Graphs with tiny vector chromatic numbers and huge chromatic numbers

Uriel Feige [*] Michael Langberg [†] Gideon Schechtman [‡]

Abstract

Karger, Motwani and Sudan (JACM 1998) introduced the notion of a vector coloring of a graph. In particular they show that every k-colorable graph is also vector k-colorable, and that for constant k, graphs that are vector k-colorable can be colored by roughly $\Delta^{1-2/k}$ colors. Here Δ is the maximum degree in the graph, and assumed to be of the order of n^δ for some $0 < \delta < 1$. Their results play a major role in the best approximation algorithms for coloring and for maximum independent set.

We show that for every positive integer k there are graphs that are vector k-colorable but do not have independent sets significantly larger than $n/\Delta^{1-2/k}$ (and hence cannot be colored with significantly less than $\Delta^{1-2/k}$ colors). For $k = O(\log n/\log\log n)$ we show vector k-colorable graphs that do not have independent sets of size $(\log n)^c$, for some constant c. This shows that the vector chromatic number does not approximate the chromatic number within factors better than $n/\text{polylog}n$.

As part of our proof, we analyze "property testing" algorithms that distinguish between graphs that have an independent set of size n/k, and graphs that are "far" from having such an independent set. Our bounds on the sample size improve previous bounds of Goldreich, Goldwasser and Ron (JACM 1998) for this problem.

1 Introduction

An independent set in a graph G is a set of vertices that induce a subgraph which does not contain any edges. The size of the maximum independent set in G is denoted by $\alpha(G)$. For an integer k, a k coloring of G is a function $\sigma : V \to [1 \dots k]$ which assigns colors to the vertices of G. A valid k coloring of G is a coloring in which each color class is an independent set. The chromatic number $\chi(G)$ of G is the smallest k for which there exists a valid k coloring of G.

Finding $\alpha(G)$ and $\chi(G)$ are fundamental NP-hard problems, closely related by the inequality $\alpha(G)\chi(G) \geq n$. Given G, the question of estimating the value of $\alpha(G)$ ($\chi(G)$) or finding large independent sets (small colorings) in G have been studied extensively. Let G be a graph of size n. Both $\chi(G)$ and $\alpha(G)$ can be approximated within a ratio of $O(\frac{n(\log\log n)^2}{\log^3 n})$ [Hal93, Fei02]. It is known

[*]Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot, Israel 76100. Supported in part by the Israel Science Foundation (grant no. 236/02). E-mail: `uriel.feige@weizmann.ac.il`

[†]Department of Computer Science, California Institute of Technology, Pasadena, California 91125. This work was done while studying in the Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot, Israel 76100. E-mail: `mikel@cs.caltech.edu`

[‡]Department of Mathematics, Weizmann Institute of Science, Rehovot, Israel 76100. Supported in part by the Israel Science Foundation (grant no. 154/01). E-mail: `gideon.schechtman@weizmann.ac.il`

that unless NP=ZPP, neither $\alpha(G)$ nor $\chi(G)$ can be approximated within a ratio of $n^{1-\varepsilon}$ for any $\varepsilon > 0$ [Hås99, FK98]. Under stronger complexity assumptions, there is some $0 < \delta < 1$ such that neither problem can be approximated within a ratio of $n/2^{\log^\delta n}$ [Kho01].

The approximation ratios for these problems significantly improve in graphs that have very large independent sets, or very small chromatic numbers. The algorithms achieving the best ratios in these cases [KMS98, AK98, BK97, HNZ01] are all based on the idea of *vector coloring*, introduced by Karger, Motwani and Sudan [KMS98].

Definition 1.1 *A vector k-coloring of a graph is an assignment of unit vectors to its vertices, such that for every edge, the inner product of the vectors assigned to its endpoints is at most (in the sense that it can only be more negative) $-1/(k-1)$.*

Every k-colorable graph is also vector k-colorable (by identifying each color class with one vertex of a perfect $(k-1)$-dimensional simplex centered at the origin). Moreover, unlike the chromatic number, a vector k-coloring (when it exists) can be found in polynomial time using semidefinite programming (up to arbitrarily small error in the inner products). Given a vector k-coloring of a graph, Karger, Motwani and Sudan show how to color a graph with roughly $\Delta^{1-2/k}$ colors, where Δ is the maximum degree in the graph. (In comparison, the technique of inductive coloring might use $\Delta + 1$ colors.) In fact, they show how to find an independent set of size roughly $n/\Delta^{1-2/k}$. Combined with other ideas, this leads to coloring algorithms and algorithms for finding independent sets with the best currently known performance guarantees. For example, there is a polynomial time algorithm that colors 3-colorable graphs with roughly $n^{3/14}$ colors [BK97]. For nonconstant values of k, it is known how to find an independent set of size $\Omega(\log n)$ in a vector $\log n$-colorable graph.

There have also been negative results regarding vector k-colorable graphs. Examples appearing in [KMS98] (and improved by Noga Alon and Mario Szegedy) show vector 3-colorable graphs that do not have independent sets larger than roughly $n^{0.95}$. The case of nonconstant k is addressed in [Fei97] (technically, the results there deal with the Lovász theta function, which is even a stronger notion than vector coloring), where graphs that are vector $2^{O(\sqrt{\log n})}$-colorable are shown not to have independent sets larger than $2^{O(\sqrt{\log n})}$. In all these negative examples, the vertex sets of the graphs involved can be viewed as a subset of $\{0,1\}^n$, with two vertices connected by an edge if their Hamming distance is larger than some prespecified value.

Our results In this work we present a different family of graphs with stronger negative properties. For every constant $k > 2$ and every $\varepsilon > 0$, we show graphs that are vector k-colorable, with $\alpha(G) \leq n/\Delta^{1-\frac{2}{k}-\varepsilon}$. This essentially matches the positive results of [KMS98]. As a function of n rather than Δ, we show vector 3-colorable graphs with $\alpha(G) < n^{0.843}$. Moreover, for $k = O(\log n/\log\log n)$, we show vector k-colorable graphs with $\alpha(G) \leq (\log n)^c$, for some universal c. This shows that the vector coloring number by itself does not approximate the chromatic number within a ratio better than $n/\text{polylog}n$. Another consequence of this (that is touched upon in Remark 2.1 in Section 2) is that certain semidefinite programs do not approximate the size of the maximum independent set with a ratio better than $n/\text{polylog}n$.

Theorem 1.2

1. *For every constant $\varepsilon > 0$ and constant $k > 2$, there are infinitely many graphs G that are vector k-colorable and satisfy $\alpha(G) \le \frac{n}{\Delta^{1-\frac{2}{k}-\varepsilon}}$, where n is the number of vertices in G, Δ is the maximum degree in G, and $\Delta > n^{\delta}$ for some constant $\delta > 0$.*

2. *For some constant c, there are infinitely many graphs G that are vector $O(\frac{\log n}{\log\log n})$-colorable and satisfy $\alpha(G) \le (\log n)^c$.*

3. *There are infinitely many graphs G which are vector 3-colorable and satisfy $\alpha(G) \le n^{0.843}$.*

Observe that if Theorem 1.2 (1) is proved for some graph with n vertices and maximum degree Δ, then for every positive integer c it also holds for some graph with $c \cdot n$ vertices and the same maximum degree Δ. (Simply make c disjoint copies of the original graph.) Hence the theorem becomes stronger as Δ becomes larger in comparison to n. We guarantee that Δ can be taken at least as large as n^{δ} for some δ that depends on ϵ and k. For every fixed ϵ and k, the value of δ can be taken as a fixed constant bounded away from 0, and independent of n. We have not made an attempt to find the tightest possible relation between δ and ϵ and k.

Proof techniques The graphs that we use are essentially the same graphs that were used in [FS02] to show integrality gaps for semidefinite programs for Max-Cut. Namely, they are obtained by placing n points at random on a d-dimensional unit sphere, and connecting two points by an edge if the inner product of their respective vectors is below $-1/(k-1)$. Such graphs are necessarily vector k-colorable, as the embedding on the sphere is a vector k-coloring. So the bulk of the work is in proving that they have no large independent set. For this we use a two phase plan. First we consider a continuous (infinite) graph, where every point on the sphere is a vertex, and two vertices are connected by an edge if the inner product of their respective vectors is below $-1/(k-1)$. On this continuous graph we use certain symmetrization techniques in order to analyze its properties. Specifically, we prove certain inequalities regarding its expansion. In the second phase, we consider a (finite) random vertex induced subgraph of the continuous graph. Based on the expansion properties of the continuous graph, we show that the random sample has no large independent set.

Two remarks are in order here. One is that it is very important for our bounds that the final random graph does not contain too many vertices compared to the dimension d. A small number of vertices implies low degree, and allows for a more favorable relation between the maximum degree and the chromatic number. For this reason we cannot use the continuous graph (or a finite discrete approximation of the continuous graph) as is. Its degree is very large compared to its chromatic number. We thus consider the graph obtained by randomly sampling the vertices of the continuous graph. In Section 7, we follow a suggestion of Luca Trevisan (private communication) and present an alternative proof of Theorem 1.2 (1) by considering the graph obtained by randomly sampling the *edges* of the continuous graph. It appears that parts (2) and (3) of Theorem 1.2 can not be proven using edge sampling.

The other remark is that we do not get an explicit graph as our example, but rather a random graph (or a distribution on graphs). This is to some extent unavoidable, given that there are no known efficient deterministic constructions of Ramsey graphs (graphs in which the size of the maximum independent set and maximum clique are both bounded by some polylog in n). The graphs we construct (when $k = \log n / \log\log n$) are Ramsey graphs, because it can be shown that the maximum clique size is never larger than the vector coloring number.

Property testing The following problem in property testing is addressed by Goldreich, Goldwasser and Ron [GGR98]. For some value of $\rho < 1$, consider a graph with the following "promise": either it has an independent set of size ρn, or it is far from any such graph, in the sense that any vertex induced subgraph of ρn vertices induces at least εn^2 edges. One wants an algorithm that samples as few vertices as possible, looks at the subgraph induced on them, and based on the size of the maximum independent set in it decides correctly (with high probability) which of the two cases above hold. In [GGR98] it is shown that a sample of size proportional to ε^{-4} suffices. We are in a somewhat similar situation when we move from the continuous graph to our random sample. The continuous graph is far from having an independent set of *measure* ρ [1]. We want to take a small as possible sample (its size will be denoted by s) such that the induced subgraph does not contain an independent set of size ρs. Luckily in our case, we can use a stronger guarantee on the continuous graph. We know that even every set of measure $\rho/2$ induces at least the measure of ε edges [2]. In this case we show that a sample size s proportional to $1/\varepsilon$ suffices. This dramatic improvement over the [GGR98] bound is crucial to the success of our analysis. We note that this improvement is not only based on our stronger guarantee on the continuous graph G. Even in exactly the same setting of [GGR98], we show that a sample size s proportional to ε^{-3} suffices. These results are of interest in the context of property testing regardless of their applications to the vector coloring issue.

Strict vector coloring One may strengthen the notion of vector k-coloring by requiring that for every edge, the inner product of the unit vectors corresponding to its endpoints be exactly $-1/(k-1)$, rather than at most $-1/(k-1)$. This is called a strict vector coloring. This notion is known to be equivalent to the theta function of Lovász [Lov79, KMS98]. Every k-colorable graph is also strictly vector k-colorable. As strict vector coloring is a stronger requirement than vector coloring, then potentially, strict vector k-colorable graphs have smaller chromatic numbers than vector k-colorable graphs. So far, there has not been any algorithmic technique that could use this observation to further improve the approximation ratios for chromatic number or for independent set. We remark however that the negative results in this work apply only to vector coloring and not to strict vector coloring. It is an open question whether similar negative results are true for strict vector coloring, or equivalently, whether the same gaps (such as $n/\text{polylog}n$) can be shown between the value of the theta function and the size of the maximum independent set. Note that the weaker negative results of [Fei97] and some of the negative results in [KMS98] do apply also to strict vector coloring.

The remainder of this paper is organized as follows. In Section 2 we briefly review the semidefinite programs that compute the vector chromatic number and its variants. In Section 3 we present the continuous graph and analyze its properties. In Section 4 we prove Theorem 1.2. Our results on property testing are presented in Sections 5 and 6. In Section 7 we present an alternative proof of Theorem 1.2 (1). In Section 8 we discuss some problems that remain open. Finally, in Section 9 we present the proof of several technical Lemmas that appear throughout our work.

[1] As the continuous graph is infinite, terms such as the *number* of vertices are replaced by the continuous analog *measure*. The measure we use on the unit sphere is the natural measure which associates with each subset its relative area on the sphere.

[2] The measure we use on the edge set is the product measure induced by the standard uniform measure on S^{d-1}.

2 The vector chromatic number and its variants

There are many equivalent ways to define the vector chromatic number and its variants. We will follow the definitions suggested in [KMS98, Cha02]. Let $G = (V, E)$ be a graph of size n. For convenience we will assume that $V = [1, \ldots, n]$. The semidefinite relaxations below assign unit vectors to every vertex $i \in V$. These unit vectors are to satisfy certain constraints which will in turn determine the value of the relaxations.

$COL_1(G)$ Minimize k
subject to:

$$\begin{array}{ll} \langle v_i, v_j \rangle \le -\frac{1}{k-1} & \forall (i,j) \in E \\ \langle v_i, v_i \rangle = 1 & \forall i \in V \end{array}$$

$COL_2(G)$ Minimize k
subject to:

$$\begin{array}{ll} \langle v_i, v_j \rangle = -\frac{1}{k-1} & \forall (i,j) \in E \\ \langle v_i, v_i \rangle = 1 & \forall i \in V \end{array}$$

$COL_3(G)$ Minimize k
subject to:

$$\begin{array}{ll} \langle v_i, v_j \rangle = -\frac{1}{k-1} & \forall (i,j) \in E \\ \langle v_i, v_j \rangle \ge -\frac{1}{k-1} & \forall i, j \in V \\ \langle v_i, v_i \rangle = 1 & \forall i \in V \end{array}$$

The function $COL_1(G)$ is the vector chromatic number of G as defined in [KMS98]. The function $COL_2(G)$ is the strict vector chromatic number of G and is equal to the Lovász θ function on $\bar{G}$ [Lov79, KMS98], where $\bar{G}$ is the *complement* graph of G. Finally, the function $COL_3(G)$ will be referred to as the *strong* vector chromatic number as defined in [Sze94, Cha02]. Let $\omega(G)$ denote the size of the maximum clique in G, in the following we show that

$$\omega(G) \le COL_1(G) \le COL_2(G) \le COL_3(G) \le \chi(G)$$

It is not hard to verify that $COL_1(G) \le COL_2(G) \le COL_3(G)$. To show the other inequalities we need the following fact. For every integer k, the k unit vectors $\{v_1, \ldots, v_k\}$ that minimize the value of $\max_{i \ne j \in [1,\ldots,k]} \langle v_i, v_j \rangle$ are the vertices of a simplex in R^{k-1} centered at the origin. For each $i, j \in [1, \ldots, k]$, these vectors satisfy $\langle v_i, v_j \rangle = -\frac{1}{k-1}$.

Now to prove the inequality $COL_3(G) \le \chi(G)$, consider a k coloring σ of G. The coloring σ partitions the vertex set V into k color classes $\{V_1, \ldots, V_k\}$. Assigning each color class V_i with the corresponding vector v_i above, we obtain a valid assignment for $COL_3(G)$. To show that $\omega(G) \le COL_1(G)$, consider a graph G with maximum clique size $\omega(G)$. To obtain a valid assignment of vectors to $COL_1(G)$ of value k we require that all pairs of vectors corresponding to the vertices of the maximum clique will have inner product of value at most $-\frac{1}{k-1}$. As mentioned above, this can happen only if $\omega(G) \le k$.

Remark 2.1 *The results of our work show a large gap between $COL_1(G)$ and $\chi(G)$ (Theorem 1.2). Combining these result with certain proof techniques appearing in [Sze94], a similar gap between $\omega(G)$ and $COL_3(G)$ can also be obtained. Details omitted.*

3 The continuous graph

Let d be a large constant, and let $S^{d-1} = \{v \in R^d \mid \|v\| = 1\}$ be the d dimensional unit sphere. Let $G_k = (V, E)$ be the continuous graph in which: *(a)* The vertex set V consists of all the points on the unit sphere S^{d-1}. *(b)* The edge set E of G_k consists of all pairs of vertices whose respective vectors (from the origin) form an angle of at least $\arccos(-1/(k-1))$. As the size of V and E is infinite (and uncountable), terms such as the *number* of vertices in V will be replaced by the continuous analogue *measure*.

In this section we analyze various properties of G_k. Specifically, we show that G_k has certain expansion properties. We then use this fact in Section 4 to prove the main theorem of our work. In our analysis, we will assume that the dimension d is (at least) a very large constant (our proofs rely on such d). Additional constants that will be presented in the remainder of this section are to be viewed as independent of d.

Definition 3.1 (Sphere measure) *Let μ be the normalized $(d-1)$ dimensional natural measure on S^{d-1}, and let μ^2 be the induced measure on $S^{d-1} \times S^{d-1}$. For any two (not necessarily disjoint) subsets A and B of V, we define the measure of edges from A to B as*

$$E(A, B) = \mu^2 \left(\{(x, y) \mid x \in A,\ y \in B,\ (x, y) \in E\}\right).$$

Definition 3.2 (Sphere caps) *Let $a \in [0, 1]$, and $x \in S^{d-1}$. An a-cap centered at x is defined to be the set $\mathcal{C}_a = \{u \in S^{d-1} \mid \langle u, x\rangle \geq a\}$. Denote the measure of an a-cap by $\rho(a)$.*

A few remarks are in place. Notice that for every x, x', an a-cap centered at x has the same measure as an a-cap centered at x' (sphere symmetry). Furthermore, notice that *large* caps have *small* corresponding values of a and vice versa. The value of $\rho(a) = \mu(\mathcal{C}_a)$ is approximately given by the following lemma proven for example in [FS02].

Lemma 3.3 *Let $\mathcal{C}_a$ be an a-cap centered at some $x \in S^{d-1}$. There exists a constant $c > 0$ (independent of a and d) s.t.*

$$\frac{c}{\sqrt{d}} \left(1 - a^2\right)^{\frac{d-1}{2}} \leq \mu(\mathcal{C}_a) = \rho(a) \leq \frac{1}{2} \left(1 - a^2\right)^{\frac{d-1}{2}}.$$

Lemma 3.4 *Let $\rho(\frac{1}{k-1})$ be the measure of a $\frac{1}{(k-1)}$-cap. Every vertex v in the graph G_k has degree $\rho(\frac{1}{k-1})$.*

Proof: Each vertex v in G_k is adjacent to every vertex in the $\frac{1}{(k-1)}$-cap centered at $-v$ (here $-v$ is the vertex antipodal to v). Hence, the measure of vertices adjacent to a given vertex v in G_k is that of a $\frac{1}{(k-1)}$-cap. $\square$

The main property of G_k of our interest is the measure of edges between any two given subsets of V of a specified size. We first prove that the sets in G_k which share the least amount of edges are caps with the same center.

Theorem 3.5 *Let $1 \geq a > 0$ and let A and B be two (not necessarily disjoint) measurable sets in V of measure $\rho(a)$. Let x be an arbitrary vertex of S^{d-1}. The minimum of $E(A, B)$ is obtained when $A = B = \mathcal{C}_a$ where $\mathcal{C}_a$ is an a-cap of measure $\rho(a)$ centered at x.*

The proof of Theorem 3.5 is based on symmetrization techniques similar to those presented in [BT76, FS02]. We prove Theorem 3.5 in Section 9. We now turn to analyze the measure of edges between caps of measure $\rho(a)$. Namely we study the value of $E(\mathcal{C}_a, \mathcal{C}_a)$. For *large* values of a, it is not hard to see that $\mathcal{C}_a$ will not induce any edges. This follows from the fact that any two vertices u and v in $\mathcal{C}_a$ satisfy $\langle u, v\rangle \geq 2a^2 - 1$. Hence, in the following we consider values of a which satisfy $2a^2 - 1 < -\frac{1}{k-1}$. For such values of a we show that $E(\mathcal{C}_a, \mathcal{C}_a)$ is proportional to $\mu^2(E)\rho(a)^{2+\frac{2}{k-2}}$ (which we denote by $\lambda(a)$). Notice that two random subsets A and B of G_k of measure $\rho(a)$ are expected to satisfy $E(A, B) = \mu^2(E)\rho(a)^2$ which is fairly close to $\lambda(a)$. Actually, we show that $E(\mathcal{C}_a, \mathcal{C}_a)$ is in the range $\left[\left(1 - c\sqrt{\frac{\log(d)}{d}}\right)^{\frac{d-1}{2}} \lambda(a), \lambda(a)\right]$ for some large constant c. It is not hard to verify that the term $\left(1 - c\sqrt{\frac{\log(d)}{d}}\right)^{\frac{d-1}{2}}$ is negligible compared to $\lambda(a)$ once d is taken to be large enough.

Theorem 3.6 *Let* $\sqrt{\frac{k-2}{2(k-1)}} > a > 0$ *and* $k > 2$ *be constant. Let* $x \in S^{d-1}$*, let* $\mathcal{C}_a$ *be an* a*-cap centered at* x*. Let* $\varepsilon(a)$ *be the value of* $E(\mathcal{C}_a, \mathcal{C}_a)$*. Finally let*

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

Then $\varepsilon(a) \in \left[\left(1 - c\sqrt{\frac{\log(d)}{d}}\right)^{\frac{d-1}{2}} \lambda(a), \lambda(a)\right]$ *for some constant* $c > 0$*.*

Proof: Let $x \in S^{d-1}$. Let $\mathcal{C}_a$ be an a-cap centered at x. W.l.o.g. we will assume that $x = (1, 0, \ldots, 0)$. Consider a vertex $v \in \mathcal{C}_a$ on the boundary of $\mathcal{C}_a$. Let $N(v)$ be the set of vertices adjacent to v. We start by computing the measure of vertices that are neighbors of v and are in the cap $\mathcal{C}_a$, *i.e.* the measure of $N(v) \cap \mathcal{C}_a = \{u = (u_1, \ldots, u_d) \in S^{d-1} \mid u_1 \geq a \text{ and } \langle v, u\rangle \leq -1/(k-1)\}$.

Claim 3.7 *Let* a, k *be as in Theorem 3.6. Let* $v = (a, \sqrt{1-a^2}, 0, \ldots, 0)$ *be a vertex on the boundary of* $\mathcal{C}_a$*. Let* $N(v)$ *be the set of neighbors of* v*. Let* $z = \sqrt{a^2 + \frac{(1/(k-1)+a^2)^2}{1-a^2}}$*. Finally let* $\delta = c\sqrt{\frac{\log(d)}{d}}$ *for a sufficiently large constant* c*. The measure of vertices in* $N(v) \cap \mathcal{C}_a$ *satisfies*

$$(1-\delta)^{\frac{d-1}{2}}\left(1 - z^2\right)^{\frac{d-1}{2}} \leq \mu(N(v) \cap \mathcal{C}_a) \leq \left(1 - z^2\right)^{\frac{d-1}{2}}$$

Claim 3.7 addresses the measure of $\mathcal{C}_a \cap N(v)$, and states that it is essentially the measure of a z-cap (where $z = \sqrt{a^2 + \frac{(1/(k-1)+a^2)^2}{1-a^2}}$). To prove Claim 3.7 we study the measure of certain restricted sets in S^{d-1}. These sets are studied in Claim 9.7 of Section 9. Claims 3.7 and 9.7 are depicted in Figure 1 and proven in Section 9.

To complete the proof of Theorem 3.6, let $a,\ z,\ \delta$ be as in Claim 3.7. For a vertex $v \in \mathcal{C}_a$ let $N(v) \cap \mathcal{C}_a$ be the set of vertices adjacent to v in $\mathcal{C}_a$. For the upper bound, notice that of all vertices in $\mathcal{C}_a$, the vertices v in which $\mu(N(v) \cap \mathcal{C}_a)$ is largest are the vertices on the boundary of $\mathcal{C}_a$. By Claim 3.7, for these vertices $\mu(N(v) \cap \mathcal{C}_a)$ is bounded by $(1-z^2)^{\frac{d-1}{2}}$. We thus conclude, that $\varepsilon(a)$ is bounded by the measure of vertices in $\mathcal{C}_a$ times $(1-z^2)^{\frac{d-1}{2}}$. That is,

$$\varepsilon(a) \leq \rho(a)(1-z^2)^{\frac{d-1}{2}} \leq (1-a^2)\left(1 - \left(a^2 + \frac{(1/(k-1)+a^2)^2}{1-a^2}\right)\right)^{\frac{d-1}{2}} = \lambda(a).$$

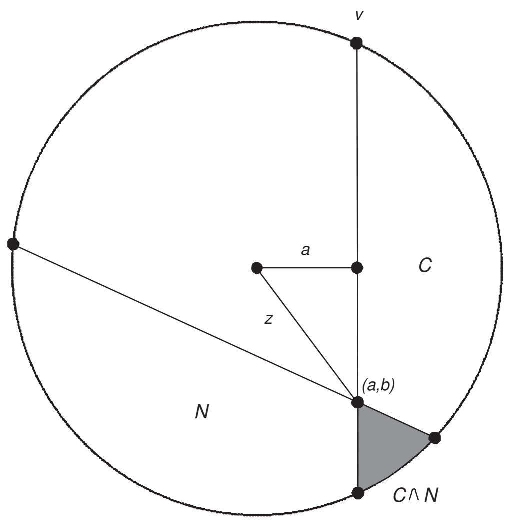

Figure 1: Projecting S^{d-1} onto the two dimensional subspace $\{(r_1, r_2, 0, \ldots, 0) \mid r_1, r_2 \in R\}$ of R^d we obtain the circle above. C is the projection of an a-cap $\mathcal{C}_a$ centered at $(1, 0, \ldots, 0)$ (where a is the distance of the cap from the origin). The vertex $v = (a, \sqrt{1-a^2}, 0, \ldots, 0) \in \mathcal{C}_a$ is on the boundary of $\mathcal{C}_a$. N is the projection of the set of points $N(v)$ on the sphere that are adjacent to v (*i.e.* form an angle of at least $\arccos(-1/(k-1))$ with v). The shaded section is $C \cap N$ (the projection of $\mathcal{C}_a \cap N(v)$). The point (a, b) is the closest point of the projection of $\mathcal{C}_a \cap N(v)$ to the origin. It is not hard to verify that $b^2 = \frac{(1/(k-1)+a^2)^2}{1-a^2}$. Finally, we denote the value of $\sqrt{a^2+b^2}$ by z. Claim 3.7 addresses the measure of $\mathcal{C}_a \cap N(v)$, and states that it is essentially the measure of a z-cap. This is done by studying the points in S^{d-1} who's projection falls *close* to (a, b). Roughly speaking, we first show that such points are in $\mathcal{C}_a \cap N(v)$, then using Claim 9.7 we show that the measure of these points is essentially the measure of a z-cap.

As for the lower bound, let $w = (w_1, w_2, \ldots, w_d)$ be a vertex in $\mathcal{C}_a$ with first coordinate w_1 of value $a + \delta$. Consider any vertex $v = (v_1, v_2, \ldots, v_d) \in \mathcal{C}_a$ with first coordinate v_1 of value less than $a + \delta$. It is not hard to verify that the measure of $N(v) \cap \mathcal{C}_a$ is greater than the measure of $N(w) \cap \mathcal{C}_a$. Using an analysis similar to that of Claim 3.7 we have that $\mu(N(w) \cap \mathcal{C}_a)$ is greater or equal to

$$(1-c\delta)^{\frac{d-1}{2}} \left(1 - \left(a^2 + \frac{(1/(k-1) + a(a+\delta))^2}{1-(a+\delta)^2}\right)\right)^{\frac{d-1}{2}} \geq (1-c\delta)^{\frac{d-1}{2}} \left(1 - z^2\right)^{\frac{d-1}{2}}$$

for a sufficiently large constant c (which changes values between both sides of the second inequality). Furthermore, for our choice of δ, the measure of vertices $v = (v_1, v_2, \ldots, v_d)$ in $\mathcal{C}_a$ with $v_1 \leq a + \delta$ is at least $\rho(a)/2$ (Claim 9.3). Hence, we conclude that $E(\mathcal{C}_a, \mathcal{C}_a)$ is at least $\frac{\rho(a)}{2}(1-c\delta)^{\frac{d-1}{2}} \left(1-z^2\right)^{\frac{d-1}{2}}$. Simplifying the above expression, we conclude our assertion. $\square$

Theorem 3.6 addresses the case in which k is constant and the caps considered are both of measure $\rho(a)$ for a constant value of a. For the proof of Theorem 1.2 (2) we also need to address nonconstant values of a and k which depend on d.

Theorem 3.8 *Let $a = \left(\frac{\log(d)}{d}\right)^{\frac{1}{4}}$. Let k satisfy $\frac{1}{k-1} = a^2$. Let $x \in S^{d-1}$, let $\mathcal{C}_a$ be an a-cap centered at x. Let $\varepsilon(a)$ be the value of $E(\mathcal{C}_a, \mathcal{C}_a)$. The value of $\varepsilon(a)$ is in the range*

$$\left[\frac{1}{poly(d)}\left(1 - \frac{2(k-1)}{k-2}a^2\right)^{\frac{d-1}{2}}, \left(1 - \frac{2(k-1)}{k-2}a^2\right)^{\frac{d-1}{2}}\right].$$

The outline of the proof of Theorem 3.8 is similar to that of Theorem 3.6. Full proof appears in Section 9.

Definition 3.9 *Let $\rho < 1$. A graph $G = (V, E)$ is said to be pairwise $\langle\rho, \varepsilon\rangle$-connected iff every two (not necessarily disjoint) subsets A and B of V of measure ρ satisfy $E(A, B) \geq \varepsilon$.*

Combining Theorems 3.5, 3.6, 3.8 we obtain:

Corollary 3.10 *Let a, k, $\varepsilon(a)$ be defined as in Theorem 3.6 or 3.8. The graph G_k is pairwise $\langle\rho(a), \varepsilon(a)\rangle$-connected.*

Roughly speaking Corollary 3.10 addresses the expansion properties of the continuous graph G_k. In Section 5, we show that these properties imply certain upper bounds on the independence number of a small random sample of G_k. Namely we prove

Theorem 3.11 *Let a, k, $\varepsilon(a)$ be defined as in Theorem 3.6 or 3.8. Let H be a random sample of s vertices of G_k (according to the uniform distribution on S^{d-1}). Let c be a sufficiently large constant. If $s \geq c\frac{\rho(a)}{\varepsilon(a)}\log^2(1/\rho(a))$ then the probability that $\alpha(H) > e^2\rho(a)s$ is at most $1/4$.*

In the following section, Theorem 3.11 is used to prove the main result of this work, Theorem 1.2. The proof of Theorem 3.11 will be presented in Section 5.2.

4 Proof of Theorem 1.2

In this section we prove Theorem 1.2. Recall that we are looking for a graph H for which both the vector chromatic number and the size of the maximum independent set are small. The graphs H we present are random subgraphs of the graphs G_k defined in Section 3. The three assertions of Theorem 1.2 are all proven similarly, the main difference between their proof is the choice of parameters used. To avoid confusion, we restate Theorem 1.2 using a slightly different notation then that appearing in the original presentation.

Theorem 1.2 (restated)

1. *For every constant $\gamma > 0$ and constant $k > 2$, there are infinitely many graphs H that are vector k-colorable and satisfy $\alpha(H) \leq \frac{s}{\Delta_H^{1-\frac{2}{k}-\gamma}}$, where s is in the number of vertices in H, Δ_H is the maximum degree in H, and $\Delta_H > s^\delta$ for some constant $\delta > 0$.*
2. *For some constant c, there are infinitely many graphs H of size s that are $O(\frac{\log s}{\log\log s})$ vector colorable and satisfy $\alpha(H) \leq (\log s)^c$.*

3. *There are infinitely many graphs H of size s which are vector 3-colorable and satisfy $\alpha(H) \leq s^{0.843}$.*

Proof: (Theorem 1.2 (1)) Let $k > 2$ be constant. Let $\gamma > 0$ be an arbitrarily small constant. Let c be a sufficiently large constant and let $a = \gamma/c$. Let $G = G_k = (V, E)$ be the continuous graph from Section 3 (here and in the remainder of the proof, we assume the dimension d of the graph G_k is taken to be significantly larger than $1/\gamma$). Finally, let $\Delta = \rho(\frac{1}{k-1})$ be the measure of vertices adjacent to any given vertex of G.

Recall (Corollary 3.10) that G is pairwise $\langle \rho(a), \varepsilon(a) \rangle$-connected. Let $\rho = \rho(a)$ and $\varepsilon = \varepsilon(a)$. This implies (Theorem 3.11) that with probability $\geq 3/4$ a random subset H of G of size $s \geq c\frac{\rho}{\varepsilon}\log^2(1/\rho)$ satisfies $\alpha(H) \leq e^2 \rho s$ (recall that c is a sufficiently large constant). We start by simplifying the expression bounding s.

Claim 4.1 *A random subset H of G of size $s = \frac{1}{\Delta\rho^{\frac{k}{k-2}+\gamma}}$ satisfies $\alpha(H) \leq e^2\rho s = \frac{e^2}{\Delta\rho^{\frac{2}{k-2}+\gamma}}$ with probability $\geq 3/4$.*

Proof: It suffices to prove that $s = \frac{1}{\Delta\rho^{\frac{k}{k-2}+\gamma}} \geq c\frac{\rho}{\varepsilon}\log^2(1/\rho)$. The claim follows (by basic calculations) from the fact that ε can be bounded by $\Delta\rho^{\frac{2(k-1)}{k-2}+\gamma}$. By Theorem 3.6, we have that

$$\varepsilon(a) = \left[\left(1 - c\left(\sqrt{\frac{\log(d)}{d}}\right)\right)\left(1 - \frac{1}{(k-1)^2}\right)\left(1 - \frac{2(k-1)}{k-2}a^2\right)\right]^{\frac{d-1}{2}}$$

It is not hard to verify that $\left(1 - c\left(\sqrt{\frac{\log(d)}{d}}\right)\right)^{\frac{d-1}{2}} > \rho^{\gamma/2}$. Furthermore, by Lemma 3.3 we have that $\left(1 - \frac{1}{(k-1)^2}\right)^{\frac{d-1}{2}} > \Delta$. Hence, $\varepsilon(a) \geq \rho^{\gamma/2}\left(1 - \frac{2(k-1)}{k-2}a^2\right)^{\frac{d-1}{2}}\Delta$. Recall that a was defined as γ/c for a sufficiently large constant c. This implies that $\gamma \geq 2\frac{a^2\left(\frac{2(k-1)}{k-2}\right)^2}{1-a^2\frac{2(k-1)}{k-2}}$ ($k > 2$ is constant). This expression is designed to fit the requirement appearing in Claim 9.1 (of Section 9). Now by Claim 9.1, it holds that $\left(1 - \frac{2(k-1)}{k-2}a^2\right)^{\frac{d-1}{2}} \geq \rho^{\frac{2(k-1)}{k-2}+\gamma/2}$. We conclude that $\varepsilon(a) \geq \Delta\rho^{\frac{2(k-1)}{k-2}+\gamma}$. □

Let H be a random subgraph of G of size $s = \frac{1}{\Delta\rho^{\frac{k}{k-2}+\gamma}}$. We will show that H satisfies the asserted conditions with probability greater than 1/2. First notice that any subgraph of G is vector k-colorable, including the subgraph H. Secondly, by Claim 4.1, $\alpha(H) \leq \frac{e^2}{\Delta\rho^{\frac{2}{k-2}+\gamma}}$ with probability $\geq 3/4$. It is left to analyze the maximum degree of H.

Claim 4.2 *With probability greater than 3/4, the maximum degree Δ_H of the subgraph H is in the range $[\frac{1}{2}\Delta s, 2\Delta s]$.*

Proof: Consider a vertex $h \in H$. Let d_h be the degree of h. As every vertex in G is of degree Δ, the expected value of d_h is $\Delta(s-1)$. Thus (using standard bounds) the probability that d_h deviates from its expectation by more than a constant fraction of its expectation is at most $2^{-\Omega(\Delta s)}$. The probability that some vertex in H has degree $\notin [\frac{1}{2}\Delta s, 2\Delta s]$ is thus at most $2^{\log(s)-\Omega(\Delta s)} \leq 3/4$ for our choice of s. □

The first assertion of Theorem 1.2 now follows using basic calculations. $\square$

Proof: (Theorem 1.2 (2)) Let d be a large constant. Let $k-1=\sqrt{\frac{d}{\log(d)}}$. Let $a^2=\frac{1}{k-1}$. Let $G=G_k=(V,E)$ be the continuous graph from Section 3.

Recall (Corollary 3.10) that G is pairwise $\langle\rho(a),\varepsilon(a)\rangle$-connected. Let $\rho=\rho(a)$ and $\varepsilon=\varepsilon(a)$. This implies (Theorem 3.11) that with probability $\geq 3/4$ a random subset H of size $s\geq c\frac{\rho}{\varepsilon}\log^2(1/\rho)$ satisfies $\alpha(H)\leq e^2\rho s$ (here c is a sufficiently large constant). As before we start by simplifying the expression bounding s.

Claim 4.3 *There exists a constant γ s.t with probability at least $3/4$ a random set H of G of size $s=\frac{d^\gamma}{\rho}\log^2(1/\rho)$ satisfies $\alpha(H)\leq e^2\rho s$ with probability $\geq 3/4$.*

Proof: Recall that $k-1=\sqrt{\frac{d}{\log(d)}}$, $a^2=\frac{1}{k-1}$. As before it suffices to bound $\varepsilon=\varepsilon(a)$. By Theorem 3.8, we have

$$\varepsilon\geq\frac{1}{poly(d)}\left(1-a^2\frac{2(k-1)}{k-2}\right)^{\frac{d-1}{2}}.$$

Furthermore using Claims 9.1 and 9.2 (of Section 9) we obtain,

$$\left(1-a^2\frac{2(k-1)}{k-2}\right)^{\frac{d-1}{2}}=\left(1-a^2\left(2+\frac{2}{k-2}\right)\right)^{\frac{d-1}{2}}\geq\frac{1}{poly(d)}\left(1-a^2\right)^{d-1}\geq\frac{\rho^2}{poly(d)}.$$

We conclude that there exists a constant γ such that $\varepsilon\geq\frac{\rho^2}{d^\gamma}$. $\square$

Let H be a random subset of vertices of G of size $s=\frac{d^\gamma}{\rho}\log^2(1/\rho)$. Notice that $\log(s)=\theta(\sqrt{d\log d})$ and $s\leq\theta(\frac{d^{\gamma+2}}{\rho})$. By definition, G is k vector colorable. This implies that any subgraph of G (including that induced by H) is $k=O\left(\frac{\log(s)}{\log(\log(s))}\right)$ vector colorable, which completes the proof of the first part of our assertion. For the second part of our assertion, by Claim 4.3 the subset H does not have an independent set of size $e^2\rho s\leq\log^\gamma(s)$ with probability at least $3/4$ (for some different constant γ). $\square$

Proof: (Theorem 1.2 (3)) The proof follows the line of proof appearing above. In general, we use the graph $G=G_3$, but this time the value of a is set to be $a=0.36$. Again, G is pairwise $\langle\rho(a),\varepsilon(a)\rangle$-connected where $\varepsilon(a)$ can be bounded by approximately $\left(\frac{3}{4}(1-4a^2)\right)^{\frac{d-1}{2}}$. Let $\rho=\rho(a)$ and $\varepsilon=\varepsilon(a)$. By Theorem 3.11, a random subset H of size $s\geq c\frac{\rho}{\varepsilon}\log^2(1/\rho)$ does not have an independent set of size $e^2\rho s$ (with probability $\geq 3/4$). Computing the value of $\log_s\rho s$ we obtain our assertion. We would like to note, that results of a similar nature can be obtained using the above techniques for any value of k. $\square$

5 Random sampling

We now turn to prove Theorem 3.11 stated in Section 3. This is done in two steps. In Section 5.1 we prove results analogous to those presented in Theorem 3.11 when the graphs considered are finite. In Section 5.2 we show that our analysis extends to the continuous case (of Section 3) as well. Finally, in

Section 5.3, we continue the study of finite graphs, and obtain results of independent interest in the context of property testing.

Let G be a graph of size n which does not have an independent set of size ρn (*i.e.* $\alpha(G) < \rho n$). Let H be a random subgraph of G of size s (*i.e.* H is the subgraph induced by a random subset of vertices in G of size s). In this section we study the minimal value of s for which $\alpha(H) \leq \rho s$ with high probability.

In general, if our only assumption on G is that $\alpha(G) < \rho n$, we cannot hope to set s to be smaller than n. Hence, we strengthen our assumption on G, to graphs G which not only satisfy $\alpha(G) \leq \rho n$ but are also *far* from having an independent set of size ρn (we defer defining the exact notion of "far" until later in this discussion). That is, given a graph G which is *far* from having an independent set of size ρn, we ask for the minimal value of s for which (with high probability) a random subgraph of size s does not have an independent set of size ρs. This question (and many other closely related ones) have been studied in [GGR98] under the title of *property testing*.

In [GGR98], a graph G of size n is said to be ε-far from having an independent set of size ρn if any set of size ρn in G has at least εn^2 induced edges. It was shown in [GGR98] that if G is ε-far from having an independent set of size ρn then with high probability a random subgraph of size $s = \frac{c \log(1/\varepsilon)\rho}{\varepsilon^4}$, for a sufficiently large constant c, does not have an independent set of size ρs.

The results of [GGR98] do not suffice for the proof (as we present it) of Theorem 1.2. We thus turn to strengthening their results. To do so, we introduce a stronger notion of being "ε-far". Roughly speaking, we prove that under our new notion of distance, choosing s to be of size $\frac{\rho}{\varepsilon}$ suffices. Furthermore, by applying our proof techniques on the original notion of distance presented in [GGR98], we improve the result of [GGR98] stated above and obtain a sample size proportional to $1/\varepsilon^3$. In Section 6 we continue to study the original notion of ε-far of [GGR98] and present a lower bound on the sample size which is proportional to $1/\varepsilon^2$. The proof techniques used in this section are based on the techniques appearing in [GGR98] and [AK02] (in the latter, property testing of the chromatic number is considered). We start with the following definitions (which are finite versions of those given in Section 3).

Definition 5.1 *Let A and B be (not necessarily disjoint) subsets of G. For each vertex $v \in A$ let $d_v(B)$ be the number of neighbors v has in B. Let $E(A, B) = \sum_{v \in A} d_v(B)$.*

Definition 5.2 *Let $\rho < 1$. A graph $G = (V, E)$ is said to be $\langle \rho, \varepsilon \rangle$-connected iff every subset A of V of size ρn satisfies $E(A, A) \geq \varepsilon n^2$ (*i.e. *the number of edges in the subgraph induced by A is greater than $\frac{\varepsilon}{2} n^2$).*

Notice that $\varepsilon \leq \rho^2$. Furthermore notice, that a graph G is $\langle \rho, \varepsilon \rangle$-connected iff G is $\varepsilon/2$-far (by the definitions presented in [GGR98]) from having an independent set of size ρn.

Definition 5.3 *Let $\rho < 1$. A graph $G = (V, E)$ is said to be* pairwise *$\langle \rho, \varepsilon \rangle$-connected iff every two (not necessarily disjoint) subsets A and B of V of size ρn satisfy $E(A, B) \geq \varepsilon n^2$.*

As mentioned above, for $\langle \rho, \varepsilon \rangle$-connected and pairwise $\langle \rho, \varepsilon \rangle$-connected graphs G, we study the minimal value of s for which a random subgraph H of G of size s satisfies $\alpha(H) \leq \rho s$ with high probability. Namely, we analyze the probability that a random subset H of G satisfies $\alpha(H) \leq \rho s$ (as a function of ρ, ε, and the sample size s). The main idea behind our proof is as follows. Given a sample size s, we start by bounding the probability that a random subset R of G of size $k > \rho s$ is an independent set. Then, using the standard union bound on all subsets R of H of size greater than ρs, we bound the probability that $\alpha(H) > \rho s$.

Throughout this section we analyze the properties of random subsets H which are assumed to be *small*. Namely, we assume that the value of s and the parameters ρ, ε and n satisfy (a) $s < c\sqrt{n}$ and (b) $s < c\rho n$ for a sufficiently small constant c. In our applications (and also in standard ones) these assumptions hold.

In Section 5.1 we analyze the above proof strategy and show that it suffices to bound a condition slightly weaker than the condition $\alpha(H) > \rho s$. Namely, using this scheme, we are able to bound the probability for which $\alpha(H) > \delta\rho s$ for sufficiently large constants δ. This result is used to prove Theorem 3.11 of Section 3. In Section 5.3 we refine our scheme and obtain the main result of this section.

Theorem 5.4 *Let G be a $\langle\rho,\varepsilon\rangle$-connected graph. Let H be a random sample of G of size s. For any constant $c_1 > 0$ there exists a constant $c_2 > 0$ (depending on c_1 alone) s.t.*

1. *If $s \geq c_2\frac{\rho^4}{\varepsilon^3}\log\left(\frac{\rho}{\varepsilon}\right)$ then the probability that H has an independent set of size $> \rho s$ is at most $e^{-c_1\frac{\rho}{\varepsilon}}$.*

2. *If G is pairwise $\langle\rho,\varepsilon\rangle$-connected, and $s \geq c_2\frac{\rho^5}{\varepsilon^3}\log\left(\frac{\rho}{\varepsilon}\right)\log(\frac{1}{\rho})$ then the probability that H has an independent set of size $> \rho s$ is at most $e^{-c_1\frac{\rho^2\log(1/\rho)}{\varepsilon}}$.*

5.1 The Naive scheme

Let $G = (V,E)$ be a $\langle\rho,\varepsilon\rangle$-connected graph (pairwise or not). In this section we study the probability that a random subset R of V of size k is an independent set. We then use this result to bound the probability that a random subset H of G of size s has a large independent set.

We would like to bound (from above) the probability that R induces an independent set. Let $\{r_1,\ldots,r_k\}$ be the vertices of R. Consider choosing the vertices of R one by one, such that at each step the random subset chosen so far is $R_i = \{r_1,\ldots,r_i\}$. Assume that at some stage R_i is an independent set. We would like to show (with high probability) that after adding the remaining vertices of $\{r_{i+1},\ldots,r_k\}$ to R_i, the final set R will not be an independent set.

Let $I(R_i)$ (for *independent*) be the set of vertices in V which are not adjacent to any vertices in R_i, and let $N(R_i)$ be the set of vertices that are adjacent to a vertex in R_i. Consider the next random vertex $r_{i+1} \in R$. If r_{i+1} is chosen from $N(R_i)$ then R_{i+1} is no longer an independent set (implying that R isn't either), and we view this round as a success. Otherwise, r_{i+1} happens to be in $I(R_i)$ and R_{i+1} is still an independent set. But if r_{i+1} also happens to have many neighbors in $I(R_i)$ then adding it to R_i will substantially reduce the size of $I(R_{i+1})$ which works in our favor. This later case is also viewed as a successful round regarding R_i.

Motivated by the discussion above, we continue with the following definitions. As before, let $G = (V,E)$ be a $\langle\rho,\varepsilon\rangle$-connected graph (pairwise or not), let $R = \{r_1,\ldots,r_k\}$ be a set of vertices in V, and let $R_i = \{r_1,\ldots,r_i\}$. Each subset R_i of V defines the following partition (LI_i, HI_i, N_i) of V:

- Let I_i be the vertices that are not adjacent to any vertex in R_i (notice that it may be the case that $R_i \cap I_i \neq \phi$). I_i is now partitioned into two parts: vertices in I_i which have *low* degree, denoted as the set LI_i, and vertices of *high* degree, denoted as HI_i. Namely LI_i is defined to be the ρn vertices of I_i with minimal degree (in the subgraph induced by I_i), and HI_i is defined to be the remaining vertices of I_i. Ties are broken arbitrarily or in *favor* of vertices in R_i (namely, vertices in R_i are placed in I_i before other vertices of identical degree). If it is the case that $|I_i| \leq \rho n$ then LI_i is defined to be I_i, and HI_i is defined to be empty.

- N_i is defined to be the remaining vertices of V (namely, the vertices that share an edge with some vertex in R_i).

We define the partition corresponding to $R_0 = \phi$ as (LI_0, HI_0, NI_0), where LI_0 are ρn vertices of G of minimal degree, HI_0 are the remaining vertices of G, and $N_0 = \phi$.

Notice, using this notation, that the subset R_i is an independent set iff $R_i \cap N_i = \phi$, or equivalently $R_i \subseteq I_i$. Moreover, in this case $R_i \subseteq LI_i$ (all vertices of R_i have degree 0 in the subgraph induced by I_i). Furthermore, each vertex r_i in an independent set $R = R_k = \{r_1, \ldots, r_k\}$ satisfies $r_i \in I_{i-1}$.

We are now ready to bound the probability that a random subset $R = \{r_1, \ldots, r_k\}$ of G is independent. Let $R_i = \{r_1, \ldots, r_i\}$, and let (LI_i, HI_i, N_i) be the corresponding partitions of V defined by R_i. Consider the case in which R is an independent set. As mentioned above, this happens iff for every i the vertex r_i is chosen to be independent from the subset R_{i-1}, or in other words $r_i \in I_{i-1} = LI_{i-1} \cup HI_{i-1}$. We would like to show that this happens with small probability (if k is large enough).

Initially, the subset I_0 is large (the entire vertex set V) and it gets smaller and smaller as we proceed in the choice of vertices in R. Each vertex in $r_i \in HI_{i-1}$ reduces the size of I_{i-1} substantially, while each vertex in LI_{i-1} may only slightly change the size of I_{i-1}. In the following, we show that there cannot be many vertices $r_i \in R$ that happen to fall into HI_{i-1} (as each such vertex reduces the size of I_{i-1} substantially). We thus turn to consider vertices r_i that fall in LI_{i-1} (there are almost k such vertices). The size of LI_i is bounded by ρn. Hence, the probability that $r_i \in LI_i$ is bounded by ρ (by our definitions $R_{i-1} \subseteq LI_{i-1}$ and the vertex r_i is random in $V \setminus R_{i-1}$). This implies that the probability that R is an independent set is roughly bounded by ρ^k. Details follow.

Lemma 5.5 *Let G be a $\langle \rho, \varepsilon \rangle$-connected graph. Let $R = \{r_1, \ldots, r_k\}$ be a set in G. The number of vertices r_i which satisfy $r_i \in HI_{i-1}$ is bounded by $t = \frac{\rho}{\varepsilon}$. If G is pairwise $\langle \rho, \varepsilon \rangle$-connected, then the number of vertices r_i which satisfy $r_i \in HI_{i-1}$ is bounded by $t = \frac{2\rho^2 \lceil \log (1/\rho) \rceil}{\varepsilon}$.*

Proof: We start with the following claim

Claim 5.6 *Let R_i be as defined above, and let (LI_i, HI_i, N_i) be its corresponding partition. Let $I_i = LI_i \cup HI_i$. If G is $\langle \rho, \varepsilon \rangle$-connected then, every vertex in HI_i has degree at least $\frac{\varepsilon}{\rho} n$ (in the subgraph induced by I_i). If G is pairwise $\langle \rho, \varepsilon \rangle$-connected then, every vertex in HI_i has degree at least $\frac{\varepsilon}{2\rho^2} |I_i|$ (in the subgraph induced by I_i).*

Proof: Assume that $|I_i| = \alpha \rho n$ for some $\alpha \geq 1$ (otherwise the set HI_i is empty, and the claim holds). Notice that this implies $|LI_i| = \rho n$. For the first part of our claim, recall by the definition of $\langle \rho, \varepsilon \rangle$-connected graphs that $E(LI_i, LI_i) \geq \varepsilon n^2$. Hence, we conclude that there exists a vertex in LI_i of degree at least $\frac{\varepsilon}{\rho} n$ (in the subgraph induced by LI_i). The set $LI_i \subseteq I_i$, so also in the subgraph induced by I_i, there exists a vertex in LI_i of degree at least $\frac{\varepsilon}{\rho} n$. As LI_i are the vertices of minimal degree in I_i, we conclude the first part of our assertion.

For the second part, let $I_i = X_1 \cup X_2 \cup, \ldots, \cup X_\ell$, where $\{X_1, \ldots, X_\ell\}$ is a partition of I_i into ℓ sets in which the size of X_j for all $j \neq \ell$ is ρn. Notice that $\ell = \lfloor \alpha \rfloor + 1$. For each $v \in I_i$ let $d_v(I_i)$ be the degree of v in the subgraph induced by I_i.

In this case our graph G is pairwise $\langle \rho, \varepsilon \rangle$-connected. This implies that the value of $E(LI_i, X_j)$ for each j (except $j = \ell$) is at least εn^2. Hence, $\sum_{v \in LI_i} d_v(I_i)$ is at least $\lfloor \alpha \rfloor \varepsilon n^2 \geq \frac{\alpha}{2} \varepsilon n^2$. This implies that LI_i must include a vertex v with degree $d_v(I_i) \geq \alpha \frac{\varepsilon}{2\rho} n$. As LI_i are the vertices of minimal degree in I_i, we conclude our assertion. $\Box$

Now to prove our lemma, consider the subsets $R_i = \{r_1, \ldots, r_i\}$ and their corresponding partitions (LI_i, HI_i, N_i). Let $I_i = LI_i \cup HI_i$. Let $N(r_i)$ be the vertices adjacent to r_i in I_{i-1}. We would like to bound the number of vertices r_i that are in HI_{i-1}. We start with the case in which G is $\langle \rho, \varepsilon \rangle$-connected. Consider a vertex r_i in HI_i. By Claim 5.6, its degree in I_{i-1} is $|N(r_i)| \geq \frac{\varepsilon}{\rho} n$. Each vertex $r_i \in HI_{i-1}$ increases the size of N_{i-1} by at least $|N(r_i)|$. Initially, N_0 is empty, and after r_k is chosen $|N_k| \leq n$. We conclude that there are at most $\frac{\rho}{\varepsilon}$ vertices r_i in R which are in HI_{i-1}. This bound can be further improved be a factor of approximately ρ to $\frac{2\rho^2 \lceil \log(1/\rho) \rceil}{\varepsilon}$ using tighter analysis when G is also pairwise $\langle \rho, \varepsilon \rangle$-connected, details follow.

Let $x \geq 0$ be an integer, and let $S_x = \{i \mid r_i \in HI_{i-1} \text{ and } |I_{i-1}| \in [\frac{n}{2^{x+1}}, \frac{n}{2^x})\}$. We would like to bound the size of S_x for all possible values of x. We start by considering values of x between 0 and $\lceil \log(1/\rho) \rceil - 1$. Consider a vertex r_i in which $i \in S_x$. That is, $r_i \in HI_{i-1}$ and $\frac{n}{2^x} \geq |I_{i-1}| > \frac{n}{2^{x+1}}$. By Claim 5.6, the degree of r_i in I_{i-1} is $|N(r_i)| \geq \frac{\varepsilon}{2\rho^2} |I_{i-1}| \geq \frac{\varepsilon}{2\rho^2} \frac{n}{2^{x+1}}$. Each vertex r_i in which $i \in S_x$ increases the size of N_{i-1} by at least $|N(r_i)|$. For such vertices, N_{i-1} is of size at least $n - \frac{n}{2^x}$ and at most of size $n - \frac{n}{2^{x+1}}$. We conclude that $|S_x|$ is of size at most $\frac{2\rho^2}{\varepsilon}$.

For $x \geq \lceil \log(1/\rho) \rceil$, the set S_x is a subset of $\{i \mid r_i \in HI_{i-1} \text{ and } |I_{i-1}| \leq \rho n\}$. Recall that $HI_{i-1} = \phi$ whenever $|I_{i-1}| \leq \rho n$. This implies that $S_x = \phi$ in these cases. All in all, we conclude that $\sum_x |S_x| \leq \frac{2\rho^2 \lceil \log(1/\rho) \rceil}{\varepsilon}$, which concludes our proof. $\Box$

Theorem 5.7 *Let G be a $\langle \rho, \varepsilon \rangle$-connected graph (pairwise or not). Let t be as in Lemma 5.5. Let $k \geq 2t$. The probability that k random vertices of G induce an independent set is at most*

$$\rho^k \left(\frac{ek}{t\rho} \right)^t$$

Proof: Let $R = \{r_1, \ldots, r_k\}$ be a set of k random vertices. As mentioned previously, the probability that $r_i \in LI_{i-1}$ is at most ρ. This follows from the fact that (1) The size of LI_{i-1} is at most ρn, (2) $R_{i-1} \subseteq LI_{i-1}$ (by our definitions), and (3) The vertex r_i is random in $V \setminus R_{i-1}$.

Now in order for R to be an independent set, every vertex r_i of R must be in the set I_{i-1}. Furthermore, by Lemma 5.5 all but t vertices r_i of R must satisfy $r_i \in LI_{i-1}$. Hence, the probability that R is an independent set is at most

$$\binom{k}{t} \rho^{k-t} \leq \left(\frac{ke}{t} \right)^t \rho^{k-t} = \rho^k \left(\frac{ek}{t\rho} \right)^t .$$

$\Box$

Let δ be a large constant. We now use Theorem 5.7 to bound the probability that a random subset H of G of size s has an independent set of size $> \delta \rho s$. The result is the following Corollary 5.8, which will be used in Section 5.2 to prove Theorem 3.11. In Section 5.3 we refine our proof techniques and *get rid* of the parameter δ. That is, we bound the probability that a random subset H of G of size s has an independent set of size $> \rho s$.

Corollary 5.8 *Let G be a $\langle \rho, \varepsilon \rangle$-connected graph (pairwise or not). Let t be as in Lemma 5.5. Let H be a random sample of G of size s. Let $\delta > e$, and let c be a sufficiently large constant. If $s \geq ct\frac{\log(1/\rho)}{\rho}$ then the probability that $\alpha(H) > \delta\rho s$ is at most $\left(\frac{e}{\delta}\right)^{\Omega(\delta\rho s)}$.*

Proof: Let $k = \delta\rho s$. Using Theorem 5.7 and the fact that a subset R of H is random in G, the probability that there is an independent set R in H of size k is at most

$$\binom{s}{k}\rho^k\left(\frac{ek}{t\rho}\right)^t \leq \left[\left(\frac{e}{\delta}\right)^{\frac{k}{t}}\frac{ek}{t\rho}\right]^t \leq \left(\frac{e}{\delta}\right)^{\Omega(k)}.$$

In the last inequality we use the fact that $\frac{k}{t}$ is greater than $c\log(1/\rho)$ for a sufficiently large constant c. $\square$

5.2 Proof of Theorem 3.11

We would now like to show that the analysis presented in Section 5.1 also holds for our continuous graph G_k of Section 3. Namely we would like to prove the following analog of Corollary 5.8.

Theorem 3.11 *Let a, k, $\varepsilon(a)$ be defined as in Theorem 3.6 or 3.8. Let H be a random sample of s vertices of G_k (according to the uniform distribution on S^{d-1}). Let c be a sufficiently large constant. If $s \geq c\frac{\rho(a)}{\varepsilon(a)}\log^2(1/\rho(a))$ then the probability that $\alpha(H) > e^2\rho(a)s$ is at most 1/4.*

Proof: Let $H = \{h_1, \ldots, h_s\}$ be s random points of the unit sphere (that is H is a random subset of G_k of size s). Let $R_k = \{r_k, \ldots, r_k\}$ be a subset of H of size k. Finally, for $i \in \{1, \ldots, k\}$ let $R_i = \{r_1, \ldots, r_i\}$ and (LI_i, HI_i, N_i) be its corresponding partition (as defined in Section 5.1). For each i the subsets LI_i, HI_i, and N_i are measurable (this follows from the fact that the neighborhood of each vertex is a cap of S^{d-1}). Hence, the proof of Lemma 5.5 and Theorem 5.7 hold under Definitions 3.1,3.2 and 3.9 of Section 3. This suffices to conclude our assertion. $\square$

5.3 An enhanced analysis

Let $G = (V, E)$ be a $\langle \rho, \varepsilon \rangle$-connected graph (pairwise or not), and let $H = \{h_1, \ldots, h_s\}$ be a set of random vertices of size s in V. In the previous section we presented a bound on the probability that $\alpha(H) > \delta\rho s$ for large constant values of δ. In this section we enhance our analysis, and bound the probability that $\alpha(H) > \rho s$ (namely, we get rid of the additional parameter δ).

Recall our proof technique from Section 5.1. We started by analyzing the probability that a subset R of H of size k is an independent set. Afterwards we bounded the probability that $\alpha(H) > \delta\rho n$ by using the standard union bound on all subsets R of H of size greater than $k = \delta\rho n$. In this section we enhance the first part of this scheme by analyzing the probability that a subset R of H of size k is a *maximum* independent set in H (rather than just an independent set of H). Then, as before, using the standard union bound on all large subsets R of H, we bound the probability that $\alpha(H) > \rho s$. We show that taking the maximality property of R into account will suffice to prove Theorem 5.4.

Let $H = \{h_1, \ldots, h_s\}$ be s random vertices in G. We would like to analyze the probability that a given subset R of H of size k is a *maximum* independent set. Recall (Section 5.1), that the probability that R is an independent set is bounded by approximately ρ^k. An independent set R is a maximum independent set in H only if adding any other vertex in H to R will yield a set which is no longer independent. Let $R = R_k$ be an independent set, and let (LI_k, HI_k, N_k) be the partition (as defined in Section 5.1) corresponding to R. Consider an additional random vertex h from H. The probability that $R \cup h$ is no longer an independent set is approximately $|N_k|/n$ (here we assume that $|R|$ is small compared to n). The probability that for every $h \in H \setminus R$ the subset $R \cup \{h\}$ is no longer independent is thus $\simeq (|N_k|/n)^{s-k}$. Hence, the probability that a given subset R of H of size k is a *maximum* independent set is bounded by approximately $\rho^k (|N_k|/n)^{s-k}$. This value is substantially smaller than ρ^k iff $|N_k|$ is substantially smaller than n. We conclude that it is in our favor to somehow ensure that $|N_k|$ is not too large. We do this in an artificial manner.

Let $R = \{r_1, \ldots, r_k\}$ be an independent set, let $R_i = \{r_1, \ldots, r_i\}$, and let (LI_i, HI_i, N_i) be the partition (as defined in Section 5.1) corresponding to R_i. Roughly speaking, in Section 5.1, every time a vertex r_i was chosen, the subset N_i was updated. If r_i was chosen in HI_{i-1}, then N_{i-1} grew substantially, and if r_i was chosen in LI_{i-1}, the subset N_{i-1} was only slightly changed. We would like to change the definition of the partition (LI_i, HI_i, N_i) corresponding to R_i as to ensure that N_i is always substantially smaller than n. This cannot be done unless we relax the definition of N_i. In our new definition, N_i will no longer represent the entire set of vertices adjacent to R_i, rather N_i will only include a subset of vertices adjacent to R_i (a subset which is substantially smaller than n). Namely, in our new definition of the partition (LI_i, HI_i, N_i) the set N_{i-1} is changed only if r_i was chosen in HI_{i-1}. In the case in which $r_i \in LI_{i-1} \cup N_{i-1}$, we do not change N_{i-1} at all. As we will see, such a definition will imply that $|N_i| \leq (1-\rho)s$, which will now suffice for our proof.

A new partition Let $H = \{h_1, \ldots, h_s\}$ be a subset of V. Let $R_i = \{r_1, \ldots, r_i\}$ be a subset of H of size i. Each such subset R_i defines a partition (LI_i, HI_i, N_i) of V. As before, let $I_i = LI_i \cup HI_i$.

1. Initially $R_0 = \phi$, LI_0 is the ρn vertices in V of minimal degree (in V), $HI_0 = V \setminus LI_0$, and $N_0 = \phi$. In the above, ties are broken by an assumed ordering on the vertices in V.

2. Let (LI_i, HI_i, N_i) be the partition corresponding to R_i, let r_{i+1} be a new random vertex. Let $R_{i+1} = R_i \cup \{r_{i+1}\}$, we define the partition $(LI_{i+1}, HI_{i+1}, N_{i+1})$. Let $N(r_{i+1})$ be the neighbors of r_{i+1} in I_i. We consider the following cases:

   - If $r_{i+1} \in LI_i$ then the partition corresponding to R_{i+1} will be exactly the partition corresponding to R_i namely $LI_{i+1} = LI_i$, $HI_{i+1} = HI_i$, and $N_{i+1} = N_i$. Notice that this implies that N_{i+1} no longer represents all neighbors of R_{i+1}. There may be vertices adjacent to R_{i+1} which are in I_{i+1}.
   - If $r_{i+1} \in HI_i$ then we consider two sub-cases:
     - If $|N_i \cup N(r_{i+1})| \leq (1-\rho)n$, then LI_{i+1}, HI_{i+1}, and N_{i+1} are defined as in Section5.1. Namely, $N_{i+1} = N_i \cup N(r_{i+1})$. I_{i+1} is defined to be $V \setminus N_{i+1}$. LI_{i+1} is defined to be the ρn vertices of I_{i+1} with minimal degree (in the subgraph induced by I_{i+1}) and HI_{i+1} is defined to be the remaining vertices of I_{i+1}. Ties are broken by the assumed ordering on V.
     - If $|N_i \cup N(r_{i+1})| > (1-\rho)n$, then let $\hat{N}(r_{i+1})$ be the first (according to the assumed ordering on V) $(1-\rho)n - |N_i|$ vertices in $N(r_{i+1})$ and set $N_{i+1} = N_i \cup \hat{N}(r_{i+1})$. Furthermore, set LI_{i+1} to be the remaining ρn vertices of G, and HI_i to be empty. Notice that in this case $|N_{i+1}|$ is of size exactly $(1-\rho)n$.

- If $r_{i+1} \in N_i$ then, once again, the partition corresponding to R_{i+1} will be exactly the partition corresponding to R_i.

A few remarks are in place. First of all it is not hard to verify that the definition above implies

Claim 5.9 *Let $i \in \{1, \ldots, k\}$. The partitions (LI_i, HI_i, N_i) corresponding to R_i as defined above satisfy (a) $I_i \subseteq I_{i-1}$. (b) $N_{i-1} \subseteq N_i$. (c) $|N_i| \leq (1-\rho)n$. (d) $|LI_i| = \rho n$. (e) The set LI_i is the ρn vertices of minimal degree in I_i.*

Secondly, due to the iterative definition of our new partition, the partitions (LI_i, HI_i, N_i) corresponding to the subsets R_i depend strongly on the specific ordering of the vertices in R_i. Namely, in contrast to the partitions used in Section 5.1, a single subset R with two different orderings may yield two different partitions. For this reason, in the remainder of this section, we will assume that the vertices of H are chosen one by one. This will imply an ordering on H and on any subset R of H. The partitions we will study will correspond to these orderings only.

Finally, in Section 5.1, an (ordered) subset $R = \{r_1, \ldots, r_k\}$ was independent iff $\forall i \;\; r_i \in I_{i-1}$ (according to the definition of I_{i-1} appearing in Section 5.1). In this section, if R is independent then it still holds that $\forall i \;\; r_i \in I_{i-1}$. However, it may be the case that $\forall i \;\; r_i \in I_{i-1}$ but R is not an independent set. In the remainder of this section, we call ordered subsets R for which $\forall i \;\; r_i \in I_{i-1}$ *free* sets. We analyze the probability that a random ordered subset H of V of size s does not have any free sets of size larger then ρs. This implies, that H does not include any independent sets of size ρs.

Definition 5.10 *An ordered subset $R_i = \{r_1, \ldots, r_i\}$ is said to be $free$ if it is the case that $r_j \in I_{j-1}$ for all $j \leq i$.*

Claim 5.11 *Let $H = \{h_1, \ldots, h_s\}$ be an ordered set of vertices in a (pairwise) $\langle \rho, \varepsilon \rangle$-connected graph G. If $\alpha(H) > \rho s$ then the maximum free set in H (w.r.t. the ordering implied by H) is of size $> \rho s$.*

Proof: Let I be an independent set of size $> \rho s$ in H. It is not hard to verify that I (under the ordering implied by H) is a free set. Let I' be a maximal free set including I, where I' is ordered by the ordering implied by H. I' is of size $> \rho s$. We conclude that the maximum free set R in H (ordered by the ordering implied by H) is of size $\geq |I'| > \rho s$. $\square$

Claim 5.11 implies that to prove Theorem 5.4 it suffices to analyze the maximum free set $R \subseteq H$. Moreover, the only ordered subsets R that we need to consider are those ordered by the ordering implied by H. We now turn to prove Theorem 5.4. Roughly speaking, we start by analyzing the probability that a random subset R is a free set. We then analyze the probability that a given subset R in H is a maximum free set. Finally, we use the union bound on all subsets R of H of size $> \rho s$ to obtain our results.

In the remainder of this section, we will assume that the subset H is chosen from G randomly *with repetitions*. That is H is a random multi-set of size s. Our results (with minor modifications) apply also to the case in which H is a random subset of G (and not a multi-set) if the size of H is not very large (here we assume that $|H| << \sqrt{n}$). As in such cases, a set H of size s which is randomly chosen from V with repetitions will not include the same vertex twice (with high probability).

We start by stating the following lemmas which are analogous to Lemma 5.5 and Theorem 5.7 from Section 5.1. The main difference between the lemmas below (and their proof), and those of the previous section is in the definition of the partition (LI_i, HI_i, N_i) and in the fact that they address free sets instead of independent sets. Proof of the lemmas is omitted.

Lemma 5.12 *Let G be a $\langle\rho,\varepsilon\rangle$-connected graph. Let $R = \{r_1,\ldots,r_k\}$ be an ordered set in G of size k. The number of vertices r_i which satisfy $r_i \in HI_{i-1}$ is bounded by $t = \frac{\rho}{\varepsilon}$. If G is pairwise $\langle\rho,\varepsilon\rangle$-connected then the number of vertices r_i which satisfy $r_i \in HI_{i-1}$ is bounded by $t = \frac{2\rho^2\lceil\log(1/\rho)\rceil}{\varepsilon}$.*

Lemma 5.13 *Let G be a $\langle\rho,\varepsilon\rangle$-connected graph (pairwise or not). Let t be as in Lemma 5.12. Let $k \geq 2t$. Let $R = \{r_1,\ldots,r_k\}$ be k random vertices of G. The probability that R induces a free set is at most*

$$\rho^k \left(\frac{ek}{t\rho}\right)^t$$

We now address the probability that a random subset R of H is a maximum free set. We will then use the union bound on all subsets R of H of size $> \rho s$ to obtain our results.

Lemma 5.14 *Let G be a $\langle\rho,\varepsilon\rangle$-connected graph (pairwise or not). Let t be as in Lemma 5.12. Let $k \geq 2t$. Let H be an ordered random sample of G of size $s \geq k$. The probability that a given subset R of H is a maximum free set is at most*

$$\rho^k \left(\frac{ek}{t\rho}\right)^t (1-\rho)^{s-k}$$

Proof: Let $R = \{r_1,\ldots,r_k\}$ (ordered by the ordering induced by H). The set R is a maximum free set in H only if (a) R is free, and (b) For each vertex $h \in H$ which is not in R, the ordered set $R^+ = \{r_1,\ldots,r_j,h,r_{j+1},\ldots,r_k\}$ is not free. Here the index j is such that r_j appears before h in the ordering of H, and r_{j+1} appears after h (*i.e.* R^+ is ordered according to the ordering of H).

The probability that R is free has been analyzed in Lemma 5.13. It is left to analyze the probability that R^+ is not free for every vertex $h \notin R$, given that R is free. Consider a vertex $h \in H$ which is not in R, and let $R^+ = \{r_1,\ldots,r_j,h,r_{j+1},\ldots,r_k\}$.

Claim 5.15 *Let $R = \{r_1,\ldots,r_k\}$ be a free set. Let $R^+ = \{r_1,\ldots,r_j,h,r_{j+1},\ldots,r_k\}$. Let the partition corresponding to $R_j = \{r_1,\ldots,r_j\}$ be (LI_j,HI_j,N_j). If $h \in LI_j$ then R^+ is also a free set.*

Proof: We will use the following notation. Let $R_i = \{r_1,\ldots,r_i\}$ denote the first i vertices of R, and let (LI_i,HI_i,N_i) be its corresponding partition. For $i > j$, let $R_i^+ = \{r_1,\ldots,r_j,h,r_{j+1},\ldots,r_i\}$ denote the first $i+1$ vertices of R^+, and let (LI_i^+,HI_i^+,N_i^+) be its corresponding partition. Finally let R_h^+ denote the subset $\{r_1,\ldots,r_j,h\}$ and (LI_h^+,HI_h^+,N_h^+) be its corresponding partition.

We would like to prove that R^+ is free. That is, we would like to show (a) that $r_i \in I_{i-1}$ for each $i \leq j$, (b) that $h \in I_j$, (c) that $r_{j+1} \in I_h^+$, and (d) that $r_i \in I_{i-1}^+$ for $i \geq j+2$. Recall that R is free and thus $r_i \in I_{i-1}$ for all $i \in \{1,\ldots,k\}$.

The first assertion follows from the fact that the first j vertices of R and R^+ are identical. The second follows from the assumption that $h \in LI_j$. For the third assumption, notice (as $h \in LI_j$) that the partition corresponding to $R_h^+ = \{r_1,\ldots,r_j,h\}$ is equal to the partition corresponding to $R_j = \{r_1,\ldots,r_j\}$. This follows from our definition of the partition (LI_h^+,HI_h^+,N_h^+). As $r_{j+1} \in I_j$ we conclude that $r_{j+1} \in I_h^+$.

For the final assertion, observe that for any $i \geq j+1$, the partition corresponding to R_i^+ is equal to the partition corresponding to R_i. This can be seen by induction (on i). We start with the partitions corresponding to R_{j+1} and R_{j+1}^+. The partition $(LI_{j+1},HI_{j+1},N_{j+1})$ is defined uniquely by the partition corresponding to R_j and the vertex r_{j+1}. Similarly, the partition $(LI^+_{j+1}, HI^+_{j+1}, N^+_{j+1})$ is defined uniquely by the partition corresponding to R^+_h and the vertex r_{j+1}. As the partition corresponding to R^+_h is equal to the partition corresponding to R_j, we conclude that the same hold for the partitions corresponding to R_{j+1} and R^+_{j+1}. The inductive step is done similarly. The partition corresponding to R_i (R^+_i) is defined uniquely by the partition corresponding to R_{i-1} (R^+_{i-1}) and the vertex r_i. As the partition corresponding to R_{i-1} equals that corresponding to R^+_{i-1} we conclude our claim. As R is free, $r_i \in I_{i-1}$ for every $i \geq j+2$. This implies, that $r_i \in I^+_{i-1}$ also, which proves the final assertion. $\square$

Claim 5.15 implies that the probability that $R^+ = \{r_1, \ldots, r_j, h, r_{j+1}, \ldots, r_k\}$ is not free given that R is free is at most $(1-\rho)$ (recall that the set LI_j is of size exactly ρn). This holds independently for every vertex h in $H \setminus R$. We conclude that the probability that R is a maximum free subset of H is at most the probability that R is free times $(1-\rho)^{s-k}$. $\square$

We now turn to analyze the probability that a random ordered subset H of G of size s has a free set of size larger than ρs. We follow the line of analysis given in Section 5.1 and analyze the probability that H has a free set of size larger than $\delta\rho s$ for any $\delta > 1$. We then get rid of the factor δ to obtain our main theorem of this section.

Corollary 5.16 *Let G be a $\langle \rho, \varepsilon \rangle$-connected graph (pairwise or not). Let t be as in Lemma 5.12. Let H be a random sample of G of size s. Let $\delta > 1$, and let c be a sufficiently large constant. Let $\Gamma = \ln\delta - \frac{\delta-1}{\delta}$. If $s \geq \frac{ct}{\rho\Gamma}(\log(1/\rho) + \log(1+1/\Gamma))$ then the probability that H has a free set of size $> \delta\rho s$ is at most*

$$s\left(\frac{1}{e^{\Gamma}}\right)^{\Omega(\delta\rho s)}$$

Proof: Let $k = \delta\rho s$, let $\delta' > \delta$ and let $k' = \delta'\rho s > k$. Using Lemma 5.14, the probability that there is a maximum free set R in H of size k' is at most

$$\begin{aligned}
\sum_{k'>k} \binom{s}{k'} \rho^{k'} \left(\frac{ek'}{t\rho}\right)^t (1-\rho)^{s-k'} &\leq \sum_{k'>k} \left(\frac{ek'}{t\rho}\right)^t \frac{s^s(1-\rho)^{s-k'}}{k'^{k'}(s-k')^{s-k'}} \rho^{k'} \\
&\leq \sum_{k'>k} \left(\frac{ek'}{t\rho}\right)^t \frac{e^{k'\frac{\delta'-1}{\delta'}}}{\delta'^{k'}} \\
&= \sum_{k'>k} \left[\frac{ek'}{t\rho}\left(\frac{1}{e^{\ln\delta' - \frac{\delta'-1}{\delta'}}}\right)^{\frac{k'}{t}}\right]^t \\
&\leq \sum_{k'>k} \left[\frac{ek'}{t\rho}\left(\frac{1}{e^{\Gamma}}\right)^{\frac{k'}{t}}\right]^t \\
&\leq \sum_{k'>k} \left(\frac{1}{e^{\Gamma}}\right)^{\Omega(k')} \leq s\left(\frac{1}{e^{\Gamma}}\right)^{\Omega(k)}
\end{aligned}$$

We use the facts that $\frac{k'}{t}$ is greater than both $c\frac{1}{\Gamma}\log(1+1/\Gamma)$ and $c\frac{1}{\Gamma}\log(1/\rho)$ for a sufficiently large constant c; and the fact that Γ is an increasing function of δ (for $\delta > 1$). $\square$

It is left to get rid of the additional parameter δ of Corollary 5.16 (namely to analyze the probability that $\alpha(H) > \rho s$).

Lemma 5.17 *If a given graph G is (pairwise) $\langle \rho, \varepsilon \rangle$-connected then G is also (pairwise) $\left\langle \rho(1 - \frac{\varepsilon}{4\rho^2}), \frac{\varepsilon}{2} \right\rangle$-connected.*

Proof: We present proof for the case in which G is $\langle \rho, \varepsilon \rangle$-connected, a similar proof holds for the case in which G is pairwise $\langle \rho, \varepsilon \rangle$-connected. Let A be some subset of G of size $\rho(1 - \frac{\varepsilon}{4\rho^2})n$. Let A^c be any set in $V \setminus A$ of size $\frac{\varepsilon}{4\rho}n$. It is known that the number of edges induced by the set $A \cup A^c$ is at least $\frac{\varepsilon}{2}n^2$ (notice that $|A \cup A^c| = \rho n$ and $E(A \cup A^c, A \cup A^c) \geq \varepsilon n^2$). The number of edges (in $A \cup A^c$) adjacent to vertices in A^c is bounded by $\frac{\varepsilon}{4\rho}\rho n^2 = \frac{\varepsilon}{4}n^2$. Hence, the number of edges induced by vertices in A is at least $\frac{\varepsilon n^2}{4}$ implying that $E(A, A) \geq \frac{\varepsilon}{2}n^2$. $\square$

Theorem 5.4 *Let G be a $\langle \rho, \varepsilon \rangle$-connected graph (pairwise or not). Let t be as in Lemma 5.12. Let H be a random sample of G of size s. Let c be a sufficiently large constant. If $s \geq ct\frac{\rho^3}{\varepsilon^2}\log\left(\frac{\rho}{\varepsilon}\right)$ then the probability that H has an independent set of size $> \rho s$ is at most $e^{-\Omega(t)}$.*

Proof: By Lemma 5.17, G is also $\left\langle \rho(1 - \frac{\varepsilon}{4\rho^2}), \frac{\varepsilon}{2} \right\rangle$-connected (pairwise or not). Let $\rho' = \rho(1 - \frac{\varepsilon}{4\rho^2})$ and $\varepsilon' = \varepsilon/2$. We would like to bound the probability that H does not have any independent sets of size greater than ρs. Let $\delta = 1 + \frac{\varepsilon}{4\rho^2}$. Notice that $\delta\rho' \leq \rho$. Hence, it suffices to bound the probability that $\alpha(H) > \delta\rho' s$. This probability, in turn, is at most the probability that H has a maximum free set of size greater than $k = \delta\rho' s$ (Claim 5.11).

Let $\Gamma = \ln(\delta) - \frac{\delta - 1}{\delta}$. It is not hard to verify that $\Gamma = \theta((\delta - 1)^2) = \theta\left(\frac{\varepsilon^2}{\rho^4}\right)$ for our value of δ. By our assumption, s is greater or equal to

$$ct\frac{\rho^3}{\varepsilon^2}\log\left(\frac{\rho}{\varepsilon}\right) \geq \frac{c_1 t}{\rho(\delta - 1)^2}\left(\log(1/\rho') + \log(1 + 1/(\delta - 1)^2)\right) \geq \frac{c_2 t}{\rho'\Gamma}\left(\log(1/\rho') + \log(1 + 1/\Gamma)\right)$$

where in the above, c_1 and c_2 are constants closely related to c. Now, by Corollary 5.16, for our choice of s, the probability that H has a maximum free set of size greater than $k = \delta\rho' s$ is at most $s\left(\frac{1}{e^\Gamma}\right)^{\Omega(\delta\rho' s)} \leq e^{-\Omega(t)}$. $\square$

Roughly speaking, Theorem 5.4 states that given a $\langle \rho, \varepsilon \rangle$-connected graph G, a random sample H of G of size s proportional to $\frac{\rho^4}{\varepsilon^3}$ (or larger) will not have an independent set of size ρs (with high probability). This improves upon the bound of $s \simeq \frac{\rho}{\varepsilon^4}$ presented in [GGR98] both in the dependence on ρ (as $\rho < 1$) and in the dependence on ε. Moreover, we present a further improvement to $s \simeq \frac{\rho^5}{\varepsilon^3}$ if our graphs are considered to be pairwise $\langle \rho, \varepsilon \rangle$-connected. In Section 6 we continue to study the minimal value of s for which $\alpha(H) < \rho s$ with high probability, and present a lower bound on the size of s which is proportional to $\frac{\rho^3}{\varepsilon^2}$.

6 Lower bounds for the testing of $\alpha(G)$

In this section we present graphs G which are $\langle \rho, \varepsilon \rangle$-connected, but with some constant probability, a random sample R of G of size $s \sim \rho^3/\varepsilon^2$ is likely to have an independent set of size greater than ρs.

Lemma 6.1 *Let ρ be a small constant and $\varepsilon < \rho^2$ s.t. $\rho^3/\varepsilon^2 << n$. For n large enough, there exists a graph G on n vertices for which (a) G is $\langle\rho,\varepsilon\rangle$ connected, (b) With constant probability (independent of ρ and ε) a random set R of size $s = \frac{\rho^3}{\varepsilon^2}$ will have an independent set of size ρs.*

Proof: Consider the graph $G = (V, E)$ in which $|V| = n$, V consists of two disjoint sets A and $V \setminus A$, where A is an independent set of size $(1 - \frac{\varepsilon}{\rho^2})\rho n$, $V \setminus A$ induces a clique, and every vertex in A is adjacent to every vertex in $V \setminus A$. On one hand, every subset of size ρn in G induces a subgraph with at least $\varepsilon n^2/2$ edges (implying that G is $\langle\rho,\varepsilon\rangle$-connected). On the other, let R be a random subset of V obtained by picking each vertex independently with probability $\frac{\rho^3}{\varepsilon^2 n}$. The expected size of R is $s = \frac{\rho^3}{\varepsilon^2}$. In the following we assume R is exactly of size s, minor modifications in the proof are needed if this assumption is not made. The set $R \cap A$ is an independent set in the subgraph induced by R. The expected size of $R \cap A$ is $(1 - \frac{\varepsilon}{\rho^2})\rho s$. Let $N(0, 1)$ denote a standard normal variable. It can be seen using the central limit theorem (for example [Fel66]) that for our choice of parameters, the probability that $|R \cap A|$ deviates from its expectation by more than a square root of its expectation is at least

$$\Pr\left[|R \cap A| > \left(1 - \frac{\varepsilon}{\rho^2}\right)\rho s + \sqrt{\rho s}\right] > \Pr\left[N(0, 1) > 1\right]$$

which is some constant probability independent of ε and ρ. In such a case the size of $R \cap A$ will be greater than $(1 - \frac{\varepsilon}{\rho^2})\rho s + \sqrt{\rho s} = \rho s$ for our value of s. Hence implying assertion (b) of the lemma. $\square$

7 An alternative proof of Theorem 1.2 (1)

In Theorem 1.2 (1), we are interested in presenting a vector k-colorable graph H in which a special relationship is satisfied between its maximum degree and its maximum independent set. It is not hard to verify that the graphs G_k presented in Section 3 are far from satisfying this relationship, as the maximum degree Δ of these graphs is too large. We overcome this problem in Sections 4 and 5 by considering a random (vertex induced) subgraph of G_k. We have shown that such a subgraph will suffice for proving the three parts of Theorem 1.2.

Another method for coping with the large maximum degree Δ of G_k was suggested by Luca Trevisan (private communication). Instead of sampling vertices from G_k at random in order to obtain a sparse graph, consider sampling **edges** at random. In the following section we combine the idea of edge sampling with our results from Section 3 and prove the first part of Theorem 1.2. The use of edge sampling simplifies the proof of Theorem 1.2 (1) (as the analysis presented in Section 5 is no longer needed). It appears that the remainder of Theorem 1.2 can not be proven using edge sampling (as we are interested in graphs in which the maximum independent set is small with respect to the number of vertices in the graph).

To construct the graph H we will follow a three phase plan. Our starting point will be the continuous graph G_k (of Section 3) which is vector k-colorable and proven to be pairwise $\langle\rho,\varepsilon\rangle$-connected. We then define and analyze a (finite) discrete version G_k^d of G_k which will be shown to inherit many of the properties of G_k. Namely this discrete graph will be *almost* vector k-colorable, and will be pairwise $\langle\rho,\varepsilon\rangle$-connected. Finally, we will define H to be the graph obtained by randomly removing edges from the discrete graph G_k^d.

The discrete graph G_k^d We now define a discrete analog G_k^d of the continuous graph G_k from Section 3. Recall that the vertex set of G_k is the d dimensional unit sphere S^{d-1}. It is shown in [FS02] that S^{d-1} can be partitioned into $n = 2^{\theta(d^2)}$ cells of equal size and of diameter at most 2^{-d} each. Let $\mathcal{P} = \{C_1, \ldots, C_n\}$ denote the cells obtained in the above partition. The graph G_k^d will be of size n, in which each vertex $v_i \in V$ corresponds to a cell $C_i \in \mathcal{P}$. The edge set of G_k^d consists of an edge (u, v) iff there is a positive measure of edges in G_k between their corresponding cells C_u, C_v.

Lemma 7.1 *Let A and B be subsets of G_k^d, and let A_c and B_c be the corresponding subsets of G_k. (a) The size of A (B) is ρn iff A_c (B_c) has measure ρ. (b) $E(A, B) \geq E(A_c, B_c)n^2$. The definition of $E(A, B)$ is given in Definition 5.1 of Section 5.*

Proof: For the first part of the lemma, assume that the set A has ρn vertices, thus the corresponding subset A_c consists of ρn cells each of measure n^{-1}. We conclude that the subset A_c has measure exactly ρ. On the other hand, if A_c has measure ρ and consists of the union of k cells of measure n^{-1}, then k must be ρn. For the second part, assume that $E(A, B) = \varepsilon n^2$, then by the fact that each edge between A and B corresponds to at most the measure of n^{-2} edges between A_c and B_c we conclude that $E(A_c, B_c)$ is at most ε. $\square$

Theorem 7.2 *Let a, k, $\varepsilon(a)$ be as defined in Theorem 3.6 or 3.8. The graph G_k^d is pairwise $\langle \rho(a), \varepsilon(a) \rangle$-connected.*

Proof: Let A and B be subsets (in G_k^d) of size $\rho(a)n$. The corresponding subsets A_c and B_c of G_k are also of measure $\rho(a)$ (Lemma 7.1). By Corollary 3.10, $E(A_c, B_c) \geq \varepsilon(a)$. We conclude that $E(A, B) \geq E(A_c, B_c)n^2 \geq \varepsilon(a)n^2$. $\square$

Lemma 7.3 *The graph G_k^d is vector $k\left(1 + \frac{ck^2}{2^d}\right)$-colorable for some constant $c > 0$.*

Proof: Recall that each cell in G_k^d has diameter at most 2^{-d}. Hence, two vertices in G_k^d are connected only if their inner product is less than $-1/(k-1) + \theta(1)2^{-d} \leq -\frac{1}{k\left(1+\frac{\theta(k)}{2^d}\right)-1}$. $\square$

By definition, the continuous graph G_k is vector k-colorable. In Lemma 7.3 we show that the finite approximation G_k^d to G_k is *almost* vector k-colorable. In general, this does not suffice for the proof of Theorem 1.2 as we are interested in graphs which are vector k-colorable (rather than "almost vector k-colorable"). This can be fixed by starting with a continuous graph with vector coloring number slightly less than k (e.g. $k/\left(1 + \frac{ck^2}{2^d}\right)$). In order to simplify our presentation, we neglect this point and consider the graph G_k^d to be exactly vector k-colorable. This is possible due to the fact that the properties of G_k are *continuous* in k. Namely, choosing d large enough, it can be seen that the multiplicative error of $(1 + \frac{ck^2}{2^d})$ in the value of k does not effect the analysis appearing throughout this section.

Lemma 7.4 *Let $\rho(\frac{1}{k-1})$ be the measure of a $\frac{1}{(k-1)}$-cap. Every vertex v in the graph G_k^d has degree $d_v \in [\frac{1}{poly(d)}\rho(\frac{1}{k-1})n, poly(d)\rho(\frac{1}{k-1})n]$.*

Proof: Consider a vertex v in G_k^d and its corresponding cell C_v. The degree of v is the number of cells in G_k which share a positive measure of edges with the cell C_v. The total measure of these cells is at least the measure of a $\left(\frac{1}{k-1} + \theta(2^{-d})\right)$-cap and at most the measure of a $\left(\frac{1}{k-1} - \theta(2^{-d})\right)$-cap. Hence, by Lemma 7.1 we conclude our theorem. $\square$

We now prove the first part of Theorem 1.2 by considering the graph H obtained by randomly sampling the edges of G_k^d.

Theorem 7.5 *For every constant $\gamma > 0$ and constant $k > 2$, there are infinitely many graphs H that are vector k-colorable and satisfy $\alpha(H) \leq \frac{n}{\Delta_H^{1-\frac{2}{k}-\gamma}}$, where n is the number of vertices in H and Δ_H is the maximum degree in H.*

Proof: Let $k > 2$ be constant. Let $\gamma > 0$ be an arbitrarily small constant. Let $a = \gamma/c$ for a sufficiently large constant c. Let $G = G_k^d = (V, E)$ be the discrete graph defined above. Let n be the size of the vertex set V of G, and let Δ be the maximum degree of G. Recall that $n = 2^{\theta(d^2)}$ where d is the dimension in which the corresponding graph G_k was defined. We will assume the dimension d is a very large constant determined after fixing a. Finally, let $\rho = \rho(a)$.

By Lemma 7.4 all vertices in G are of degree in the range $[\frac{1}{poly(d)}\Delta, \Delta]$ where $\Delta \sim poly(d)\rho(\frac{1}{k-1})n$. By Theorem 7.2 and the proof of Claim 4.1, every subset of vertices U in G of size ρn has at least $\Delta\rho^{\frac{2(k-1)}{k-2}+\gamma}n$ edges.

Let $p = \frac{1}{\Delta\rho^{\frac{k}{k-2}+2\gamma}}$. Let H be the subgraph of G obtained by deleting each edge of G independently with probability $(1-p)$.

Lemma 7.6 *With probability $\geq 3/4$ all vertices v of H will have degree $d_v(H)$ in the range*

$$\left[\frac{1}{poly(d)}\rho^{-\frac{k}{k-2}-2\gamma}, 2\rho^{-\frac{k}{k-2}-2\gamma}\right].$$

Proof: The expected degree $d_v(H)$ of each vertex v in H satisfies

$$d_v(H) \in \left[\frac{1}{poly(d)}\rho^{-\frac{k}{k-2}-2\gamma}, \rho^{-\frac{k}{k-2}-2\gamma}\right].$$

It is not hard to verify (using standard bounds) that with probability $\geq 3/4$ it is the case that all vertices v have degree $d_v(H)$ which does not deviate from their expectation by more than a constant fraction of their expectation. $\square$

Lemma 7.7 *With probability $\geq 3/4$ the size of the maximum independent set in H ($\alpha(H)$) is at most ρn.*

Proof: It suffices to show that every subset U of H of size ρn has at least a single edge. Using Claim 9.2, it follows that for each subset U of H, the probability that all its edges were removed is at most

$$(1-p)^{\Delta\rho^{\frac{2(k-1)}{k-2}+\gamma}n} \leq e^{-\rho^{1-\gamma}n}.$$

The number of subsets U of size ρn is

$$\binom{n}{\rho n} \leq e^{\rho n \ln n} \leq \frac{1}{4}e^{\rho^{1-\gamma}n}$$

Applying the union bound on all subsets U of H of size ρn we conclude our assertion $\square$

Now with probability at least 1/2 both Lemma 7.6 and Lemma 7.7 hold implying our theorem. $\square$

8 Discussion

In our work we have presented *tight* bounds on the chromatic number of vector k-colorable graphs. Tight in the sense that they match the upper bounds presented in [KMS98]. Many questions still remain open.

Stronger coloring relaxations As mentioned in the Introduction and Section 2, there are stronger relaxations to the minimum coloring problem that also have a geometrical interpretation. For example, one such relaxation is the well known (and extensively studied) Lovász theta function [Lov79]. It is not hard to verify that these relaxations can be used *as is* in the coloring algorithm presented in [KMS98]. One may speculate that using such stronger relaxations will yield improved coloring results. At the moment this is not known to be true.

One may consider proving that even the use of such stronger relaxations in the [KMS98] algorithm cannot yield stronger coloring results. Or in other words, one may try to extend our negative results to stronger relaxations as well. A few remarks are in place. It seems as if the techniques we use in this work do not extend to the stronger coloring relaxations presented in Section 2. Our graphs (continuous and random), or to be precise, their embeddings, are not valid with respect to these stronger coloring relaxations. Therefore, in order to extend our negative results, one must change these embeddings appropriately without increasing their vector chromatic number. It seems as if changing the embedding of our graphs as to satisfy these stronger coloring relaxations has a large effect on their vector chromatic number. This implies that such an approach will yield weaker negative results.

Alternatively, one may consider using our proof techniques on graphs other than the ones presented. One natural candidate is the graph $G = (V, E)$ in which the vertex set V consists of the set $\{0, 1\}^n$ and two vertices are connected by an edge if their Hamming distance is equal to some prespecified value. The graph G and certain subgraphs of G have been used in the past in the context under discussion (*e.g* [KMS98, Fei97, GK98, Cha02]). Using our proof technique on such graphs involves the analysis of certain edge isoperimetric inequalities (analogous to those presented in Theorem 3.5). Unfortunately, little is known regarding the edge isoperimetric inequalities of the above graphs G. Such inequalities have been studied in the past ([Bez02, KKL88]) yielding partial results. However, these results do not suffice to extend our proof techniques. A better understanding of edge isoperimetric inequalities of these graphs is of great interest regardless their application to the vector coloring issue.

Property testing In Section 5 and Section 6 we study the property testing paradigm with respect to the independent set problem. We present improved results on the sample size needed when testing graphs which are *far* from having large independent sets. Namely, in Theorem 5.4 we prove that if a graph G of size n is $\langle \rho, \varepsilon \rangle$-connected then w.h.p. a random induced subgraph of G of size $s \sim 1/\varepsilon^3$ will not have an independent set of size ρs (this improves upon the sample size of $s \sim 1/\varepsilon^4$ presented in [GGR98]). Moreover, in Lemma 6.1 we show that the sample size s must be of size at least $\sim 1/\varepsilon^2$ if we wish the probability of failure to be nonconstant. It would be interesting if the factor ε gap between the upper and lower bounds presented above could be settled.

9 Proof of claims

Claim 9.1 *For $1 \geq p \geq 0$ and $y \geq 1$, we have that $(1-p)^y \geq 1 - py \geq (1-p)^{y + \frac{py^2}{1-py}}$.*

Proof: It is well known that $e^y \geq 1 + y$ for any real y. Hence,

$$1 - py = \frac{1}{1 + py + \frac{p^2y^2}{1-py}} \geq e^{-p(y+\frac{py^2}{1-py})} \geq (1-p)^{y+\frac{py^2}{1-py}}.$$

As for the other direction, for every $y \geq 1$

$$\frac{d}{dp}((1-p)^y - 1 + py) = y(1 - (1-p)^{y-1}) \geq 0.$$

□

Claim 9.2 *For all $x > 1$*

$$\left(1 - \frac{1}{x}\right)\frac{1}{e} \leq \left(1 - \frac{1}{x}\right)^x \leq \frac{1}{e}$$

Proof: It is known that for all $x > 1$

$$\left(1 - \frac{1}{x}\right)^x \leq \frac{1}{e} \leq \left(1 - \frac{1}{x}\right)^{x-1}$$

□

Claim 9.3 *Let $a > \delta > 0$ such that $a\delta \geq 2\log(d)/d$ then*

$$\rho(a) - \rho(a + \delta) \geq (1 - 1/d)\rho(a).$$

Proof: Recall that $\frac{c}{\sqrt{d}}\left(1 - a^2\right)^{\frac{d-1}{2}} \leq \rho(a) \leq \left(1 - a^2\right)^{\frac{d-1}{2}}$, for some constant $c > 0$. Thus,

$$\rho(a+\delta) \leq \left(1 - a^2 - \delta^2 - 2a\delta\right)^{\frac{d-1}{2}} \leq \left((1-a^2)(1-2a\delta)\right)^{\frac{d-1}{2}} \leq \left(1 - a^2\right)^{\frac{d-1}{2}}\frac{2}{d^2} \leq \rho(a)/d.$$

□

Theorem 3.5 *Let $1 \geq a > 0$ and let A and B be two (not necessarily disjoint) measurable sets in V of measure $\rho(a)$. Let r be an arbitrary vertex of S^{d-1}. The minimum of $E(A, B)$ is obtained when $A = B = \mathcal{C}_a$ where $\mathcal{C}_a$ is an a-cap of measure $\rho(a)$ centered at r.*

Proof: Our proof is similar to that presented in [FS02]. We start by reviewing the *two point symmetrization* procedure introduced in [BT76]. Let $r \in S^{d-1}$ and H be the hyperplane passing through the origin with normal r. The hyperplane H partitions S^{d-1} into three sets: points *above*, *below*, and *on* the hyperplane. Define S_0 to be the set $\{u \in S^{d-1} \mid \langle u, r\rangle = 0\}$, S^+ to be $\{u \in S^{d-1} \mid \langle u, r\rangle > 0\}$, and S^- to be $\{u \in S^{d-1} \mid \langle u, r\rangle < 0\}$. For any $x \in S^{d-1}$ denote its reflection with respect to H as $\sigma(x)$. Finally, given any (measurable) subset X of S^{d-1}, we define the *symmetrization* X^* of X w.r.t. H as follows. If x is a point in $X \cap S^-$ such that $\sigma(x) \notin X$ replace x with $\sigma(x)$. All other points $x \in X$ remain in X^*, and no other new points are added to X^*. Formally

$$X^* = X \setminus \{x \mid x \in X \cap S^-,\ \sigma(x) \notin X\} \cup \{\sigma(x) \mid x \in X \cap S^-,\ \sigma(x) \notin X\}$$

Notice that the measure of X and X^* are identical.

Now, let Λ be the set consisting of all pairs of closed subsets in S^{d-1}. Given two closed subsets A and B of measure $\rho(a)$, let $\lambda(A, B) = \lambda \subset \Lambda$ be the set of pairs $(\alpha, \beta) \in \Lambda$ that satisfy:

1. $\mu(A) = \mu(\alpha)$, $\mu(B) = \mu(\beta)$.
2. $\forall \varepsilon > 0 \quad \mu(A_\varepsilon) \geq \mu(\alpha_\varepsilon)$, $\mu(B_\varepsilon) \geq \mu(\beta_\varepsilon)$.
3. $E(A, B) \geq E(\alpha, \beta)$.

where for any set A, the set A_ε is defined as $\{x \in S^{d-1} \mid \exists y \in A \;\; s.t. \;\; \|x - y\| \leq \varepsilon\}$.

We start by showing that λ is closed under the two point symmetrization technique. Next we prove that λ is a closed subset of Λ (under the Hausdorff topology). Afterwards, we show that there exists a cap C of measure $\rho(a)$ such that $(C, C) \in \lambda = \lambda(A, B)$, implying that $E(A, B) \geq E(C, C)$. This concludes our proof for subsets A and B which are closed. Finally, we turn to the general case.

Lemma 9.4 *The set λ is closed under the two point symmetrization procedure (w.r.t. any $r \in S^{d-1}$ and the hyperplane H passing through the origin defined by r).*

Proof: Let H be some hyperplane passing through the origin. Let (α, β) be a pair in λ. Let α^* and β^* be the sets obtained after the symmetrization procedure w.r.t. H applied to the sets α and β. It suffices to show that $(\alpha^*, \beta^*) \in \lambda$. In [Ben84] it is shown that both α^* and β^* preserve the first two properties of λ. Hence, we turn to the third property. We need to show that the subsets α^* and β^* obtained after the symmetrization procedure share fewer edges than the original sets α and β (which in turn share fewer edges than A and B). Let x^+ and y^+ be two points in S^+, and let x^- and y^- be their corresponding mirror images. Let α' be the restriction of α to these four points (that is $\alpha' = \alpha \cap \{x^+, x^-, y^+, y^-\}$). Define β', $\alpha^{*'}$ and $\beta^{*'}$ similarly. We will show that $E(\alpha', \beta') \geq E(\alpha^{*'}, \beta^{*'})$. It is not hard to verify that this implies our assertion.

Let E' be the edge set induced by the vertices $\{x^+, x^-, y^+, y^-\}$. Recall by the definition of our continuous graph G_k, the edge set E consists of pairs of points x and y in S^{d-1} which satisfy $\langle x, y\rangle \leq -\frac{1}{k-1}$. This implies that E' is symmetric and has the following properties:

1. $(x^+, y^+) \in E' \leftrightarrow (x^-, y^-) \in E'$.
2. $(x^+, y^-) \in E' \leftrightarrow (x^-, y^+) \in E'$.
3. $(x^+, y^+) \in E' \rightarrow (x^+, y^-) \in E'$.

Consider the case in which E' consists of the edges (x^+, y^-) and (x^-, y^+) only, the subset α' is equal to $\{x^+, x^-\}$ and the subset β' is equal to $\{y^-\}$. This case in depicted in Figure 2 (a). In this specific case we have that $\alpha^{*'} = \alpha'$, and $\beta^{*'} = \{y^+\}$, implying that $E(\alpha', \beta') = E(\alpha^{*'}, \beta^{*'}) = 1$. Notice that as α' and β' are finite sets, we measure the amount of edges between α' and β' using the discrete analog of $E(\alpha', \beta')$ defined in Definition 5.1 of Section 5 (the same goes for $\alpha^{*'}$ and $\beta^{*'}$).

There are of course several other cases to consider (12 edge configurations and 256 cases of different subsets α' and β'). Some of these cases are depicted in Figure 2. This large case analysis may be significantly reduced using various observations (involving the equivalence of many different cases). We have checked our assertion on the full case analysis (using a computer program). A similar proof holds when x^+ or y^+ are in S_0. Details omitted. $\square$

Lemma 9.5 *λ is closed in Λ.*

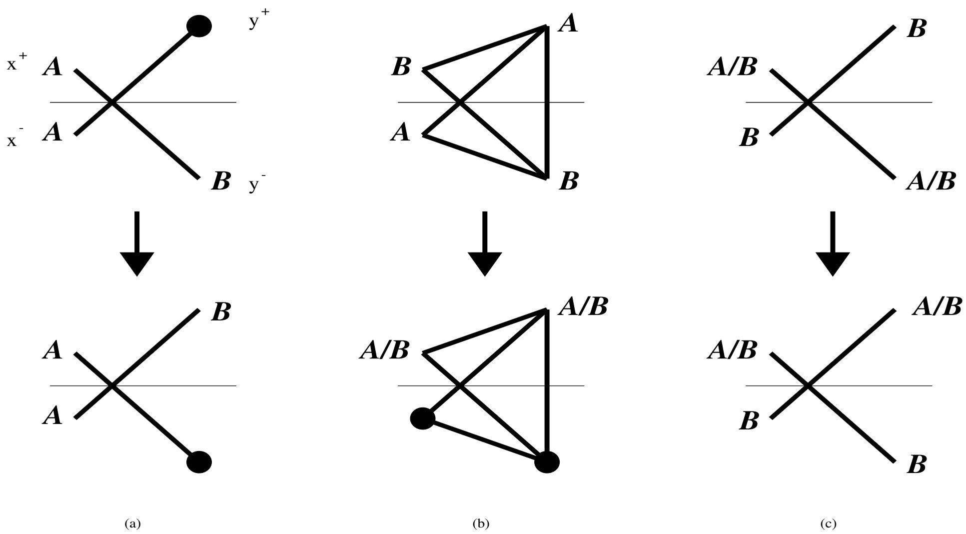

Figure 2: Three cases of E', α' and β' are presented. The set $\alpha' = \alpha \cap \{x^+, x^-, y^+, y^-\}$ and $\alpha^{*'}$ are presented by the letter A. The set $\beta' = \beta \cap \{x^+, x^-, y^+, y^-\}$ and $\beta^{*'}$ are presented by the letter B. A point in $\alpha' \cap \beta'$ ($\alpha^{*'} \cap \beta^{*'}$) is presented by A/B. A point neither in α or β is represented by a solid dot. The edge set E' is depicted by solid lines. Finally, the hyperplane H is represented as a horizontal line. Each configuration is presented before (above) and after (below) the symmetrization procedure. In case (a) the set α' is closed under the symmetrization procedure w.r.t H and $-H$. In such cases it holds that $E(\alpha', \beta') = E(\alpha^{*'}, \beta^{*'})$. In cases (b) and (c) notice that there is a difference between the value of $E(\alpha^{*'}, \beta^{*'})$ and the number of undirected edges between the sets $\alpha^{*'}$ and $\beta^{*'}$ (for example in case (b) $E(\alpha^{*'}, \beta^{*'}) = \{(x^+, y^+), (y^+, x^+)\}$ is of size 2).

Proof: Let (α_n, β_n) be a sequence in λ tending to (α, β) (in the Hausdorff topology). We will show that $(\alpha, \beta) \in \lambda$. For every $\varepsilon, \delta > 0$ we have for large enough values of n that

$$\alpha_\varepsilon \subseteq (\alpha_n)_{\varepsilon+\delta}, \quad \beta_\varepsilon \subseteq (\beta_n)_{\varepsilon+\delta}; \quad \mu((\alpha_n)_{\varepsilon+\delta}) \leq \mu(A_{\varepsilon+\delta}), \quad \mu((\beta_n)_{\varepsilon+\delta}) \leq \mu(B_{\varepsilon+\delta}).$$

Sending δ to zero and using the fact that for all closed sets α, $\mu(\alpha_\varepsilon) \to \mu(\alpha)$ we conclude the second property of λ. Namely that $\mu(\alpha_\varepsilon) \leq \mu(A_\varepsilon)$ and $\mu(\beta_\varepsilon) \leq \mu(B_\varepsilon)$.

For the first property, sending ε to zero we have on one hand that $\mu(\alpha) \leq \mu(A)$ and $\mu(\beta) \leq \mu(B)$. For the other direction observe that $\mu(\alpha_\varepsilon) \geq \mu(\alpha_n) = \mu(A)$ and $\mu(\beta_\varepsilon) \geq \mu(\beta_n) = \mu(B)$ for any $\varepsilon > 0$ provided n is large enough.

For the final property of λ, let $\theta(\varepsilon) = \mu^2(\{(x, y) \mid x \in \alpha_\varepsilon,\ y \in \beta_\varepsilon,\ (x, y) \in E\})$. Notice that $E(\alpha, \beta) \leq \theta(\varepsilon)$. Let n be large enough to ensure that $\alpha_\varepsilon \subseteq (\alpha_n)_{2\varepsilon}$, $\beta_\varepsilon \subseteq (\beta_n)_{2\varepsilon}$. For such n it is the case that $\theta(\varepsilon) \leq E((\alpha_n)_{2\varepsilon}, (\beta_n)_{2\varepsilon})$. Furthermore, as $(\alpha_n, \beta_n) \in \lambda$ for all values of n we have that $\mu((\alpha_n)_{2\varepsilon} \setminus \alpha_n) \leq \mu(A_{2\varepsilon} \setminus A)$ and $\mu((\beta_n)_{2\varepsilon} \setminus \beta_n) \leq \mu(B_{2\varepsilon} \setminus B)$. Which both tend to zero as ε tends to zero. We conclude that $E((\alpha_n)_{2\varepsilon}, (\beta_n)_{2\varepsilon})$ tends to $E(\alpha_n, \beta_n)$ as ε tends to zero *uniformly* in n.

Hence, given $\delta > 0$ one can find $\varepsilon > 0$ such that $E((\alpha_n)_{2\varepsilon}, (\beta_n)_{2\varepsilon}) \leq E(\alpha_n, \beta_n) + \delta$ for all values of n. We conclude, that for n large enough

$$E(\alpha, \beta) \leq \theta(\varepsilon) \leq E((\alpha_n)_{2\varepsilon}, (\beta_n)_{2\varepsilon}) \leq E(\alpha_n, \beta_n) + \delta \leq E(A, B) + \delta.$$

Since this holds for all δ we conclude our assertion. $\square$

Lemma 9.6 *There exists a cap C of measure $\rho(a)$ such that $(C, C) \in \lambda$.*

Proof: Fix a point $x_0 \in S^{d-1}$ and let C be the cap centered at x_0 of measure $\rho(a)$. The function $f_1(\alpha, \beta) = \mu(C \cap \alpha)$ is upper semi continuous on Λ [3]. Λ with the Hausdorff topology is a compact set, since λ is closed (Lemma 9.5) the function f_1 achieves its maximum over λ on some pair $(\alpha_{max}, \beta_{max}) \in \lambda$. In [FS02] it is shown (using Lemma 9.4) that $\alpha_{max} = C$. This is done by showing that $f_1(\alpha^*_{max}, \beta) \geq f_1(\alpha_{max}, \beta)$ with equality iff $\alpha_{max} = C$ where α^*_{max} is the subset obtained by the symmetrization procedure with respect to a hyperplane H s.t. $x_0 \notin H$; and β is arbitrary.

Now consider the function $f_2(C, \beta) = \mu(C \cap \beta)$ defined on $\Lambda' = \{(C, \beta) \mid (C, \beta) \in \Lambda\}$. The function f_2 is also upper semi continuous (this time on Λ'). Furthermore, it is shown in [FS02] that Λ' with the Hausdorff topology is a compact set and that $\lambda' = \{(C, \beta) \mid (C, \beta) \in \lambda\}$ is closed in Λ'. Therefore, the function f_2 achieves its maximum over λ' on some pair (C, β_{max}). As before it can be seen that $\beta_{max} = C$. We conclude that $(C, C) \in \lambda$. $\square$

As mentioned previously, Lemma 9.6 implies our theorem for subsets A and B which are closed. The proof for general subsets A and B follows by considering any series $\{A_i\}$ and $\{B_i\}$ of closed sets that satisfy: (a) $\forall n \;\; A_n \subseteq A$ and $B_n \subseteq B$, (b) $\mu(A_n) \to \mu(A)$ and $\mu(B_n) \to \mu(B)$, and (c) $\forall n \;\; \mu(A_n) = \mu(B_n)$. Now let C, $\{C_i\}$ be caps, all centered at the same point $x \in S^{d-1}$ such that $\mu(C) = \mu(A) = \mu(B)$ and $\forall n \;\; \mu(C_n) = \mu(A_n) = \mu(B_n)$. It now holds that

$$E(A, B) \geq E(A_n, B_n) \geq E(C_n, C_n) \to E(C, C)$$

$\square$

Claim 9.7 *Let a, z, δ be as in Claim 3.7. Let $b^2 = z^2 - a^2$. Let $N = \{(u_1, \ldots, u_d) \in S^{d-1} \mid u_1 = a,\; u_2 = b\}$. Let N_δ be a δ neighborhood of N (*i.e.* all points in S^{d-1} which are of distance less that δ from the set N). The measure of the set N_δ is at least $(1 - z^2)^{\frac{d-1}{2}}$.*

Proof: Recall that $z = \sqrt{a^2 + b^2}$. Let $\hat{N} = \{(u_1, \ldots, u_d) \in S^{d-1} \mid u_1 = z,\; u_2 = 0\}$ be the set obtained from N by rotating the unit sphere by an angle of $\arccos \frac{a}{z}$. Using spherical symmetry, we have that the measure of the set N_δ is equal to the measure of the δ neighborhood of the set $\hat{N}$ (which will be denoted as $\hat{N}_\delta$). Hence, in the following we will present a lower bound on the measure of $\hat{N}_\delta$. We start by computing the measure of the strip

$$S = \{(u_1, \ldots, u_d) \in S^{d-1} \mid u_1 \in [z - \delta^2/1000, z + \delta^2/1000]\}.$$

Using Claims 9.2 and 9.3 we conclude (for sufficiently large d)

$$\mu(S) = \rho(z - \delta^2/1000) - \rho(z + \delta^2/1000) \geq \frac{1}{2}\rho(z - \delta^2/1000) \geq \frac{1}{\delta^3}\left(1 - z^2\right)^{\frac{d-1}{2}}.$$

Let γ be a constant such that $\sqrt{1 - z^2} \geq |\gamma| > \delta/10$. Consider the set $R = \{(u_1, \ldots, u_d) \in S^{d-1} \mid u_1 = z,\; u_2 = \gamma\}$. Denote the $\delta^2/200$ neighborhood of R as R_γ. Using spherical symmetry, one can

[3] $f(\alpha)$ is upper semi continuous on Λ if for every sequence $\alpha_n \to \alpha$ in Λ it is the case that $f(\alpha)$ is at least $\limsup f(\alpha_n)$.

upper bound the measure of the set R_γ by the measure of the cap $\{(u_1,\ldots,u_d) \in S^{d-1} \mid u_1 > \sqrt{z^2+\gamma^2} - \delta^2/200\}$ which is at most

$$\left(1 - \left(\sqrt{z^2+\gamma^2} - \delta^2/200\right)^2\right)^{\frac{d-1}{2}} \le \left(1-z^2\right)^{\frac{d-1}{2}} \le \delta^3 \mu(s).$$

Let $\mathcal{R}$ be the union of the sets R_γ for $|\gamma| = \delta/10, \delta/10 + \delta^2/1000, \delta/2 + 2\delta^2/1000, \ldots, \sqrt{1-z^2}$. We conclude that the measure of $\mathcal{R}$ is at most $\frac{1}{2}\mu(S)$.

Finally, each $v = (v_1,\ldots,v_d) \in S \setminus \mathcal{R}$ satisfies $v_1 \in [z - \delta^2/1000, z + \delta^2/1000]$ and $v_2 \in [-\delta/10, \delta/10]$ and thus is in $\hat{N}_\delta$. Hence $S \setminus \mathcal{R} \subseteq \hat{N}_\delta$, implying that the measure of $\hat{N}_\delta$ is at least $\frac{1}{2}\mu(S) \ge (1-z^2)^{\frac{d-1}{2}}$. $\square$

Claim 3.7 *Let a, k be as in Theorem 3.6. Let $v = (a, \sqrt{1-a^2}, 0, \ldots, 0)$ be a vertex on the boundary of $\mathcal{C}_a$. Let $N(v)$ be the set of neighbors of v. Let $z = \sqrt{a^2 + \frac{(1/(k-1)+a^2)^2}{1-a^2}}$. Finally let $\delta = c\sqrt{\frac{\log(d)}{d}}$ for a sufficiently large constant c. The measure of vertices in $N(v) \cap \mathcal{C}_a$ satisfies*

$$(1-\delta)^{\frac{d-1}{2}} \left(1-z^2\right)^{\frac{d-1}{2}} \le \mu(N(v) \cap \mathcal{C}_a) \le \left(1-z^2\right)^{\frac{d-1}{2}}$$

Proof: The set $N(v) \cap \mathcal{C}_a$ is equal to

$$\left\{(u_1,\ldots,u_d) \in S^{d-1} \mid u_1 \ge a \text{ and } u_2 \le -\frac{1/(k-1) + a u_1}{\sqrt{1-a^2}}\right\}.$$

Denote $\frac{1/(k-1)+a^2}{\sqrt{1-a^2}}$ by $f(a)$. Let $z = \sqrt{a^2 + f(a)^2}$. By spherical symmetry, it can be seen that the above set is of measure at most $(1-z^2)^{\frac{d-1}{2}}$, and of measure greater or equal to the measure of $\Gamma = \{(u_1,\ldots,u_d) \in S^{d-1} \mid u_1 \ge z \text{ and } 0 \le u_2 \le c(u_1 - z)\}$, where $1 \ge c \ge \tan\left(\frac{\arccos(a)}{2}\right)$ (for any $a \in [0, 1/2]$, c is in the range $(1/2, 1)$). Hence, it suffices to bound the measure of Γ by below.

Let $\delta = c\sqrt{\frac{\log(d)}{d}}$ for a sufficiently large constant c. The set Γ above is of measure larger than the measure of the δ neighborhood of

$$N = \{(u_1,\ldots,u_d) \in S^{d-1} \mid u_1 = z + \delta\left(1 + \frac{2}{c}\right) \text{ and } u_2 = \delta\}$$

for c as above (the δ neighborhood of N is defined as in Claim 9.7). We denote this set as N_δ. By Claim 9.7, the measure of N_δ is at least of value $\left(1 - z^2 - \theta(\delta)\right)^{\frac{d-1}{2}}$ (notice that $z > 1/k$ which in turn is independent of d, thus Claim 9.7 can be applied in our case). Hence, the measure of $N(v) \cap \mathcal{C}_a$ is at least $(1-\theta(\delta))^{\frac{d-1}{2}} (1-z^2)^{\frac{d-1}{2}}$ which concludes our proof. $\square$

Theorem 3.8 *Let $a = \left(\frac{\log(d)}{d}\right)^{\frac{1}{4}}$. Let k satisfy $\frac{1}{k-1} = a^2$. Let $r \in S^{d-1}$, let $\mathcal{C}_a$ be an a-cap centered at r. Let $\varepsilon(a)$ be the value of $E(\mathcal{C}_a, \mathcal{C}_a)$. The value of $\varepsilon(a)$ is in the range*

$$\left[\frac{1}{poly(d)}\left(1 - \frac{2(k-1)}{k-2}a^2\right)^{\frac{d-1}{2}}, \left(1 - \frac{2(k-1)}{k-2}a^2\right)^{\frac{d-1}{2}}\right].$$

Proof: We start by stating the following claim.

Claim 9.8 *Let $\mathcal{C}_a$ be an a-cap centered at $r = (1, 0, \ldots, 0)$. Let v be the vertex $(a, \sqrt{1-a^2}, 0, \ldots, 0)$ on the boundary of $\mathcal{C}_a$. Let $N(v)$ be the set of neighbors of v. Let $z = \sqrt{a^2 + \frac{(1/(k-1)+a^2)^2}{1-a^2}}$. If $z^4 = O(\log(d)/d)$ then the measure of vertices in $N(v)$ which are in the cap $\mathcal{C}_a$ satisfies*

$$\frac{1}{poly(d)} \left(1 - z^2\right)^{\frac{d-1}{2}} \leq \mu(N(v) \cap \mathcal{C}_a) \leq \left(1 - z^2\right)^{\frac{d-1}{2}}$$

Proof: The set $N(v) \cap \mathcal{C}_a$ is equal to

$$\left\{(u_1, \ldots, u_d) \in S^{d-1} \mid u_1 \geq a \text{ and } u_2 \leq -\frac{1/(k-1) + au_1}{\sqrt{1-a^2}}\right\}.$$

Denote $\frac{1/(k-1)+a^2}{\sqrt{1-a^2}}$ by $f(a)$. Let $z = \sqrt{a^2 + f(a)^2}$. By spherical symmetry, it can be seen that the above set is of measure at most $(1-z^2)^{\frac{d-1}{2}}$, and of measure greater or equal to the measure of $\Gamma = \{(u_1, \ldots, u_d) \in S^{d-1} \mid u_1 \geq z \text{ and } 0 \leq u_2 \leq c(u_1 - z)\}$, where $1 \geq c \geq \tan\left(\frac{\arccos(a)}{2}\right)$. Hence, it suffices to bound the measure of Γ by below.

Let $r_1, \ldots, r_d$ be independent standard normal variables, let $s = \sqrt{\sum_{i=1}^d r_i^2}$, and let $s_3 = \sqrt{\sum_{i=3}^d r_i^2}$. Let $\phi(x)$ be the standard normal density function. The measure of Γ is given by the probability

$$\begin{aligned}
\mu(\Gamma) &= \Pr\left[\frac{r_1}{s} \geq z \text{ and } 0 \leq \frac{r_2}{s} \leq c(\frac{r_1}{s} - z)\right] \\
&\geq \Pr\left[r_1 \geq \sqrt{\frac{z^2(r_2^2 + s_3^2)}{1-z^2}} \text{ and } 0 \leq r_2 \leq A\right] \\
&\geq \Pr[s_3^2 \leq d] \Pr\left[r_1 \geq \sqrt{\frac{z^2(r_2^2 + d)}{1-z^2}} \text{ and } 0 \leq r_2 \leq A_d\right] \\
&\geq \frac{1}{4} \Pr\left[r_1 \geq \sqrt{\frac{z^2 d}{1-z^2}} \text{ and } 0 \leq r_2 \leq \min(A_d, B)\right] \\
&\geq \frac{1}{4} \Pr\left[\sqrt{\frac{z^2 d}{1-z^2}} \leq r_1 \leq \sqrt{\frac{z^2(d+1)}{1-z^2}} \text{ and } 0 \leq r_2 \leq \min(A_d, B)\right]
\end{aligned}$$

Where $A = c\frac{r_1 - z\sqrt{r_1^2(1+c^2-c^2z^2)+s_3^2(1-c^2z^2)}}{1-c^2z^2}$, $A_d = c\frac{r_1 - z\sqrt{r_1^2(1+c^2-c^2z^2)+d(1-c^2z^2)}}{1-c^2z^2}$, and B is of value $\sqrt{\frac{r_1^2(1-z^2)}{z^2} - d}$. Notice that we have used the fact that $\Pr[s_3^2 \leq d] \geq 1/4$.

We proceed by evaluating $\min(A_d, B)$. Both A_d and B are functions of r_1. Furthermore r_1^2 is in the range $\left[\frac{z^2 d}{1-z^2}, \frac{z^2(d+1)}{1-z^2}\right]$. Let $x \in [0, 1]$, we can represent A_d and B as functions of x by setting r_1 to be $\sqrt{\frac{z^2(d+x)}{1-z^2}}$. That is $A_d = z\frac{\sqrt{d+x} - \sqrt{d+xz^2(1+c^2-c^2z^2)}}{\sqrt{1-z^2}(1-c^2z^2)}$ and $B = \sqrt{x}$. It is not hard to verify that for our value of z it is the case that $\min(A_d, B) = A_d$.

$$
\begin{aligned}
\mu(\Gamma) &\geq \frac{1}{4}\Pr\left[\sqrt{\frac{z^2 d}{1-z^2}} \leq r_1 \leq \sqrt{\frac{z^2(d+1)}{1-z^2}} \text{ and } 0 \leq r_2 \leq A_d\right] \\
&= w\int_{r_1\in R_1}\int_{r_2=0}^{A_d} e^{-\frac{r_1^2+r_2^2}{2}} \geq w\int_{r_1\in R_1}\frac{e^{-\frac{r_1^2}{2}}}{A_d}\int_{r_2=0}^{A_d} r_2 e^{-\frac{r_2^2}{2}} \\
&= w\int_{r_1\in R_1}\frac{e^{-\frac{r_1^2}{2}}}{A_d}(1-e^{-\frac{A_d^2}{2}}) \geq w\int_{r_1\in R_1} A_d e^{-\frac{r_1^2}{2}} \\
&\geq w\int_{x=0}^{1} A_d e^{-\frac{z^2(d+x)}{2(1-z^2)}}\frac{z}{2\sqrt{1-z^2}\sqrt{d+x}}dx \\
&\geq w\frac{z^2}{d(1-z^2)(1-c^2z^2)}e^{-\frac{dz^2}{2(1-z^2)}}\int_{x=0}^{1} xe^{-\frac{z^2x}{2(1-z^2)}} \geq w\frac{z^2}{d(1-z^2)(1-c^2z^2)}e^{-\frac{dz^2}{2(1-z^2)}} \\
&\geq w\frac{z^2}{d(1-z^2)(1-c^2z^2)}\left(1-\frac{z^2}{1-z^2}\right)^{\frac{d-1}{2}} \\
&\geq w\frac{z^2}{d(1-z^2)(1-c^2z^2)}\left(1-z^2\right)^{\frac{d-1}{2}}\left(1-2z^4\right)^{\frac{d-1}{2}}
\end{aligned}
$$

Where R_1 is the range $\left[\sqrt{\frac{z^2 d}{1-z^2}}, \sqrt{\frac{z^2(d+1)}{1-z^2}}\right]$, and w is some constant independent of d (the value of w may change from line to line). Above we use the fact that $e^x > 1+x$, and the fact that A_d is bounded by above by a constant independent of d. In the final inequalities we use a change of variables $r_1 = \sqrt{\frac{z^2(d+x)}{1-z^2}}$, and use the fact that $1+\alpha/3 \leq \sqrt{1+\alpha} \leq 1+\alpha/2$ for $\alpha < 3$. As $\mu(\Gamma) \geq \frac{1}{poly(d)}\left(1-z^2\right)^{\frac{d-1}{2}}$ for $z^4 = O(\log(d)/d)$, we conclude our assertion. □

For the proof of Theorem 3.8, let $z = \sqrt{a^2 + \frac{(1/(k-1)+a^2)^2}{1-a^2}}$ (as in Claim 9.8), thus $z^4 = O(\log(d)/d)$. For the upper bound, use a bound on the measure of $\mathcal{C}_a$, and the upper bound in Claim 9.8. As for the lower bound, let $\delta = 2\left(\frac{\log(d)}{d}\right)^{\frac{3}{4}}$. Such a choice of δ will satisfy $a\delta \geq 2\frac{\log(d)}{d}$, and $z\delta = O(\log(d)/d)$.

Let $w = (w_1, w_2, \ldots, w_d) \in \mathcal{C}_a$ with first coordinate w_1 of value $a+\delta$. Consider a vertex $v = (v_1, v_2, \ldots, v_d) \in \mathcal{C}_a$ with first coordinate v_1 of value less than $a+\delta$. By Claim 9.8, it is not hard to verify that the measure of $N(v)\cap\mathcal{C}_a$ is greater than the measure of $N(w)\cap\mathcal{C}_a$ which is greater than $\frac{1}{poly(d)}\left(1-cz\delta)\right)^{\frac{d-1}{2}}\left(1-z^2\right)^{\frac{d-1}{2}} \geq \frac{1}{poly(d)}\left(1-z^2\right)^{\frac{d-1}{2}}$ for some constant c (we use Claim 9.2 and the fact that $z\delta = O(\log(d)/d)$). As $a\delta \geq 2\log(d)/d$, we conclude using Corollary 9.3, that the measure of edges in $E(\mathcal{C}_a, \mathcal{C}_a)$ is at least

$$(\rho(a)-\rho(a+\delta))\frac{1}{poly(d)}\left(1-z^2\right)^{\frac{d-1}{2}} \geq \frac{\rho(a)}{poly(d)}\left(1-z^2\right)^{\frac{d-1}{2}} \geq \frac{1}{poly(d)}\left(1-a^2\right)^{\frac{d-1}{2}}\left(1-z^2\right)^{\frac{d-1}{2}}.$$

Simplifying the above expression with the help of Claim 9.2, we conclude our assertion. □

Acknowledgments

We would like to thank Luca Trevisan for his suggestion to analyze *edge* sampling.

References

[AK98] N. Alon and N. Kahale. Approximating the independence number via the θ function. *Mathematical Programming*, 80(2):253–264, 1998.

[AK02] N. Alon and M. Krivelevich. Testing k-colorability. *SIAM Journal on Discrete Mathematics*, 15(2):211–227, 2002.

[Ben84] Y. Benyamini. Two point symmetrization, the isoperimetric inequality on the shpere and some applications. *Longhorn Notes, Univ. of Texas, Texas Funct. Anal. Seminar*, pages 53–76, 1983-1984.

[Bez02] S. Bezrukov. Edge-isoperimetric problems of graphs. *To appear in J'anos Bolyai Math. Series*, 2002.

[BK97] A. Blum and D. Karger. An $o(n^{3/14})$-coloring for 3-colorable graphs. *Information Processing Letters*, 61(6):49–53, 1997.

[BT76] A. Baernstein and A. Taylor. Spherical rearrangments, subharmonic functions, and *-functions in n-space. *Duke Math J.*, 43:245–268, 1976.

[Cha02] M. Charikar. On semidefinite programming relaxations for graph coloring and vertex cover. *Proceedings of SODA*, pages 616–620, 2002.

[Fei97] U. Feige. Randomized graph products, chromatic numbers, and the Lovász theta function. *Combinatorica*, 17(1):79–90, 1997.

[Fei02] U. Feige. Approximating maximum clique by removing subgraphs. *Manuscript*, 2002.

[Fel66] W. Feller. *An introduction to probability theory and its applications*, volume 2. John Wiley & Sons, 1966.

[FK98] U. Feige and J. Kilian. Zero knowledge and the chromatic number. *Journal of Computer and System Sciences*, 57:187–199, 1998.

[FS02] U. Feige and G. Schechtman. On the optimality of the random hyperplane rounding technique for MAX CUT. *Random Structures and Algorithms*, 20(3):403–440, 2002.

[GGR98] O. Goldreich, S. Goldwasser, and D. Ron. Property testing and its connection to learning and approximation. *Journal of ACM*, 45(4):653–750, 1998.

[GK98] M. Goemans and J. Kleinberg. The Lovász theta function and a semidefinite programming relaxation of Vertex Cover. *SIAM Journal on Discrete Mathematics*, 11:196–204, 1998.

[Hal93] M. Halldorsson. A still better performance guarantee for approximate graph coloring. *Information Processing Letters*, 45:19–23, 1993.

[Hås99] J. Håstad. Clique is hard to approximate within $n^{1-\varepsilon}$. *Acta Math*, 182(1):105–142, 1999.

[HNZ01] E. Halperin, R. Nathaniel, and U. Zwick. Coloring k-colorable graphs using smaller palettes. *Proceedings of SODA*, pages 319–326, 2001.

[Kho01] S. Khot. Improved inapproximability results for max clique, chromatic number and approximate graph coloring. *In Proceedings of 42nd IEEE Symposium on Foundations of Computer Science*, 2001.

[KKL88] J. Kahn, G. Kalai, and N. Linial. The influence of variables on boolean functions. *In proceedings of 29th Symposium on the Foundations of Computer Science*, pages 68–80, 1988.

[KMS98] D. Karger, R. Motwani, and M. Sudan. Approximate graph coloring by semidefinite programming. *Journal of ACM*, 45(2):246–265, 1998.

[Lov79] L. Lovász. On the Shannon capacity of a graph. *IEEE Transactions on Information Theory*, 25:2–13, 1979.

[Sze94] M. Szegedy. A note on the θ number of Lovász and the generalized Delsarte bound. *In proceedings of 35th Annual Symposium on Foundations of Computer Science*, pages 36–39, 1994.