2 < \hat{t_2}$ then the proof follows. □

Set ϵ'', t_1, t_2 as in the proof of the previous lemma.

Corollary 6.5. *For every v it holds that* $\left\{u|\vec{u} \in C_{\frac{1}{2}+\epsilon''}(-\vec{v})\right\} \subseteq T_v^1 \subseteq \left\{u|\vec{u} \in C_{\frac{1}{2}}(-\vec{v})\right\}$.

Let $\epsilon > 0$ be minimal such that for $d \leq (c_{GRD} - \epsilon)\log(n)$ it holds that $G \in K'_{\frac{n}{3},\frac{1}{2}+\epsilon'',d}$ is connected with probability. Similar proof as in Lemma 5.11 can show that there exists some small constant ϵ such that this graph is connected with high probability. Note that ϵ is a function of ϵ'' that satisfies $\lim_{\epsilon'' \to 0} \epsilon(\epsilon'') = 0$.

In order to prove that the algorithm colors the graph with high probability it suffices to prove that the merging process works. Lets look on a new graph L with the same set of vertices as G but now $(u,v) \in E(L)$ if $u \in T_v^2$ before the merging process. From Corollary 6.5 L is a disjoint union of 3 components L_i each of which has a subgraph distributed the same as $K'_{\frac{n}{3},\frac{1}{2}+\epsilon'',d}$. And since these are all connected with high probability then the merging process finds the original coloring with high probability.

7 Uniquely colorable graphs

Recall Definition 4.1 of a $(d, 30° - \epsilon)$-dense graph.

Theorem 7.1. *If $G \in G_{n,3,d}$ is a $(d, 30° - \epsilon)$-dense graph then G is uniquely 3-colorable, where $\epsilon < 30°$.*

7.1 Definitions

Definition 7.2. (The coloring equivalence relation). Given a graph G and a proper coloring of it χ, we denote by $CER(G,\chi)$ the equivalence relation on vertices induced by the coloring i.e $(u,v) \in CER(G,\chi) \iff Col(u) = Col(v)$.

Two legal coloring are considered equivalent if $CER(G,\chi) = CER(G,\chi')$ (namely they are the same up to permuting the names of the colors).

7.2 Proof of Theorem 7.1

We assume toward a contradiction that we have two different 3-coloring of $G:\ \chi$ (the planted coloring) and χ' such that $CER(G,\chi) \neq CER(G,\chi')$.

It must be the case that there exists $(u,v) \notin CER(G,\chi)$ but $(u,v) \in CER(G,\chi')$. Otherwise χ' uses more colors then χ or $CER(G,\chi) = CER(G,\chi')$. We call such pair (u,v) uniting vertices.

Lemma 7.3. *Let $G \in G_{n,3,d}$ be a $(d, 30° - \epsilon)$-dense graph,* where $\epsilon < 30°$. *If there exist two different coloring for G: χ and χ' then there exist uniting vertices (u,v) such that* $\vec{u},\vec{v} \in CR(\vec{x}, 30° - \epsilon)$ for some $\vec{x} \in S^d$.

Proof. Let (u',v') be the closest (in terms of angular distance) uniting vertices in the graph. Suppose towards a contradiction that there is no $CR(30° - \epsilon)$ sphere cap containing them both. Hence $\vec{u'},\vec{v'} \notin CR(mid, 30° - \epsilon)$, where $mid = \frac{\vec{u'}+\vec{v'}}{2} / \left\| \frac{\vec{u'}+\vec{v'}}{2} \right\|$. Since

G is a $(d, 30°)$-dense graph there are 3 vertices c_1, c_2, c_3 in $CR(mid, 30° - \epsilon)$ with different colors in χ. If $\exists i\ (u', c_i) \in CER(G, \chi')$ then since $CER(G, \chi')$ is an equivalence relation then also $(v', c_i) \in CER(G, \chi')$ and we got a contradiction since c_i is closer to v' than u'. Hence $\forall 1 \leq i \leq 3 : (u', c_i) \notin CER(G, \chi')$ but this also leads to a contradiction because on the one hand $1 \leq i \leq j \leq 3 : (c_i, c_j) \notin CER(G, \chi)$ but on the other hand $\exists i, j : (c_i, c_j) \in CER(G, \chi')$ otherwise χ' uses too many colors. □

Let $\mathcal{C}, \epsilon, \epsilon'$ be defined as in Section 4.1 and let H be defined as in Definition 4.3. Now we show a proof of Theorem 7.1

Proof. Suppose for the sake of contradiction that in addition to the planted 3-coloring χ there is a different 3-coloring χ'. Then by Lemma 7.3 there is some cap $C_x \in \mathcal{C}$ with two vertices u and v with different colors in χ but the same color in χ'. Any cap $C_y \in \mathcal{C}$ with y a neighbor of x in H has at most two colors in χ'. This is because any vertex in C_x is at distance at least $180° - \epsilon' - 2(30° - \epsilon) \geq 120°$ from any vertex in C_y. One of the color classes of χ' in C_y then includes two vertices with different colors in χ but the same color in χ'. Hence due to the connectivity of H, every cap in $\mathcal{C}$ is colored by at most two colors in χ' and has two vertices u and v with different colors in χ but the same color in χ'. Let *col* be a coloring of X defined as follows: color every $x \in X$ by the set of the colors of C_x according to χ' (each set of colors in χ' is a color class in *col*). This gives a 6-coloring of X because every cap $C_z \in \mathcal{C}$ contains some vertices of G and these vertices are colored by at least one color and by at most two colors. Note that$|\{x \subseteq \{c_1, c_2, c_3\}\, |0 < |x| < 3\}| = 6$. The coloring of X is also a 6-coloring of H. Let $x, y \in X$ be neighbors in H and let x_1, x_2 be the two vertices in C_x that are colored with some color (say c_1) in χ' although they have different colors in χ. Because as before any vertex in C_y is a neighbor of one of the vertices x_1, x_2 then every vertex in C_y is not colored with c_1 in χ'. Therefore the set of colors of vertices in C_y differs from the set of color of vertices in C_x. Hence *col* is a legal 6-coloring of H. By the results of Feige et al. [5], when ϵ is sufficiently small, then H is not 6-colorable which is a contradiction, see Theorem E.1. □

It remains to compute for which values of n and d is $G \in G_{n,3,d}$ likely to be dense. Observe that Theorem 7.1 works for any value of ϵ which satisfy the conditions of Theorem E.1, and hence ϵ may be a function that tends to zero as n grows. This is equivalent to requiring that every cap of radius 30° has vertices from the three original colors. Therefore by Lemma 4.4 if $G \in G_{n,3,d}$ with $d = c \log n$, where c is an arbitrary positive constant that satisfies $c \leq c_0 = -\frac{1}{\log(\sin(30°))} \approx 1.4426950$, then G is uniquely colorable graph with high probability.

References

[1] N. Alon and N. Kahale. A spectral technique for coloring random 3-colorable graphs. *SIAM Journal on Computing*, 26(6):1733–1748, 1997.

[2] N. Alon and J.H. Spencer. *The probabilistic method*, volume 73. Wiley-Interscience, 2011.

[3] S. Arora, D. Karger, and M. Karpinski. Polynomial time approximation schemes for dense instances of np-hard problems. In *Proceedings of the twenty-seventh annual ACM symposium on Theory of computing*, pages 284–293. ACM, 1995.

[4] Avrim Blum and Joel Spencer. Coloring random and semi-random k-colorable graphs. *Journal of Algorithms*, 19(2):204–234, September 1995.

[5] Feige, Langberg, and Schechtman. Graphs with tiny vector chromatic numbers and huge chromatic numbers. *SICOMP: SIAM Journal on Computing*, 33, 2004.

[6] U. Feige and J. Kilian. Heuristics for semirandom graph problems. *Journal of Computer and System Sciences*, 63(4):639–671, 2001.

[7] U. Feige and G. Schechtman. On the optimality of the random hyperplane rounding technique for max cut. *Random Structures & Algorithms*, 20(3):403–440, 2002.

[8] D. Haussler and E. Welzl. ϵ-nets and simplex range queries. *Discrete & Computational Geometry*, 2(1):127–151, 1987.

[9] D. Karger, R. Motwani, and M. Sudan. Approximate graph coloring by semidefinite programming. *Journal of the ACM (JACM)*, 45(2):246–265, 1998.

[10] Richard M. Karp. Reducibility among combinatorial problems. In *Complexity of Computer Computations*, pages 85–103, 1972.

[11] D.E. Knuth. The art of computer programming (volume 2-seminumerical algorithms) author: Donald e. knuth, publisher: Addison-wesle. 1969.

[12] S. Li. Concise formulas for the area and volume of a hyperspherical cap. *Asian Journal of Mathematics and Statistics*, 4(1):66–70, 2011.

[13] H. Maehara. A threshold for the size of random caps to cover a sphere. *Annals of the Institute of Statistical Mathematics*, 40(4):665–670, 1988.

[14] V.V. Vazirani. *Approximation algorithms.* springer, 2004.

[15] Avi Wigderson. A new approximate graph coloring algorithm. In *STOC*, pages 325–329, 1982.

A Sphere caps cover appendix

Lemma A.1. *There exists a cover of S^d with sphere caps C_a^d containing at most $\frac{1}{\mu(C_{a'})}$ caps where $a' = cos\left(\frac{1}{2}arccos\left(a\right)\right)$.*

Proof. Note that $R\left(C_{a'}\right) = \frac{1}{2}arccos\left(a\right) = \frac{1}{2}R\left(C_a\right)$. The biggest number of caps $C_{a'}$ that could be placed on the sphere without intersecting is at most $\frac{1}{\mu(C_{a'})}$. For any maximal non-intersecting arrangement of caps $C_{a'}$ on the sphere (in a sense that no space for other cap) replacing the original caps with with caps of radius 2 times bigger will cover all the sphere. a Point left uncovered in the original arrangement means one could add a new cap $C_{a'}$ exactly in this point contradicting the maximality of the arrangement. □

Lemma A.2. *Fixing a sphere cap $C_{a'}^d$ on the sphere the probability for a random cap C_a^d to cover it is $\mu\left(C_{a'}\right)$ where $a' = cos\left(\frac{1}{2}arccos\left(a\right)\right)$.*

Proof. Fix a $C_{a'}^d$ cap on the sphere and denoted it by t. Any other cap C_a^d whose center in t will cover it , and any other cap C_a^d whose center not in t will not cover it . □

Lemma A.3. *t random sphere caps C_a will cover the unit sphere with probability at least $1 - \frac{1}{\mu(a'')}\left(1-\mu(a')\right)^t$, where $a' = cos\left(\frac{1}{2}arccos\left(a\right)\right)$ and $a'' = cos\left(\frac{1}{2}arccos\left(a'\right)\right)$*

Proof. We bound the probability of failure. Fix a deterministic cover of $C_{a'}$ of size $\frac{1}{\mu(C_{a''})}$, by Lemma A.1. For every one of these caps the probability that it is not covered by the random C_a caps is $(1 - C_{a'})$, by Lemma A.2. So the probability it's left uncovered is $(1 - C_{a'})^t$. The lemma follows using union bound. □

In Maehara [13] certain threshold of the size of n random spherical caps was shown that causes them to cover the all unit sphere. We note that Lemma A.3 is an alternative proof of the case the sphere caps size passes the threshold using a more standard technique. The difference is that the result here is general for any dimension where in Maehara [13] the result is for dimension 3.

B Algorithm 2 Appendix

In the calculation through out the appendix we denote $Int = cos\left(\arccos\left(\frac{1}{k-1}\right) - \frac{\arccos(1-\epsilon)}{2}\right)$. Note that $\epsilon = \frac{1}{d^2}$.

Lemma B.1. If $\epsilon = \frac{1}{d^2}$ and k is a constant then $\frac{\check{\mu}_1}{\mu\left(C_{\frac{1}{k-1}}\right)} \approx 1$.

Proof. Denote $Int = cos\left(\arccos\left(\frac{1}{k-1}\right) - \frac{\arccos(1-\epsilon)}{2}\right)$.

$$\frac{\check{\mu}_1}{\mu\left(C_{\frac{1}{k-1}}\right)} = \frac{\mu\left(C_{Int}\right)}{\mu\left(C_{\frac{1}{k-1}}\right)} = \frac{S\left(C_{Int}\right)}{S\left(C_{\frac{1}{k-1}}\right)} =$$

$$1 - \frac{S\left(C_{\frac{1}{k-1}}\right) - S\left(C_{Int}\right)}{S\left(C_{\frac{1}{k-1}}\right)} \geq$$

Using 3.6.

$$1 - \frac{\left(Int - \frac{1}{k-1}\right) S\left(\mathbb{S}^{d-1}_{\sqrt{1-\left(\frac{1}{k-1}\right)^2}}\right) + V\left(\mathbb{S}^{d-1}_{\sqrt{1-\left(\frac{1}{k-1}\right)^2}}\right) - V\left(\mathbb{S}^{d-1}_{\sqrt{1-(Int)^2}}\right)}{S\left(C_{\frac{1}{k-1}}\right)} \geq$$

Using 1.9

$$1 - \frac{\left(Int - \frac{1}{k-1}\right) S\left(\mathbb{S}^{d-1}_{\sqrt{1-\left(\frac{1}{k-1}\right)^2}}\right) + V\left(\mathbb{S}^{d-1}_{\sqrt{1-\left(\frac{1}{k-1}\right)^2}}\right) - V\left(\mathbb{S}^{d-1}_{\sqrt{1-(Int)^2}}\right)}{V\left(\mathbb{S}^{d-1}_{\sqrt{1-\left(\frac{1}{k-1}\right)^2}}\right)} =$$

$$= \frac{V\left(\mathbb{S}^{d-1}_{\sqrt{1-(Int)^2}}\right)}{V\left(\mathbb{S}^{d-1}_{\sqrt{1-\left(\frac{1}{k-1}\right)^2}}\right)} - \frac{\left(Int - \frac{1}{k-1}\right) S\left(\mathbb{S}^{d-1}_{\sqrt{1-\left(\frac{1}{k-1}\right)^2}}\right)}{V\left(\mathbb{S}^{d-1}_{\sqrt{1-\left(\frac{1}{k-1}\right)^2}}\right)} =$$

Using 1.3 and 1.4:

$$\frac{\frac{\pi^{\frac{d-1}{2}}}{\Gamma(\frac{d-1}{2}+1)}\left(\sqrt{1-(Int)^2}\right)^{d-1}}{\frac{\pi^{\frac{d-1}{2}}}{\Gamma(\frac{d-1}{2}+1)}\left(\sqrt{1-\left(\frac{1}{k-1}\right)^2}\right)^{d-1}} - \frac{\left(Int - \frac{1}{k-1}\right)\frac{2\pi^{\frac{d-1}{2}}}{\Gamma(\frac{d-1}{2})}\left(\sqrt{1-\left(\frac{1}{k-1}\right)^2}\right)^{d-2}}{\frac{\pi^{\frac{d-1}{2}}}{\Gamma(\frac{d-1}{2}+1)}\left(\sqrt{1-\left(\frac{1}{k-1}\right)^2}\right)^{d-1}}$$

Since $\frac{\Gamma(\frac{d-2}{2}+1)}{\Gamma(\frac{d-2}{2})} \approx \frac{d-1}{2} + 1$ For $k(n) = O(1)$ the last term is $O(1)\left(Int - \frac{1}{k-1}\right)\left(\frac{d-1}{2} + 1\right)$ which tends to zero as n grows, see Lemma B.4.

$$= \frac{\frac{\pi^{\frac{d-1}{2}}}{\Gamma(\frac{d-1}{2}+1)}\left(\sqrt{1-(Int)^2}\right)^{d-1}}{\frac{\pi^{\frac{d-1}{2}}}{\Gamma(\frac{d-1}{2}+1)}\left(\sqrt{1-\left(\frac{1}{k-1}\right)^2}\right)^{d-1}}$$

$$= \left(\frac{1-(Int)^2}{1-\left(\frac{1}{k-1}\right)^2}\right)^{\frac{d-1}{2}}$$

This term tends to 1 as n grows, see Lemma B.5 : □

$$\approx 1$$

Lemma B.2. $lim_{d\to\infty}\left[\frac{cos\left[arccos\left[\frac{1}{k-1}\right]-\frac{arccos[1-eps]}{2}\right]}{\frac{1}{2}+\frac{1}{d^2}}\right] = 1$ *i.e* $Int \approx \frac{1}{2} + \frac{1}{d^2}$.

Proof. The proof uses continuity of the trigonometric functions.

$\lim_{d\to\infty}\left[\frac{\cos\left[\arccos\left[\frac{1}{k-1}\right]-\frac{\arccos[1-\mathrm{eps}]}{2}\right]}{\frac{1}{2}+\frac{1}{d^2}}\right]$

$= 2\lim_{d\to\infty}\left[\cos\left[\arccos\left[\frac{1}{k-1}\right] - \frac{\arccos[1-\mathrm{eps}]}{2}\right]\right]$

$= 2\cos\left[\lim_{d\to\infty}\left[\arccos\left[\frac{1}{k-1}\right] - \frac{\arccos[1-\mathrm{eps}]}{2}\right]\right]$

$= 2\left[\cos\left[\arccos\left[\frac{1}{k-1}\right] - \lim_{d\to\infty}\left[\frac{\arccos[1-\mathrm{eps}]}{2}\right]\right]\right]$

$= 2\left[\cos\left[\arccos\left[\frac{1}{k-1}\right] - 0\right]\right] =_{k=3} 1$ □

Lemma B.3. $lim_{d\to\infty}\left(\frac{1-\left(\frac{1}{2}+\frac{1}{d^2}\right)}{1-Int}\right)^d = 1$

Proof. $\lim_{d\to\infty}\left(\frac{1-\left(\frac{1}{2}+\frac{1}{d^2}\right)}{1-Int}\right)^d = \lim_{d\to\infty}e^{d\log\left(\frac{1-\left(\frac{1}{2}+\frac{1}{d^2}\right)}{1-Int}\right)}$

Now with the similar proof as Lemma B.2 we can see $\lim_{d\to\infty}\left(\frac{1-\left(\frac{1}{2}+\frac{1}{d^2}\right)}{1-Int}\right) = 1$ so using $\lim_{x\to 1}\frac{\log(x)}{x-1} = 1$ we get:

$$= \lim_{d\to\infty}e^{d\left(\frac{1-\left(\frac{1}{2}+\frac{1}{d^2}\right)}{1-Int}-1\right)} = e^{\mathrm{Lim}_{d\to\infty}d\left(\frac{1-\left(\frac{1}{2}+\frac{1}{d^2}\right)}{1-Int}-1\right)}$$

But

$$\lim_{d\to\infty}d\left(\frac{1-\left(\frac{1}{2}+\frac{1}{d^2}\right)}{1-Int}-1\right) = 0$$

We apply l'Hôpital's rule:

$$\lim_{d\to\infty}\frac{\frac{d}{dd}\left(\frac{1-\left(\frac{1}{2}+\frac{1}{d^2}\right)}{1-Int}-1\right)}{\frac{d}{dd}\frac{1}{d}}$$

$$\lim_{d\to\infty}\frac{\Theta\left(\frac{1}{d^3}\right)}{-\frac{1}{d^2}} = 0$$

□

Lemma B.4. $lim_{d\to\infty}\left(Int - \frac{1}{k-1}\right)\left(\frac{d-1}{2}+1\right) = 0.$

Proof. $lim_{d\to\infty}\left(Int - \frac{1}{k-1}\right)\left(\frac{d-1}{2}+1\right)$

By LemmaB.2 (k=3).

$= lim_{d\to\infty}\left(\frac{1}{d^2}\right)\left(\frac{d-1}{2}+1\right) \leq lim_{d\to\infty}\frac{d}{d^2} = 0$ □

Lemma B.5. $lim_{d\to\infty}\left(\frac{1-(Int)^2}{1-\left(\frac{1}{k-1}\right)^2}\right)^{\frac{d-1}{2}} = 1.$

Proof. $lim_{d\to\infty}\left(\frac{1-(Int)^2}{1-\left(\frac{1}{k-1}\right)^2}\right)^{\frac{d-1}{2}} = lim_{d\to\infty}\left(\frac{1-(Int)^2}{1-\left(\frac{1}{k-1}\right)^2}\right)^{-\frac{1}{2}} lim_{d\to\infty}\left(\frac{1-(Int)^2}{1-\left(\frac{1}{k-1}\right)^2}\right)^{-\frac{d}{2}}$

By continuity of the the functions involve:

$$= lim_{d\to\infty}\left(\frac{1-(Int)^2}{1-\left(\frac{1}{k-1}\right)^2}\right)^{-\frac{d}{2}} = *$$

Now we replace Int by $\frac{1}{2}+\frac{1}{d^2}$ so we have to prove that:

$$lim_{d\to\infty}\frac{\left(\frac{1-(Int)^2}{1-\left(\frac{1}{k-1}\right)^2}\right)^{-\frac{d}{2}}}{\left(\frac{1-\left(\frac{1}{2}+\frac{1}{d^2}\right)^2}{1-\left(\frac{1}{k-1}\right)^2}\right)^{-\frac{d}{2}}} = 1$$

but this hold by simple manipulation and using Lemma B.3. So we continue:

$$* = lim_{d\to\infty}\left(\frac{1-\left(\frac{1}{2}+\frac{1}{d^2}\right)^2}{1-\left(\frac{1}{k-1}\right)^2}\right)^{-\frac{d}{2}}$$

$k = 3$

$$= \left(lim_{d\to\infty}\left(\left(\frac{4}{3}\right)\left(1-\left(\frac{1}{2}+\frac{1}{d^2}\right)^2\right)\right)^d\right)^{-\frac{1}{2}}$$

$$= \left(lim_{d\to\infty}e^{d\left(\log\left(\frac{4}{3}\right)+\log\left(1-\left(\frac{1}{2}+\frac{1}{d^2}\right)^2\right)\right)}\right)^{-\frac{1}{2}}$$

$$= \left(e^{lim_{d\to\infty}d\left(\log\left(\frac{4}{3}\right)+\log\left(1-\left(\frac{1}{2}+\frac{1}{d^2}\right)^2\right)\right)}\right)^{-\frac{1}{2}}$$

Now we left to show :

$$lim_{d\to\infty}d\left(\log\left(\frac{4}{3}\right)+\log\left(1-\left(\frac{1}{2}+\frac{1}{d^2}\right)^2\right)\right) = 0$$

Let $t = \frac{1}{d}$ then we get:

$$lim_{t\to 0}\frac{\left(\log\left(\frac{4}{3}\right)+\log\left(1-\left(\frac{1}{2}+t^2\right)^2\right)\right)}{t}$$

we can use l'Hôpital's rule

$$lim_{t\to 0}\frac{-2\left(\frac{1}{2}+t^2\right)2t}{1-\left(\frac{1}{2}+t^2\right)^2}=0$$

□

Lemma B.6. $lim_{d\to\infty}n^2e^{-\Omega(\delta_1^2\frac{k-1}{k}\check{\mu}_1 n)}=0$

Proof. Substituting $\check{\mu}_1$ and using lower bounds on sphere caps if the following expression tends to zero then the claim holds:
$lim_{d\to\infty}n^2e^{-\frac{k-1}{k}n\left(\frac{1}{2}\left(1-\text{int}^2\right)^{\frac{d-1}{2}}\right)}=0$
Note that $d=c_0\frac{\log n}{\log\log n}$ and therefore proving for $d=\frac{\log n}{\log\log n}$ is suffice.

$$e^{lim_{d\to\infty}2\log n-\frac{k-1}{k}n\left(\frac{1}{2}\left(1-\text{int}^2\right)^{\frac{d-1}{2}}\right)}$$

We have to prove:

$$lim_{d\to\infty}2\log n-\frac{k-1}{k}n\left(\frac{1}{2}\left(1-\text{int}^2\right)^{\frac{d-1}{2}}\right)=-\infty$$

$$lim_{d\to\infty}2\log n-lim_{d\to\infty}\frac{k-1}{k}n\left(\frac{1}{2}\left(1-\text{int}^2\right)^{\frac{d-1}{2}}\right)$$

Using Lemma B.5 we can replace $(1-\text{int}^2)$ by $1-\left(\frac{1}{k-1}\right)^2$.

$$lim_{d\to\infty}2\log n-lim_{d\to\infty}\frac{k-1}{k}n\left(\frac{1}{2}\left(1-\left(\frac{1}{k-1}\right)^2\right)^{\frac{d-1}{2}}\right)$$

$$lim_{d\to\infty}2\log n-\Omega\left(n^{c'}\right)=-\infty$$

Where c' is any constant strictly larger than 1. □

Lemma B.7. *If* $d=c\frac{\log n}{\log\log n}$ *then* $\lim_{n\to\infty}\frac{n'}{\frac{n}{k}}=0$*, where* $c<1$.

Proof. Note that $n'=max\left(\frac{4}{\mu(C_a)}\log\left(\frac{2}{\gamma}\right),\frac{8d}{\mu(C_a)}\log\left(\frac{8d}{\mu(C_a)}\right)\right)$ and that $a=\cos\left(\frac{1}{2}\arccos\left(1-\frac{1}{d^2}\right)\right)$. One can check that in terms of asymptotic behavior we only need to show that $\lim_{n\to\infty}\frac{\frac{1}{\mu(C_a)}}{n}=0$. We use Fact 1.8

$$\lim_{n\to\infty} \frac{\left(1 - \cos\left[\frac{1}{2}\text{arccos}\left[1 - \frac{1}{d^2}\right]\right]^2\right)^{-\frac{1}{2}d}}{n} =$$

$$\lim_{n\to\infty} \frac{2^{d/2}d^d}{n} =$$

$$\lim_{n\to\infty} \frac{2^{c\frac{\log n}{\log\log n}/2} c\frac{\log n}{\log\log n}^{c\frac{\log n}{\log\log n}}}{n} =$$

$$\lim_{n\to\infty} \frac{2^{c\frac{\log n}{\log\log n}/2} n^c}{n}$$

Since $c < 1$ the claim follows. □

C Algorithm 3 Appendix

Lemma C.1. *Let $c < -\frac{1}{\log(\sin(30°))}$ be a constant, if $d = c\log n$ then $\lim_{n\to\infty} \frac{n'}{\frac{n}{k}} = 0$.*

Proof. Note that $n' = max\left(\frac{4}{\mu(CR(30°))}\log\left(\frac{2}{\gamma}\right), \frac{8d}{\mu(CR(30°))}\log\left(\frac{8d}{\mu(CR(30°))}\right)\right)$. One can check that in terms of asymptotic behavior it suffices to show that $\lim_{n\to\infty} \frac{\frac{1}{\mu(CR(30°))}}{n} = 0$. We use Fact 1.8

$$\lim_{n\to\infty} \frac{(1 - \cos(30°)^2)^{-\frac{1}{2}d}}{n} =$$

$$\lim_{n\to\infty} \frac{\left(\sin(30°)^2\right)^{-\frac{1}{2}c\log n}}{n} =$$

$$\lim_{n\to\infty} \frac{n^{-\frac{1}{2}c\log\left(\sin(30°)^2\right)}}{n}$$

Therefore if $c\log\left(\frac{1}{\sin(30°)}\right) < 1 \Rightarrow c < -\frac{1}{\log(\sin(30°))} = 1.4426950$ then the claim follows. □

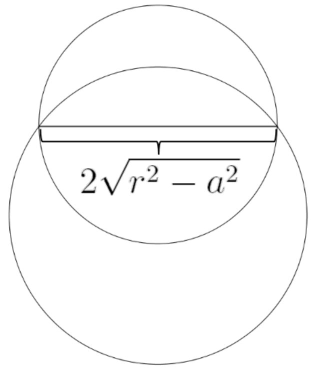


Figure 9: Proof of Lemma D.1

D Intersection of sphere caps appendix

Lemma D.1. *Let C_a be a sphere cap of $\mathbb{S}_r^d$, where $a > 0$. Let C_b be a sphere cap of $\mathbb{S}_{r'}^d$, where $b > 0$. If $r < r'$ and $\sqrt{r^2 - a^2} = \sqrt{r'^2 - b^2}$ then the surface area of C_a is larger than this of C_b.*

Proof. See Figure 9 □

Corollary D.2. *Let CC_a be a sphere cap of B_r^d, where $a > 0$. Let CC_b be a sphere cap of $B_{r'}^d$, where $b > 0$. If $r < r'$ and $\sqrt{r^2 - a^2} = \sqrt{r'^2 - b^2}$ then the volume of CC_a is larger than this of CC_b.*

Lemma D.3. *Let $a > 0$. The ratio between the volume of C_a^d and the surface area of C_a^{d+2} is $O(1)$*

Proof. Let $a > 0$ and let x be $(\sin \arccos a)^2$. $\mathrm{I}_c(a, b)$ is the regularized incomplete beta function . By Li [12] The volume of a sphere cap C_a^d is given by

$$\frac{1}{2} \frac{\pi^{d/2}}{\Gamma(d/2 + 1)} \mathrm{I}_x \left(\frac{d+1}{2}, \frac{1}{2} \right)$$

Again by Li [12] the surface area of a sphere cap C_a^{d+2} is given by

$$\frac{1}{2}\frac{\pi^{(d+2)/2}}{\Gamma\left((d+2)/2\right)}\mathrm{I}_x\left(\frac{d+1}{2},\frac{1}{2}\right)$$

Therefore The ratio between the volume of C_a^d and the surface area of C_a^{d+2} is $O(1)$. □

D.1 Proof of Lemma 5.4

Proof. We show a lower bound on the volume of z_2. Note that $H_b(\vec{v}_1)\cap B^d$ is a $d-1$ ball with a radius smaller than 1. Note that z_2 is a sphere cap of $H_b(\vec{v}_1)\cap B^d$. Assume that after some translation $\vec{0}$ is the center of $H_b(\vec{v}_1)\cap B^d$. Let $\left(H_b(\vec{v}_1)\cap B^d\right)'$ be the ball with the smallest radius that contains $H_b(\vec{v}_1)\cap B^d$ in dimension $d+1$. Let z_2' be a cap of $\left(H_b(\vec{v}_1)\cap B^d\right)'$ with the same radius as z_2.

Let $V(z_2)$ be the volume of the cap z_2 and let $S(z_2')$ be the surface area of the cap z_2', by Lemma D.3.

$$V(z_2)=O(1)\,S(z_2')$$

Let r be the radius of $H_b(\vec{v}_1)\cap B^d$. Let a be such that $C_a^{d-1}(\vec{x})=z_2$ for some vector $\vec{x}$. Note that $\sqrt{r^2-a^2}=\sqrt{1-\|\bar{x}\|^2}$ and by Lemma D.1

$$O(1)\,S(z_2)\geq O(1)\,S\left(C_{\|\bar{x}\|}^{d+1}\right)$$

$$=O(1)\,\mu\left(C_{\|\bar{x}\|}^{d+1}\right)S\left(\mathbb{S}^{d-1}\right)$$

By lemma 5.3

$$\mu\left(C_b(\vec{v}_1)\cap C_b(\vec{v}_2)\right)\geq\frac{V(z_2)}{S(\mathbb{S}^d)}$$

$$\geq\frac{O(1)\,\mu\left(C_{\|\bar{x}\|}^{d+1}\right)S\left(\mathbb{S}^{d-1}\right)}{S(\mathbb{S}^d)}=\frac{O(1)}{\sqrt{d}}\mu\left(C_{\|\bar{x}\|}^{d+1}\right)$$

and Fact 1.8

$$\frac{O(1)}{\sqrt{d}}\mu\left(C_{\|\bar{x}\|}^{d+1}\right)\geq\frac{O(1)}{d}\left(\sqrt{1-\left(\frac{b}{\cos\left(\frac{\alpha}{2}\right)}\right)^2}\right)^d$$

□

E Uniquely colorable graphs Appendix

The size of the maximum independent set in G is denoted by $\alpha(G)$. Note that for every graph G it holds that $\chi(G) \geq \frac{n}{\alpha(G)}$. Let H be defined as in Section 4.2. The measure of the maximum independent set in H (which is an infinite graph) is also denoted by $\alpha(H)$ and it holds that $\chi(H) \geq \frac{1}{\alpha(H)}$.

Theorem E.1. *If* $\epsilon' < 90°$ *then* $\chi(H) \gg 6$.

Proof. Suppose for the sake of contradiction that $\chi(H) \leq 6$. By a simple variation Theorem 3.5 in Feige et al. [5] $\alpha(H) \leq \mu(CR(\epsilon'))$.

By a simple calculation $\mu(CR(\epsilon'))$ tends to zero as d grows. But if $\chi(H) \leq 6$ then $\alpha(H) \geq \frac{1}{6}$ which is a contradiction. □

F A note on the dimension range

Assume that the dimension d of our graphs $G'_{n,3,d}$ (the graphs before the modification) is $c\log(n)$. When c gets larger each vertex has fewer neighbors. We would like to determine the values of c for which $G \in G'_{n,\frac{1}{2},d}$ has isolated vertices with high probability. Let v be a vertex in G it holds that $E[|N(v)|] = (n-1)\mu\left(C_{1/2}\right)$.Note that by Fact 1.8 $\mu\left(C_{1/2}\right)$ can be upper bounded by $\frac{1}{2}\left(1-\frac{1}{2}^2\right)^{\frac{d-1}{2}}$. Note that:

$$(n-1)\mu\left(C_{1/2}\right) \leq n\frac{1}{2}\left(1-\frac{1}{2}^2\right)^{\frac{d-1}{2}} = n\frac{1}{1}e^{\log\left(\frac{3}{4}\right)\frac{d-1}{2}} =$$

$$= n\frac{1}{2}e^{\log\left(\frac{3}{4}\right)\frac{d}{2}}e^{-\log\left(\frac{3}{4}\right)\frac{1}{2}} = n\frac{1}{2}n^{\log\left(\frac{3}{4}\right)\frac{c}{2}}e^{-\log\left(\frac{3}{4}\right)\frac{1}{2}}$$

$$= e^{-\log\left(\frac{3}{4}\right)\frac{1}{2}}\frac{1}{2}n^{1+\log\left(\frac{3}{4}\right)\frac{c}{2}}$$

Therefore if $1+\log\left(\frac{3}{4}\right)\frac{c}{2} < 0 \Rightarrow c > \frac{2}{\log\left(\frac{4}{3}\right)} = 6.95212$ then by applying the Markov's inequality the probability that v is isolated tends to one.

We would like to determine the values of c for which $G \in G'_{n,\frac{1}{2},d}$ has no isolated vertices with high probability. Note that by Fact 1.8 $\mu\left(C_{1/2}\right)$ can be lower bounded by

$\frac{O(1)}{\sqrt{d}}\left(1-\frac{1}{2}^2\right)^{\frac{d-1}{2}}$. Note that:

$$(n-1)\,\mu\left(C_{1/2}\right) \geq n\frac{O\left(1\right)}{\sqrt{d}}\left(1-\frac{1}{2}^2\right)^{\frac{d-1}{2}} = n\frac{O\left(1\right)}{\sqrt{d}}e^{\log\left(\frac{3}{4}\right)\frac{d-1}{2}} =$$

$$= n\frac{O\left(1\right)}{\sqrt{d}}e^{\log\left(\frac{3}{4}\right)\frac{d}{2}}e^{-\log\left(\frac{3}{4}\right)\frac{1}{2}} = n\frac{O\left(1\right)}{\sqrt{d}}n^{\log\left(\frac{3}{4}\right)\frac{c}{2}}e^{-\log\left(\frac{3}{4}\right)\frac{1}{2}}$$

$$= e^{-\log\left(\frac{3}{4}\right)\frac{1}{2}}\frac{O\left(1\right)}{\sqrt{d}}n^{1+\log\left(\frac{3}{4}\right)\frac{c}{2}}$$

Therefore if $1+\log\left(\frac{3}{4}\right)\frac{c}{2} > 0 \Rightarrow c < \frac{2}{\log\left(\frac{4}{3}\right)} = 6.95212$ then by applying the Chernoff and the union bounds the probability that $\forall v \in G$ there are neighbors tends to 1.