New Approximation Guarantee for Chromatic Number

Sanjeev Arora*
arora@cs.princeton.edu

Eden Chlamtac†
chlamtac@cs.princeton.edu

Moses Charikar‡
moses@cs.princeton.edu

Department of Computer Science
Princeton University
Princeton, NJ 08540

ABSTRACT

We describe how to color every 3-colorable graph with $O(n^{0.2111})$ colors, thus improving an algorithm of Blum and Karger from almost a decade ago. Our analysis uses new geometric ideas inspired by the recent work of Arora, Rao, and Vazirani on SPARSEST CUT, and these ideas show promise of leading to further improvements.

Categories and Subject Descriptors: F.2 [Theory of Computation]: Analysis of Algorithms and Problem Complexity

General Terms: Algorithms, Theory.

Keywords: graph coloring, chromatic number, semidefinite programming, approximation algorithms.

1. INTRODUCTION

In the graph k-coloring problem we wish to assign each vertex one of k colors such that every pair of vertices connected with an edge are assigned different colors. This problem arises in a host of applications, and was one of the 22 NP-complete problems on Karp's list in 1972. Subsequently, much effort was spent on trying to design efficient approximation algorithms, namely, given a k-colorable graph to try to color it with as few colors as possible. Unfortunately, probabilistically checkable proofs (PCPs) have been used to show that if certain reasonable complexity conjectures are true then in general we may need as many as $n^{1-\epsilon}k$ colors[9]. But the case of small k, including $k = 3$ is still wide open, and this is the problem we consider. It is conceivable that a polynomial-time algorithm could color every 3-colorable graph with say 6 colors. (Recently Dinur et. al. [6] gave some evidence that such an algorithm may require an unbounded number of colors, but they still do not rule out the possibility that $\log\log n$ colors may suffice.)

Wigderson [17] showed how to color 3-colorable graphs with $\tilde{O}(\sqrt{n})$ colors. (Throughout the paper, $\tilde{O}(\cdot)$ notation is used to suppress polylogarithmic factors.) Blum [2] improved this to $\tilde{O}(n^{3/8})$. After Goemans and Williamson [8] used semidefinite programming (SDP) to design approximation algorithms for MAX-CUT and MAX-SAT [8], Karger, Motwani, and Sudan [11] used SDP to improve Blum's guarantee to $\tilde{O}(n^{1/4})$. Blum and Karger [3] then noticed that the ideas of KMS and Blum could be combined to further improve this guarantee to $\tilde{O}(n^{3/14})$, which is where things have stood for a decade. Our paper improves this bound to $O(n^{0.2133})$ using a new analysis of the SDPs introduced by Karger, Motwani, Sudan. Using stronger SDPs we further improve the bound to $O(n^{0.2111})$.

The most important contribution of this paper is not so much the improvement in the exponent, but a demonstration that sophisticated geometric reasoning can allow further progress on this problem. A priori this was not clear. For example, given the simplicity of the Goemans-Williamson analyses of SDPs, one might conjecture that it does not give the best possible algorithms for most problems. Therefore it has been a surprise to discover that the algorithms obtained for MAX-CUT (namely, 0.878..-approximation [8]) and MAX-3SAT (7/8-approximation [12]) are actually best possible assuming certain plausible complexity assumptions [10, 13]. On the other hand, for EXPANSION and SPARSEST CUT problems, the longstanding approximation ratio of $O(\log n)$ [15] turned out to be not tight. Arora, Rao and Vazirani [1] recently improved the ratio to $O(\sqrt{\log n})$ by a more sophisticated analysis of SDPs with triangle inequality constraints. Though this algorithmic breakthrough was quickly followed by new hardness results [14, 5], a large gap remains and the exact status of the problems is still open.

Our new algorithm for graph coloring and its analysis is directly inspired by techniques of Arora, Rao and Vazirani. In particular, our analysis is nonlocal and uses their "walk" argument (also called "chaining" argument). We recall that the analysis of Karger, Motwani, Sudan (which was inspired by the Goemans-Williamson analysis of MAX-CUT) is local in the sense that it analyses the effect of rounding on each edge and then just uses linearity of expectations to estimate the effect on the entire graph. By contrast, our nonlocal analysis shows that the local analysis of KMS cannot be simultaneously tight for many edges. (See Section 3.) To state our improvement, we need the following definition.

*Supported by NSF grants MSPA-MCS 0528414, CCF 0514993, ITR 0205594.

†Supported by a Francis Upton fellowship.

‡Supported by NSF ITR grant CCR-0205594, NSF CAREER award CCR-0237113, MSPA-MCS 0528414 and an Alfred P. Sloan Fellowship.

STOC'06, May21–23, 2006, Seattle, Washington, USA.

Definition 1. For a graph $G = (V, E)$ with vertex set $V = \{1, 2, \ldots, n\}$, a *vector θ-coloring* is an assignment of unit vectors $v_1, \ldots, v_n \in \mathbb{R}^n$ to the vertices, such that:

$$\forall (i,j) \in E: \quad (v_i, v_j) \leq -\frac{1}{\theta - 1}. \tag{1}$$

The vector θ-coloring is *strict* when equality holds in condition (1).

As is shown in [11], for any $k \geq 2$, every k-colorable graph is also vector k-colorable, and moreover a vector k-coloring can be found in polynomial time using semidefinite programming. Their main new result was that given a graph of maximum degree Δ and a vector 3-coloring, their rounding algorithm (see algorithm KMS in Section 2) can be used to produce an $\tilde{O}(\Delta^{1/3})$-coloring. They derive their $n^{1/4}$-coloring algorithm by applying the above algorithm for graphs where $\Delta < n^{3/4}$ and applying Wigderson's algorithm on the other graphs. Blum and Karger achieved a better performance by applying the above algorithms for graphs where $\Delta < n^{9/14}$ and switching to Blum's algorithm for other graphs.

Feige, Langberg, and Schechtman [7] subsequently showed that the KMS bound of $\tilde{O}(\Delta^{1/3})$ cannot be improved, but they could only show this for values of Δ far smaller than $n^{9/14}$, the value used in Blum-Karger's algorithm. This left open the question whether the KMS bound is tight for all Δ. We answer this in the negative: we show that the KMS bound can be improved to $\Delta^{0.3308}$ for $\Delta \geq n^{0.6446}$. To do this we need to slightly change the rounding algorithm (compare algorithms KMS and KMS$'$ in Section 2), and apply our nonlocal analysis; see Section 4. Plugging into Blum and Karger's scheme, this already gives a slightly better coloring.

Then we observe that our analysis can be strengthened if we make more fundamental changes to the SDP and rounding algorithm. This improved algorithm and its analysis are described in Section 6.

The rest of the paper is organized as follows. In Section 2 we describe the KMS approach and state our first improvement. In Section 3 we give a high-level description of our nonlocal analysis; this description is not precise. In Section 4 we give the simplest attempt to formalize our intuition, which already gives some improvement to KMS. This analysis is first done for the case of strict vector 3-coloring, which is then generalized in Section 5 to cover the case of nonstrict vector θ-coloring for any $\theta \leq 3$. In Section 6 we describe a rounding scheme for a stronger SDP and its analysis. Finally, in Appendix B we state a geometric conjecture, which, if true, would imply an $O(n^{0.1991})$ coloring algorithm.

We use Blum's standard terminology to discuss coloring algorithms, specifically, the idea that it is sufficient for the algorithm to make one of three types of *progress*, as described in Appendix A. As an aside, we note that we had to slightly modify Blum's ideas to get our best algorithms, and our improved (and somewhat simplified) version of Blum's technique is summarized in Theorem 12.

2. KMS ROUNDING AND OUR RESULTS

Recall that the *standard normal distribution* has density function $\frac{1}{\sqrt{2\pi}} e^{-x^2/2}$. A random vector $\zeta = (\zeta_1, \ldots, \zeta_n)$ is said to have the *n-dimensional standard normal distribution* if the components ζ_i are independent and each have the standard normal distribution. Note that this distribution is invariant under rotation. In particular, for any unit vector $x \in \Re^n$, the projection (ζ, x) has the standard normal distribution. Moreover, for any orthogonal subspaces $U, W \subset \Re^n$, the projections of ζ onto U, W, respectively, are independent.

We use the following notation for the tail bound of the standard normal distribution.

$$N(x) \stackrel{\text{def}}{=} \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \tag{2}$$

The following property of the normal distribution will be crucial.

LEMMA 1. *For $\varepsilon > 0$, and for fixed constant $\kappa > 1$, we have*

$$N(\kappa \cdot \varepsilon) = \Theta(\text{polylog}(1/\text{N}(\varepsilon)) \cdot \text{N}(\varepsilon)^{\kappa^2}).$$

For the remainder of this paper we will concern ourselves with vector coloring of 3-colorable graphs, but the results for vector k coloring can be easily generalized for any fixed $k > 3$. For $k = 3$ we will also consider stronger SDPs that incorporate the strict vector 3-coloring constraints as well as the following *odd cycle constraints*:

$$\forall i: \text{ odd cycle } C, \quad \sum_{j \in C} (v_i, v_j) \leq \frac{|C| - 3}{4} \tag{3}$$

These constraints are valid since in the intended (3-coloring) solution, every odd cycle must have three consecutive vertices with three different colors, and of the $|C| - 3$ vertices in the remaining path, at most half can have the same color as v_i.

In this section we recall the KMS rounding algorithm and its analysis. As is standard, we assume that in order to find colorings with $\tilde{O}(s(n))$ colors, it suffices to find independent sets of size $n/s(n))$. Furthermore, we only concentrate on the case where there is a bound Δ on the maximum degree; see Section A in the appendix for how to turn such a guarantee into an algorithm whose performance is stated in terms of n. From now on $\{v_i\}_{i \in V}$ will denote a vector coloring of graph G, and Δ denotes the maximum degree. For simplicity, v_i will stand both for the vector and the corresponding graph vertex.

The KMS rounding algorithm is as follows:

KMS$(G, \{v_i\}, \varepsilon)$

- Choose $\zeta \in \mathbb{R}^n$ from the n-dimensional standard normal distribution.
- $V_\zeta(\varepsilon) \stackrel{\text{def}}{=} \{i | (\zeta, v_i) \geq \varepsilon\}$. Return all $v \in V_\zeta(\varepsilon)$ with no neighbors in $V_\zeta(\varepsilon)$.

THEOREM 1 (KMS). *There exists some $\varepsilon = \varepsilon(n, \Delta) > 0$ such that the expected size of the independent set returned by algorithm KMS $(G, \{v_i\}, \varepsilon)$ is $\tilde{\Omega}(\Delta^{-1/3} \cdot n)$.*

We consider the following slight variation of the KMS rounding algorithm:

KMS$'(G, \{v_i\})$

- For "all" $\varepsilon \in \mathbb{R}$, do the following, and return the largest independent set found:
  - Choose $\zeta \in \mathbb{R}^n$ from the n-dimensional standard normal distribution.
  - Pick any edge $< v_i, v_j >$ with both endpoints in $V_\zeta(\varepsilon)$, and eliminate both v_i and v_j. Repeat until no such edges are left.
  - Return all remaining vertices in $V_\zeta(\varepsilon)$.

Remark 1. Equivalently, we can first choose ζ, and then enumerate over all relevant values of ε (that is, over $\varepsilon_i = (\zeta, v_i)$). However, for the purposes of the analysis, we will consider the first formulation.

Note that the set returned by KMS$'$ contains the set returned by KMS, so Theorem 1 holds also for KMS$'$. The crucial difference is that in KMS$'$, the vertices removed from $V_\zeta(\varepsilon)$ form a matching, and this will be used in the simple "pruning" argument of Lemma 3.

Crucially, when $\{v_i\}$ is a (possibly non-strict) vector 3-coloring, and the maximum degree is bounded by $\Delta = n^{0.6446}$, KMS$'(G, \{v_i\})$ returns an independent set of size $\Omega(\Delta^{0.3308})$ (by Theorem 4). Combining this result with the Blum coloring tools (see Theorem 12), immediately yields the following result:

THEOREM 2. *For 3-colorable graphs, one can find an $O(n^{0.2133})$ coloring in polynomial time.*

Finally, using the odd-cycle constraints (3), we give a better rounding algorithm as described in Section 6. Setting $\Delta = n^{0.6481}$, we combine the analysis of this final algorithm, summarized in Theorem 11, with the Blum coloring tools to obtain the following:

THEOREM 3. *For 3-colorable graphs, one can find an $O(n^{0.2111})$ coloring in polynomial time.*

As in [1], our analysis involves the analysis of certain *well spread-out* sets of vectors called (η, δ)-covers (defined in Section 4) and the use of measure concentration. We are interested in the structure of *efficient* covers, i.e. those where the number of vectors is close to the minimum possible. In contrast to [1] however, we believe that new geometric theorems related to measure concentration of such covers would improve the current analysis of our rounding algorithms. In Appendix B, we state such a conjecture, which, if true, would yield an improvement to our bounds.

3. KMS ANALYSIS AND HIGH-LEVEL DESCRIPTION OF OURS

Now we recall the proof of Theorem 1 from [11] —but rephrased in our terminology. Then we outline our improved analysis. Details of our analysis appear in later sections.

Note that, for any choice of ε (in either KMS or KMS$'$), for any v in the vector coloring, $\Pr[v \in V_\zeta(\varepsilon)] = N(\varepsilon)$. Say a vertex/vector is *good* for a certain value of ε if in the KMS algorithm,

$$\Pr[v \text{ is eliminated} \mid v \in V_\zeta(\varepsilon)] \leq 1/2, \tag{4}$$

and otherwise call the vertex *bad*. If v is good, then the probability it ends up in the final independent set is at least $\frac{1}{2}\Pr[v \in V_\zeta(\varepsilon)] = N(\varepsilon)/2$.

Now we analyze what makes a vertex good. For any $v \in V$, let $\{u_i\}$ be its neighbors in G. Then

$$\Pr[v \text{ is eliminated} \mid v \in V_\zeta(\varepsilon)] = \Pr[\exists i (\zeta, u_i) \geq \varepsilon | (\zeta, v) \geq \varepsilon]. \tag{5}$$

Since v and u_i are neighbors in a strict vector 3-coloring, we can write $u_i = -\frac{1}{2}v + \frac{\sqrt{3}}{2}u_i'$ where u_i' is a unit vector orthogonal to v. Writing $u_i' = \frac{2}{\sqrt{3}}(u_i + \frac{1}{2}v)$, we see that for any vector ζ:

$$(\zeta, v) \geq \varepsilon \text{ and } \quad (\zeta, u_i) \geq \varepsilon \quad \Longrightarrow \quad (\zeta, u_i') \geq \sqrt{3}\varepsilon.$$

Hence the right hand side of (5) is bounded above by

$$\begin{aligned}
\Pr[\exists i (\zeta, u_i) \geq \varepsilon | (\zeta, v) \geq \varepsilon] &\leq \Pr[\exists i : (\zeta, u_i') \geq \sqrt{3}\varepsilon | (\zeta, v) \geq \varepsilon)] \\
&= \Pr[\exists i : (\zeta, u_i') \geq \sqrt{3}\varepsilon] \\
&\leq \sum_i \Pr[(\zeta, u_i') \geq \sqrt{3}\varepsilon] \\
&\leq \Delta \cdot N(\sqrt{3}\varepsilon) \\
&= \tilde{O}(\Delta \cdot N(\varepsilon)^3)
\end{aligned}$$

Choose ε so that $N(\sqrt{3}\varepsilon) = \tilde{\Theta}(N(\varepsilon)^3)$ is less than $1/2\Delta$ (equivalently, Δ is less than $\tilde{\Theta}(N(\varepsilon))^{-3}$), in which case every vertex is good. Therefore the output independent set has expected size at least $N(\varepsilon)n/2 = \tilde{\Omega}(\Delta^{-1/3}n)$.

Before proceeding, we state one corollary of the above proof that will be useful in our analysis of KMS$'$.

LEMMA 2. *Let $\{u_i\}$ be the neighbors of v in a strict 3-vector coloring. Then writing $u_i = -\frac{1}{2}v + \frac{\sqrt{3}}{2}u_i'$ for each i, we have*

$$\Pr[\exists j : (\zeta, u_i) \geq \varepsilon | (\zeta, v) \geq \varepsilon] \leq \Pr[\exists j : (\zeta, u_i') \geq \sqrt{3}\varepsilon].$$

Under what conditions is the above analysis tight? The analysis of when a vertex is good is locally tight, even though it uses the union bound. Our main contribution is a nonlocal argument that shows that the local analysis cannot be simultaneously tight for all vertices for this value of ε. Thus the KMS$'$ algorithm can use a smaller ε than the KMS paper did, which increases $\Omega(nN(\varepsilon))$, the size of the final independent set. If less than $n/2$ vertices are bad, the expected size of the independent set is at least $nN(\varepsilon)/4$. We show below that there is such an ε satisfying $\Delta > N(\sqrt{3}\varepsilon)^{-(1+c)}$ for some $c > 0$. Thus the size of the independent set $N(\varepsilon)n = \tilde{\Omega}(\Delta^{-1/3(1+c)}n)$, an improvement over KMS.

Our nonlocal argument is directly inspired by the "walk" argument of Arora, Rao, and Vazirani. However, our walks are only of length $O(1)$ whereas theirs were longer. To illustrate our idea, let us first assume that the vectors in the SDP solution are "nondegenerate," by which we mean that their pairwise inner products do not exhibit any statistically significant patterns apart from those implied by the SDP constraints. To give an example, if v is any vertex/vector, then the constraints of strict vector coloring require the vectors u_i' defined above to be orthogonal to v. In a nondegenerate solution, we also expect that for any arbitrary unit vector v_0, most of the vectors u_i' should only have negligible projection on v_0. In fact, if we assume the KMS analysis is tight, we know this is the case (see Lemma 5).

We give a heuristic argument why the KMS$'$ algorithm should return an independent set of size $\tilde{\Omega}(n^{8/9})$ in a non-degenerate solution[1]. Specifically, let ε be the smallest value such that for at least 1/2 the vertices v:

$$\Pr_{\zeta}[v \text{ gets eliminated} | v \in V_\zeta(\varepsilon)] \geq 1/2. \tag{6}$$

We show that $N(\varepsilon) \geq \tilde{\Omega}(n^{-1/9})$.

First, a simple pruning argument (see Lemma 3; the only place where we use the difference between KMS and KMS$'$)) allows us to assume that condition (6) holds for *all* vertices in the graph rather than just half the vertices (with the probability 1/2 replaced by a smaller constant). Hence Lemma 2 implies that for every vertex v, its neighbors $\{u_i\}$ satisfy

$$\Pr[\exists i : (\zeta, u_i') \geq \sqrt{3}\varepsilon] = \Omega(1). \tag{7}$$

Sets of vectors of the above form are called $(\sqrt{3}\varepsilon, \Omega(1))$-covers in [1] (see Definition 3 later). For a $1 - o(1)$ fraction of gaussian vectors ζ, (ζ, v) is negligible compared to ε (since v is a fixed vector) so we conclude that $(\zeta, u_i) \geq \frac{3}{2}\varepsilon(1-o(1))$ with high probability.

Now consider all the neighbors of such a u_i, say $\{w_{ij}\}$, and write each of them as $w_{ij} = -\frac{1}{2}u_i + \frac{\sqrt{3}}{2}w_{ij}'$. By our assumption, the vectors w_{ij}' satisfy the condition in (7). Symmetry implies that for most gaussians ζ, there is some w_{ij}' such that $(\zeta, w_{ij}') \leq -\sqrt{3}\varepsilon$. Such a w_{ij} satisfies:

$$(\zeta, w_{ij}) = -\frac{1}{2}(\zeta, u_i) + \frac{\sqrt{3}}{2}(\zeta, w_{ij}') \tag{8}$$

$$= -\frac{3\varepsilon(1-o(1))}{4} - \frac{3\varepsilon}{2} = -2.25\varepsilon(1-o(1)). \tag{9}$$

Note that here we are using nondegeneracy strongly, since we are assuming that the union of the events "$[(\zeta, u_i') \geq \sqrt{3}\varepsilon] \wedge [\exists j : (\zeta, w_{ij}') \leq -\sqrt{3}\varepsilon]$" happens with good probability when the individual events have large probability. This would be true if the events "$(\zeta, u_i') \geq \sqrt{3}\varepsilon$" were disjoint. It turns out (see Theorem 6) that the nondegeneracy assumption is enough to make a slightly weaker claim of the same form.

One can continue this argument with neighbors of w_{ij} and so on, ultimately deducing that for $\Omega(1)$ fraction of gaussians ζ, there is some vector y such that $(\zeta, y) \geq 3\varepsilon(1-o(1))$. (In general, given that some vector has a projection s on ζ, this argument shows the existence of a neighbor whose projection is $\geq (s/2 + 3\varepsilon/2)(1-o(1))$, which is larger than s so long as $s < 3\varepsilon$.)

Since the number of vertices in the graph is only n, the union bound implies that for a standard normal vector ζ the expected number of y such that $(\zeta, y) \geq 3\varepsilon$ is at most $nN(3\varepsilon)$. We conclude that $nN(3\varepsilon) = \Omega(1)$, and hence $nN(\varepsilon)^9 = \tilde{\Omega}(1)$.

Of course, the above analysis ignores all conditioning between the probability calculations in successive steps of the argument, which is justifiable only when the vectors are non-degenerate. In the correct argument such conditioning cannot be ignored, though intuition suggests that if the vectors are not degenerate the KMS algorithm does better anyway.

[1]The integrality gaps of [7] rule out the possibility of finding independent sets of size $\Omega(n^{0.843})$ in vector 3-colorable graphs in general. Thus, there must be an inherent loss in going from this heuristic argument to a rigorous proof.

Formalizing this intuition seems to require difficult geometric arguments in general, but a simpler analysis is possible in the regime where all parameters are close to tight, which happens to be the case in the KMS analysis as noted earlier. Specifically, the $(\sqrt{3}\varepsilon, \delta)$-covers are almost "efficient," in the sense that the number of vectors in them, namely, Δ, is almost as low as it could be, which is $N(\sqrt{3}\varepsilon)^{-1}$. This fact allows some handle on the conditioning.

We believe that it should be possible to improve our analysis, although doing so seems to call for some new geometric theorems.

4. DETAILS OF OUR SIMPLER ANALYSIS

In this section we prove Theorem 2 using a two-step walk analysis of KMS$'$, which will also serve as the basis for a more elaborate analysis in later sections. For simplicity, we will assume strict vector 3-coloring here, a condition which we relax later. We will use the following definition to quantify the improvement.

Definition 2. Given a graph G with vector 3-coloring $\{v_i\}$ and maximum degree $\leq \Delta$, the parameter $\varepsilon > 0$ is *c-inefficient for* $(G, \{v_i\})$ if $N(\sqrt{3}\varepsilon) = (8\Delta)^{-1/(1+c)}$.

Using this terminology, we can give the following guarantee on the performance of KMS', which in turn implies Theorem 2.

THEOREM 4. *For every $\tau > \frac{5}{9}$ there exists $c_1(\tau) > 0$ such that for all $c < c_1(\tau) - o(1)$, and any graph G with maximum degree $\leq n^\tau$, if the parameter ε is (at most) c-inefficient for $(G, \{v_i\})$, then $KMS'(G, \{v_i\})$ returns an independent set of expected size $\Omega(N(\varepsilon)n)$. Furthermore, $c_1(\tau)$ satisfies*

$$c_1(\tau) \stackrel{\text{def}}{=} \sup\left\{ c \,\middle|\, \min_{|\alpha| \leq \sqrt{c/(1+c)}} \lambda_c(\alpha) > \sqrt{\frac{1+c}{\tau}} \right\}, \tag{10}$$

where

$$\lambda_c(\alpha) \stackrel{\text{def}}{=} \left(3 - \alpha - 2\sqrt{1-\alpha^2}\sqrt{c}\right) / \sqrt{5 - 2\alpha - 3\alpha^2}. \tag{11}$$

The rest of this section is devoted to proving Theorem 4. The proof is by contradiction: if $\Delta < N(\sqrt{3}\varepsilon)^{1/(1+c)}$ then, as in Section 3, we use a chaining argument to exhibit a high-probability event that is actually very unlikely.

We will simplify things by first arguing that if at least half the vertices $v \in V$ are bad (i.e., their probability of being eliminated from $V_\zeta(\varepsilon)$ is more than 1/2), then we can focus on a subgraph in which all vertices are almost-bad.

LEMMA 3. *In $KMS'(G, \{v\}, \varepsilon)$, if*

$$\Pr[v \textit{ is eliminated} \mid v \in V_\zeta(\varepsilon)] \geq 1/2$$

for at least $n/2$ vertices $v \in V$, then there is a non-empty induced subgraph $G' = (V', E')$ of G such that for all $v \in V'$ we have

$$\Pr[v \textit{ is eliminated with a neighbor in } G' \mid v \in V_\zeta(\varepsilon)] \geq 1/8.$$

PROOF. The proof uses a pruning argument from [1]. Consider the graph of all matching edges (for all choices of ζ), with edges weighted by probability of elimination. Denote by $\gamma(v)$ the total weight of edges incident to v. Eliminate, one after the other, vertices with $\gamma(v) < N(\varepsilon)/8$. Since the initial total edge weight in the graph is $\frac{1}{2}\sum_v \gamma(v) \geq$

$\frac{1}{2} \cdot \frac{n}{2} \cdot N(\varepsilon)/2 = n \cdot N(\varepsilon)/8$, and the total edge weight eliminated is $< n \cdot N(\varepsilon)/8$, there must be positive edge weight left. The remaining graph is therefore non-empty and has the desired property. $\square$

We adapt the following definition from [1], to use normal distributions rather than unit vectors.

Definition 3. A set of vectors $\{x_1, \ldots, x_k\}$ is said to be a (η, δ)*-cover*, if for $\zeta \in \Re^n$ chosen from the standard normal distribution,

$$\Pr[\exists i : (\zeta, x_i) \geq \eta] \geq \delta.$$

The cover is *uniform* if the $\{x_i\}$ are all unit vectors.

Definition 4. A uniform (η, δ)-cover $\{x_1, \ldots, x_k\}$ is said to be (at most) *c-inefficient*, if $\frac{k}{\delta} \leq N(\eta)^{-(1+c)}$.

Note that any uniform (η, δ)-cover must contain at least $\delta \cdot N(\eta)^{-1}$ vectors, by union bound. Hence, a cover is efficient when the number of vectors is only slightly larger than the minimum required. Thus we have the following corollary of Lemma 3 (using Lemma 2 and its notation).

LEMMA 4. *If the input to KMS′ has degree of inefficiency c, and at least half the vertices are bad, then there is a non-empty subgraph in which the neighbors $\{u_i\}$ of any vertex v are such that $\{u_i'\}$ form a c-inefficient $(\sqrt{3}\varepsilon, \frac{1}{8})$-cover.*

Efficient covers have the property that their vectors are not concentrated near any single vector, as the following lemma shows.

LEMMA 5. *Let $\{x_i\}$ be a uniform (η, δ)-cover which is at most c-inefficient, where $N(\eta)$ is small relative to δ (specifically $N(\eta) \leq \delta^{f(\eta)}$ where f is a superconstant function), and let v_0 be any unit vector. Then the subset $\{x_i : |(x_i, v_0)| \leq \sqrt{c/1+c} + o(1)\}$ is a $(\eta, \delta - o(\delta))$-cover.*

PROOF. For every i, let $\alpha_i = (x_i, v_0)$, and write $x_i = \alpha_i v_0 + \sqrt{1-\alpha_i^2} x_i'$, where $x_i' \perp v_0$. Let X be the subset of $\{x_i\}$ in the lemma statement, and consider the set of indices not participating in the cover in X, that is, $I_c = \{i : |\alpha_i| \geq \sqrt{c/(1+c)} + o(1)\}$. We will show that the set $\{x_i : i \in I_c\}$ is at most a $(\eta, o(\delta))$-cover. For any $\rho \geq 0$ we have

$$\begin{aligned}
\Pr[\exists i \in I_c : (\zeta, x_i) \geq \eta] &\leq 2\Pr[(\zeta, v_0) \geq \rho\eta] \\
&\quad + \Pr[|(\zeta, v_0)| \leq \rho\eta \wedge \exists i : (\zeta, x_i) \geq \eta] \\
&\leq 2N(\rho\eta) \\
&\quad + \Pr\left[\exists i : (\zeta, x_i') \geq \frac{(1-|\rho|)\eta}{\sqrt{1-\alpha_i^2}}\right] \\
&\leq 2N(\rho\eta) \\
&\quad + \delta N(\eta)^{1+c} \max_{i \in I_c} N\left(\frac{(1-|\rho|)\eta}{\sqrt{1-\alpha_i^2}}\right)
\end{aligned}$$

where the last inequality follows from efficiency of $\{x_i\}$. Now, let $\rho = \frac{(\log(1/\delta))^{1/4}}{\sqrt{\eta}}$. Since we have assumed that $\delta = N(\eta)^{o(1)}$, we have $\rho = o(1)$ and $N(\rho\eta) = o(\delta)$. Also by the lower bound on $|\alpha_i|$ for $i \in I_c$, and by Lemma 1, it follows that $\max_{i \in I_c} N\left(\frac{(1-|\rho|)\eta}{\sqrt{1-\alpha_i^2}}\right) = o(N(\eta)^{1+c})$, which completes the proof. $\square$

The above lemma directly implies the following structural theorem.

THEOREM 5. *Let v, u be unit vectors with $(v, u) = -\frac{1}{2}$, and $u = -\frac{1}{2}v + \frac{\sqrt{3}}{2}u'$, $u' \perp v$, and let $\{t_j\}$ be a c-inefficient (η, δ)-cover, for some $\delta \geq \eta^{-o(1)}$, and $t_j \perp u$ for all j. Then rewriting each t_j as $t_j = \alpha_j(\frac{\sqrt{3}}{2}v + \frac{1}{2}u') + \sqrt{1-\alpha_j^2}t_j'$ for some $\alpha \in [-1, 1]$ and unit vector t_j' orthogonal to both v and u', we have*

$$\Pr\left[\exists j : (\zeta, t_j') \geq \frac{(1-o(1))\eta}{\sqrt{1-\alpha_j^2}}\right] \geq (1-o(1))\delta.$$

This bound holds even when the probability is restricted to j s.t. $|\alpha_j| \leq \sqrt{\frac{c}{1+c}} + o(1)$.

PROOF. The decomposition of t_j follows immediately from the observation that $0 = (t_j, u) = -\frac{1}{2}(t_j, v) + \frac{\sqrt{3}}{2}(t_j, u')$. Lemma 5 implies that $\{t_j\}$ is a $(\eta, (1-o(1))\delta)$-cover even when restricted to indices $J = \{j : |\alpha_j| \leq \sqrt{\frac{c}{1+c}} + o(1)\}$. Now let $v' = \frac{\sqrt{3}}{2}v + \frac{1}{2}u'$, and choose $\rho = (\log\eta)^{3/4}$. Then we have

$$\begin{aligned}
\delta &\leq \Pr[\exists j \in J : (\zeta, t_j) \geq \eta] \\
&\leq 2\Pr[(\zeta, v') \geq \rho] + \Pr\left[\exists j \in J : (\zeta, t_j') \geq \frac{\eta - |\alpha_j|\rho}{\sqrt{1-\alpha_j^2}}\right] \\
&\leq 2N(\rho) + \Pr\left[\exists j \in J : (\zeta, t_j') \geq \frac{\eta - |\alpha_j|\rho}{\sqrt{1-\alpha_j^2}}\right] \\
&\leq o(\delta) + \Pr\left[\exists j \in J : (\zeta, t_j') \geq \frac{(1-o(1))\eta}{\sqrt{1-\alpha_j^2}}\right].
\end{aligned}$$

$\square$

Consider $v \in V$ with neighbors $\{u_i\}$ and 2-neighborhood $\{w_j\}$ (all neighbors of various u_i). Write $u_i = \frac{1}{2}v + \frac{\sqrt{3}}{2}u_i'$ and, for u_i and w_j neighbors, $w_j = \frac{1}{2}u_i + \frac{\sqrt{3}}{2}w_{ij}'$. Applying Theorem 5 for fixed u_i and $t_j = w_{ij}'$ gives

$$\begin{aligned}
w_j &= -\frac{1}{2}u_i + \frac{\sqrt{3}}{2}w_{ij}' \\
&= \frac{1}{4}v - \frac{\sqrt{3}}{4}u_i' + \frac{\sqrt{3}}{2}w_{ij}' \\
&= \frac{1}{4}v - \frac{\sqrt{3}}{4}u_i' + \frac{\sqrt{3}}{2}\left(\alpha_j\left(\frac{\sqrt{3}}{2}v + \frac{1}{2}u_i'\right) + \sqrt{1-\alpha_j^2}w_{ij}''\right) \\
&= \left(\frac{1}{4} + \frac{3}{4}\alpha_j\right)v - \frac{\sqrt{3}}{4}(1-\alpha_j)u_i' + \frac{\sqrt{3}}{2}\sqrt{1-\alpha_j^2}w_{ij}''
\end{aligned} \tag{12}$$

for some unit vectors $\{w_{ij}''\}$ all orthogonal to v and u_i'. Note that by this decomposition $(w_j, v) = \frac{1}{4} + \frac{3}{4}\alpha_j$, so α_j depends only on v and w_j, justifying the notation α_j instead of α_{ij}. Theorem 5 also tells us that, for any fixed u_i, if the vectors $\{w_{ij}'\}$ constitute a uniform $(\sqrt{3}\varepsilon, \frac{1}{8})$-cover, then $\{\sqrt{1-\alpha_j^2}w_{ij}''\}$ form a (non-uniform) $(\sqrt{3}\varepsilon - o(\varepsilon), \frac{1}{8} - o(1))$-cover. We now use measure concentration, and the efficiency of the cover of $\{u_i'\}$, to compose the $\{u_i'\}$ cover with the

$\{w''_{ij}\}$ covers as in the last section. We start by showing how covers can be "boosted" using measure concentration.

LEMMA 6 (MEASURE CONCENTRATION). *Let $\{y_j\}$ be a non-uniform $(\eta, N(\theta))$-cover. Then for any $\zeta \in \Re^n$ having standard normal distribution, and $\rho \geq 0$,*

$$\Pr[\exists j : (\zeta, y_j) \geq \eta - \|y_j\| \rho] \geq N(\theta - \rho).$$

PROOF. Let $\gamma_n(\cdot)$ denote the normalized Gaussian measure on R^n. The theorem of measure concentration for Gauss space ([4], [16]) states that for any measurable set $A \subseteq \Re^n$, if $\gamma_n(A) = N(\theta)$ for $\theta \in R$, then for any $\rho \geq 0$ the set $A_\rho = \{\zeta | \exists z \in A : \|\zeta - z\| \leq \rho\}$ has measure at least $N(\theta - \rho)$.

Let $A = \{\zeta | \exists i : (\zeta, y_j / \|y_j\|) \geq \eta / \|y_j\|\}$. By our assumption, this set has measure at least $N(\theta)$. Since $\{y_j / \|y_j\|\}$ are unit vectors, one can readily verify that in this case $A_\rho = \{\zeta | \exists i : (\zeta, y_j / \|y_j\|) \geq \eta / \|y_j\| - \rho\}$. Applying measure concentration, the claim follows immediately. □

We now use this lemma to prove a cover composition theorem.

THEOREM 6 (COVER COMPOSITION). *Let $\{x_i\}$ be a (uniform) c-inefficient (η_1, δ_1)-cover, and for each i, let $\{y_{ij}\}_j$ be a (non-uniform) (η_2, δ_2)-cover such that $y_{ij} \perp x_i$. Then*

$$\begin{aligned}&\Pr\left[\exists i, j : ((\zeta, x_i) \geq \eta_1) \wedge \right.\\ &\qquad \left.\left((\zeta, y_{ij}) \geq \eta_2 - \|y_{ij}\| \cdot \left(N^{-1}(\delta_2) + (\sqrt{c} + o(1)) \cdot \eta_1\right)\right)\right]\\ &\geq \delta_1 - o(\delta_1).\end{aligned}$$

PROOF. If we associate with every vector x_i the halfspace $\{\zeta | (\zeta, x_i) \geq \eta_1\}$, then the set corresponding to the cover is just the union of halfspaces. The idea is to upper-bound the measure of points in each halfspace not participating in the relevant set. Formally, let $\rho = \rho(c, \eta_1)$ to be determined later. Then for each i we define

- $H_i = \{\zeta | (\zeta, x_i) \geq \eta_1\}$
- $Z_i = \{\zeta | \forall j : (\zeta, y_{ij}) \leq \eta_2 - \|y_{ij}\| \cdot \left(N^{-1}(\delta_2) + \rho\eta_1\right)\}$

Note that the theorem concerns a lower bound on the probability of the event $\zeta \in \cup_i (H_i \setminus Z_i)$. First, we have to bound the measure of each Z_i. For any fixed i, applying Lemma 6 to the cover of $\{y_{ij}\}_j$, we get that $\gamma_n(Z_i) \leq N(\rho\eta_1)$.

Using this bound, the independence of orthogonal components of a gaussian vector, and the efficiency of $\{x_i\}$, we get

$$\begin{aligned}&\Pr\left[\exists i, j : ((\zeta, x_i) \geq \eta_1) \wedge \right.\\ &\qquad \left.\left((\zeta, y_{ij}) \geq \eta_2 - \|y_{ij}\| \cdot \left(N^{-1}(\delta_2) + \rho\eta_1\right)\right)\right]\\ &= \gamma_n\left(\cup_i (H_i \setminus Z_i)\right)\\ &\geq \gamma_n\left(\cup_i H_i\right) - \gamma_n\left(\cup_i \left(H_i \cap Z_i\right)\right)\\ &\geq \delta_1 - \sum_i \gamma_n\left(H_i \cap Z_i\right)\\ &= \delta_1 - \sum_i N(\eta_1) \cdot \gamma_n\left(Z_i\right)\\ &\geq \delta_1 - \delta_1 \cdot N(\eta_1)^{-(1+c)} \cdot N(\eta_1) \cdot \max_i \gamma_n\left(Z_i\right)\\ &\geq \delta_1 - \delta_1 \cdot N(\eta_1)^{-c} \cdot N(\rho\eta_1)\end{aligned}$$

From Lemma 1 we have that $N(\rho\eta_1) = o(N(\eta_1)^c)$ for $\rho = \sqrt{c} + \Theta\left(\log \eta_1 / {\eta_1}^2\right)$, and the theorem follows. □

Remark 2. It is worth mentioning at this point that we think a much stronger version of Theorem 6 may be true. Specifically, the analysis in the proof may be tight, however, we believe, not when the $\{y_{ij}\}$ covers are themselves efficient. In that case, we believe that the $\sqrt{c}$ factor could be replaced by $O(c)$, which would be a much smaller price to pay for small constant c. The improvement relies on a conjecture regarding measure concentration of efficient covers which would replace Lemma 6. We discuss this conjecture later.

We can use the composition theorem to obtain a result reminiscent of the chaining argument in [1]: it shows that whenever KMS′ fails (in expectation), for some large η we can find a $(\eta, \Omega(1))$-cover containing few vectors.

THEOREM 7. *Let $G = (V, E)$ be an n-vertex graph with maximum degree Δ, and identify the vertices V with a strict vector 3-coloring. Then if at least half the vertices v are bad for KMS′(G, V, ε) with degree of inefficiency c, then there is some $v \in V$ and some subset of its 2-neighborhood $\{w_j\} \subseteq \Gamma(\Gamma(v))$ such that*

1. *Each w_j can be written $w_j = (\frac{1}{4} + \frac{3}{4}\alpha_j)v + t_j$ for some vector $t_j \perp v$, and α_j satisfying $|\alpha_j| \leq \sqrt{c/(1+c)} + o(1)$.*
2. *With probability at least $\frac{1}{8} - o(1)$, there is some j for which*

$$(\zeta, t_j) \geq (3 - \alpha_j - 2\sqrt{1 - \alpha_j^2}\sqrt{c} - o(1))\sqrt{3}\varepsilon.$$

PROOF. Take any $v \in V$, and let $\{u_i\}$ and $\{w_j\}$ be its neighbors, and neighbors of neighbors, respectively. Write $u_i = \frac{1}{2}v + \frac{\sqrt{3}}{2}u'_i$ and, for u_i and w_j neighbors, $w_j = \frac{1}{2} + \frac{\sqrt{3}}{2}w'_{ij}$. Then we may assume, by Lemma 4, that $\{u'_i\}$ and the sets $\{w'_{ij}\}_j$ for every i are all uniform $(\sqrt{3}\varepsilon, \frac{1}{8})$-covers which are at most c-inefficient.

For each i, we now apply Theorem 5 for the cover $\{w'_{ij}\}_j$, to obtain the vectors $\{\sqrt{1 - \alpha_j^2}w''_{ij}\}_j$ which are orthogonal to v and u_i, and form a non-uniform $((1-o(1))\sqrt{3}\varepsilon, \frac{1}{8} - o(1))$-cover when restricted to j for which $|\alpha_j| \leq \sqrt{c/(1+c)} + o(1)$.

Applying Theorem 6 for $x_i = -u'_i$ and $y_{ij} = \sqrt{1 - \alpha_j^2}w''_{ij}$, and incorporating low order terms, we get

$$\begin{aligned}&\Pr\left[\exists i, j : ((\zeta, u'_i) \leq -\sqrt{3}\varepsilon) \wedge \right.\\ &\qquad \left.\left((\zeta, w''_{ij}) \geq \left(1/\sqrt{1 - \alpha_j^2} - \sqrt{c} - o(1)\right)\sqrt{3}\varepsilon\right)\right]\\ &\geq \frac{1}{8} - o(1)\end{aligned}$$

Using the decomposition in equation 12, for any w_j in the 2-neighborhood of v we can write $w_j = \left(\frac{1}{4} + \frac{3}{4}\alpha_j\right)v + t_j$, where $t_j = -\frac{\sqrt{3}}{4}(1 - \alpha_j)u'_i + \frac{\sqrt{3}}{2}\sqrt{1 - \alpha_j^2}w''_{ij}$ for any u_i which is a neighbor of both v and w_j, and the theorem then follows.

□

Now we prove Theorem 4 for the case of strict vector 3-coloring.

PROOF OF THEOREM 4 (FOR STRICT COLORING). Let c be the degree of inefficiency of the input, and suppose, for the

sake of contradiction, that at least half the vertices v are bad. Applying Theorem 7, we obtain a set of at most n vectors $\{t_j\}$ so that with constant probability, some t_j has projection at least $\frac{\sqrt{3}}{4}(3-\alpha_j-2\sqrt{1-\alpha_j^2}\sqrt{c}-o(1))\sqrt{3}\varepsilon$ for some α_j s.t. $|\alpha_j| \le \sqrt{c/(1+c)}+o(1)$.

Let $\hat{t}_j = t_j/\|t_j\|$. Noting that $\|t_j\| = \frac{\sqrt{3}}{4}\sqrt{5-2\alpha_j-3\alpha_j^2}$, we have

$$\begin{aligned}\Omega(1) &\le \Pr[\exists j : (\zeta, \hat{t}_j) \ge (\lambda_c(\alpha_j)-o(1))\sqrt{3}\varepsilon] \\ &\le \sum_j N((\lambda_c(\alpha_j)-o(1))\sqrt{3}\varepsilon) \\ &\le n \cdot \max_j N((\lambda_c(\alpha_j)-o(1))\sqrt{3}\varepsilon) \\ &\le n \cdot N\left(\min_{|\alpha|\le\sqrt{c/(1+c)}}(\lambda_c(\alpha)-o(1))\cdot\sqrt{3}\varepsilon\right)\end{aligned}$$

Define $t(c) = \min_{|\alpha|\le\sqrt{c/(1+c)}} \lambda_c(\alpha)$, and note that $\lim_{c\to 0} t(c) = \frac{3}{\sqrt{5}}$. Therefore, for $\tau > \frac{5}{9}$, there is some solution $c_1 > 0$ to equation (10). Moreover, for c which is at most $c_1(\tau)-o(1)$, we have $t(c)^2 \ge \frac{1+c}{\tau}+o(1)$. Thus, using the degree of inefficiency of the input, the above inequality gives

$$\begin{aligned}\Omega(1) &\le n \cdot N\left(\sqrt{3}\varepsilon\right)^{t(c)^2-o(1)} \\ &\le n \cdot \Delta^{-\frac{t(c)^2-o(1)}{1+c}} \\ &\le n \cdot \Delta^{-1/\tau}\cdot o(1) = o(1)\end{aligned}$$

which is a contradiction. □

3-step walks. The above technique can be generalized to a three step argument as follows: Consider Theorem 7 as one step (ignoring the existence of intermediate vertices $\{u_i\}$), and use a variant of the two step walk argument above where the second step is as in Theorem 7. This approach yields the following result when combined with Theorem 12:

THEOREM 8. *For 3-colorable graphs, one can find an $O(n^{0.212299..})$ coloring in polynomial time.*

5. NON-STRICT VECTOR COLORING AND VECTOR CHROMATIC NUMBER < 3

We sketch a generalization of our analysis (specifically, the analysis in section 4) which applies to KMS′ when the vector 3-coloring in the input is not necessarily strict. The result is the same as given in Section 4 for strict vector 3 coloring. When the input graph has vector chromatic number strictly < 3, this generalization yields a further improvement.

First, we adapt the notion of inefficiency of input of algorithm KMS' to the present setting:

Definition 5. Given a graph G with vector $(1+\frac{1}{a})$-coloring $\{v_i\}$ and maximum degree $\le \Delta$, the parameter $\varepsilon > 0$ is *c-inefficient for* $(G, \{v_i\})$ if $N(\sqrt{\frac{1+a}{1-a}}\varepsilon) = (8\Delta)^{-1/(1+c)}$.

THEOREM 9. *Let $a \ge \frac{1}{2}$. For every $\tau > \frac{1+a^2}{(1+a)^2}$ there exists $c_1(\tau) > 0$ such that for all $c < c_1^a(\tau) - o(1)$, and any graph G with maximum degree $\le n^\tau$, if the parameter ε is (at most) c-inefficient for $(G, \{v_i\})$, then $KMS'(G, \{v_i\})$ returns an independent set of expected size $\Omega(N(\varepsilon)n)$. Furthermore, $c_1^a(\tau)$ satisfies*

$$c_1^a(\tau) \stackrel{\text{def}}{=} \sup\left\{c \,\middle|\, \min_{|\alpha|\le\sqrt{c/(1+c)}} \lambda_c^a(\alpha) > \sqrt{\frac{1+c}{\tau}}\right\}, \tag{13}$$

where

$$\lambda_c^a(\alpha) \stackrel{\text{def}}{=} \frac{1+a(1-\alpha)-\sqrt{1-\alpha^2}\sqrt{c}}{\sqrt{(1-\alpha)(a^2(1-\alpha)+1+\alpha)}}.$$

For a "one-step analysis" (e.g. the original KMS result), it is clear if the inner product between neighbors is $< -\frac{1}{2}$ then the analysis only improves. The difficult case in the two step walk argument seems to occur when we walk from v to u to w, where $(v,u) < (u,w)$. This case can be avoided using a simple binning argument.

PROOF OF THEOREM 9. As before, we may assume, by Lemma 3, that every vertex can be eliminated with probability $\Omega(1)$. Consider a partition of the edges into $\log n$ bins $\{E_k\}$ by inner product of endpoints, i.e. $E_k = \{vu \in E : -(v,u) \in [\frac{1}{2}+\frac{k}{\log n}, \frac{1}{2}+\frac{k+1}{\log n})\}$. For every v there is one bin E_k such that

$$\Pr[v \text{ is eliminated with } u \in \Gamma_{E_k}(v) \mid v \in V_\zeta(\varepsilon)] = \Omega\left(\frac{1}{\log n}\right).$$

Let us concentrate on the (directed) subgraph of these edges.

To simplify the argument, assume that for each v there is a fixed value $a_v \in [a,1)$ such that $(v,u) = -a_v$ for all neighbors u of v in the relevant subgraph. We can do this because the resulting error terms (projections along vectors of norm $O(\frac{1}{\log n})$) are negligible. Now, let v be the vector for which this a_v is smallest (w.l.o.g. $a_v = a$), let $\{u_i\}$ be its neighbors, and $\{w_j\}$ the neighbors of $\{u_i\}$, where $w_j = -b_i u_i + \sqrt{1-b_i^2}w'_{ij}$ (for w_j neighbors of u_i). Arguing as before, one can show

$$\begin{aligned}w_j = &\left(ab_i + \alpha_{ij}\sqrt{(1-a^2)(1-b_i^2)}\right)v - \\ &\left(b_i\sqrt{1-a^2} - \alpha_{ij}a\sqrt{1-b_i^2}\right)u'_i + \sqrt{(1-\alpha_{ij}^2)(1-b_i^2)}w''_{ij}\end{aligned}$$

where w.l.o.g. $|\alpha_{ij}| \le \sqrt{c/(1+c)}$. Moreover, the above decomposition satisfies

$$\begin{aligned}\Pr\Bigg[\exists i,j : &\left((\zeta, u'_i) \le -\sqrt{\frac{1+a}{1-a}}\,\varepsilon\right) \wedge \\ &\left((\zeta, w''_{ij}) \ge \rho_{ij}\sqrt{\frac{1+b_i}{1-b_i}}\,\varepsilon\right)\Bigg] = \tilde{\Omega}(1)\end{aligned}$$

for some

$$\rho_{ij} = 1/\sqrt{1-\alpha_{ij}^2} - \sqrt{c} - o(1).$$

Now, letting t_j be the component of w_j orthogonal to v, it suffices to show that the above probability implies a projection of at least $\min_{i,j}\lambda_c(\alpha_{ij})$ for $t_j/\|t_j\|$. For brevity, write $t_j = -\theta_{ij}u'_i + \kappa_{ij}w''_{ij}$. Since the individual components $-u'_i$ and w''_{ij} have projection at least $\sqrt{(1+a)/(1-a)}\varepsilon$ and $\rho_{ij}\sqrt{(1+a)/(1-a)}\varepsilon$, respectively, and noting that $\rho_{ij} < 1$,

we only need to show that

$$\frac{\theta_{ij}}{\kappa_{ij}} \geq a\sqrt{\frac{1-\alpha_{ij}}{1-\alpha_{ij}^2}}$$

(the corresponding ratio for strict vector $(1+\frac{1}{a})$-coloring). But this follows easily from the assumption that $a \leq b_i$. $\square$

6. A BETTER ROUNDING ALGORITHM

In this section we present a rounding algorithm for strict vector 3-coloring with odd-cycle constraints, for which we are able to prove a better guarantee, namely, Theorem 3.

Let us first introduce some notation. Let $\{v_i\}$ be a strict vector 3-coloring of graph G. For any set of vertices S, we define $\Gamma(S) = \cup_{v\in S}\Gamma(v)$. For every v_i and $\theta > 0$, define

$$V_i^\theta \stackrel{\text{def}}{=} \{v_j \in \Gamma(\Gamma(v_i)) \setminus \{v_i\} : (v_i, v_j) \geq \theta\}.$$

We will denote by G_i^θ the subgraph of G induced on V_i^θ. For all $v_i \neq v_j$, we define the components normal to v_i as

$$t_j^i = \frac{v_j - (v_i, v_j)v_i}{\|v_j - (v_i, v_j)v_i\|}.$$

Observation 1. Let $\theta > 0$ and vectors $\{v_i\}$ be a vector 3-coloring. Then for any vector v_i, the set $\{t_j^i | v_j \in V_i^\theta\}$ is a (non-strict) vector $3/(1+2\theta^2)$-coloring of V_i^θ.

Observation 2. Let $\alpha \in \mathbb{R}$ and vectors $\{v_i\}$ be a vector 3-coloring satisfying the odd cycle constraints, and let $\theta = \frac{1}{4} + \frac{3}{4}\alpha$. Then we have the following:

1. If $\alpha > 0$ then V_i^θ is an independent set.
2. For any odd positive integer $k \geq 3$, if $\alpha > -1/k$ then V_i^θ contains no odd cycles of length $\leq k$.

Proof. In case (1), an edge w_1w_2 (where w_1 and w_2 are neighbors of $u1, u2 \in \Gamma(v_i)$, respectively) would violate the 5-cycle constraint with respect to cycle $\{v, u_1, w_1, w_2, u_2\}$ and vector v_i. Case (2) is an immediate application of the odd cycle constraints to potential odd cycles in V_i^θ and vector v_i. $\square$

So far, our analysis has relied on constructing $(\eta, \Omega(1))$-covers for large η, in order to contradict the assumption that KMS$'$ fails on certain inputs. We have always made the pessimistic assumption that these covers may contain up to $\Omega(n)$ vectors. Clearly the analysis is improved if the number of these vectors is much less. Otherwise, assuming some such cover has many vectors, we use the fact that these covers are associated with original vectors (from the vector coloring) $\{w_j\}$ of the form $w_j = \theta_j v + \sqrt{1-\theta_j^2}w_j'$ for $\theta_j \approx \frac{1}{4}$. This is a kind of non-degeneracy, namely, the vectors $\{w_j\}$ have a large projection in a common direction. In this case G_i^θ has no short odd cycles and a vector $3-\delta$-coloring (for some $\delta = \Omega(1)$) and we exploit that.

Now, we present the following rounding algorithm, which gives a better guarantee than KMS$'$ when the graph contains no short cycles of odd length. Here, G denotes an n-vertex graph.

KMS-Aux$(G, \{v_i\})$

- If G is an independent set, output all vertices. Otherwise, let $V' = \{v_i\}$, and construct set I as follows:
- For every v_i, let

  $k_i = \max\{k | v_i, \Gamma(v_i), \ldots, \Gamma^k(v_i)$ are all ind. sets$\}$.
- Let v_i and $l \in \{0, \ldots, k_i\}$ be such that the ratio $\left|\Gamma^{l+1}(v_i)\right| / \left|\Gamma^l(v_i)\right|$ is minimized. Add $\Gamma^l(v_i)$ to I. If at most $n/2$ vertices remain in V' after removing $\Gamma^l(v_i)$ and $\Gamma^{l+1}(v_i)$, remove only $\Gamma^l(v_i)$. Otherwise, remove both sets and repeat.
- Let G' be the subgraph of G induced on V'. Return the larger of I and KMS$'(G', V')$.

The following result describes the performance guarantee of algorithm KMS-Aux.

Theorem 10. *Let $a \geq \frac{1}{2}$, and let $k \geq 2$ be an integer. Then for any graph G with no odd cycles of length $\leq 2k+1$ we have the following:*

1. *The set I produced by KMS-Aux$(G, \{v_i\})$ has size $\Omega(n^{1-1/(k+1)})$.*
2. *For every $\tau > \left(\frac{1+a^2}{(1+a)^2} + (k-1)\cdot\frac{1-a}{1+a}\right)^{-1}$ there exists $c_k^a(\tau) > 0$ such that for all $c < c_1(\tau)$ if graph G has non-strict $1+\frac{1}{a}$-coloring $\{v_i\}$, maximum degree $\leq n^\tau$, then KMS-Aux produces an independent set of size $\Omega(N(\varepsilon)n)$, for ε which is c-inefficienct for $(G, \{v_i\})$.*

Furthermore, $c_k^a(\tau)$ satisfies

$$c_k^a \stackrel{\text{def}}{=} \sup\left\{c \,\middle|\, \min_{\alpha^2 \leq c/(1+c)} \lambda_c^a(\alpha)^2 > \frac{1+c}{\tau} - (k-1)\cdot\frac{1-a}{1+a}\right\}.$$

Proof. To prove (1), note that for any vector v, the sets $\{v\}$, $\Gamma(v)$,...,$\Gamma^k(v)$ are all independent sets (otherwise, an edge between vertices of $\Gamma^j(v)$ together with paths to v constitute an odd cycle of length $2j+1$). Moreover one of the $k+1$ ratios $\frac{|\Gamma(v)|}{|\{v\}|}, \ldots, \frac{\left|\Gamma^{k+1}(v)\right|}{\left|\Gamma^k(v)\right|}$ must be at most $n^{1/(k+1)}$. Hence all the sets added to I are a $O(n^{-1/k})$-fraction of the sets removed from V'. Since $\left|\cup(\Gamma^l(v_i) \cup \Gamma^{l+1}(v_i))\right| \geq n/2$, it must be the case that $I = \Omega(n^{1-1/(k+1)})$.

To prove (2), assume first that while constructing I, there is always some vector v with 2-neighborhood of size $|\Gamma(\Gamma(v))| \geq N(\varepsilon)^{k-1}\cdot n$. This implies that for such a vector v, one of the $k-1$ ratios $\frac{\left|\Gamma^3(v)\right|}{\left|\Gamma^2(v)\right|}, \ldots, \frac{\left|\Gamma^{k+1}(v)\right|}{\left|\Gamma^k(v)\right|}$ is at most $1/N(\varepsilon)$. As above, this implies that $|I| = \Omega(N(\varepsilon)n)$.

Otherwise, after I is constructed, all 2-neighborhoods in G' have size at most $N(\varepsilon)^{k-1}\cdot n$. We may assume that $|V'| \geq n/4$, since otherwise $|I| = \Omega(n)$. Assume for the sake of contradiction that KMS$'(G', V')$ does not produce an independent set of expected size $\Omega(N(\varepsilon)n)$. Then, following the analysis in Section 5, there exists a set $\{t_i\}$ of unit vectors associated with a 2-neighborhood in G', such that $\Pr\left[\exists i : (\zeta, i) \geq \lambda_c^a(\alpha_i)\cdot\sqrt{\frac{1+a}{1-a}}\right] = \Omega(1)$ for some $|\alpha_i| \leq \sqrt{c/(1+c)}$. By our choice of c, we have that $\lambda_c^a(\alpha) > \sqrt{\frac{1+c}{\tau} - (k-1)\cdot\frac{1-a}{1+a} + \rho}$ for some constant $\rho > 0$. Hence,

we have $|\{t_i\}| \cdot N(\varepsilon)^{\frac{1+c}{\tau}-(k-1)\cdot\frac{1-a}{1+a}+\rho} = \tilde{\Omega}(1)$. But this contradicts our assumptions regarding the inefficiency of ε and size of 2-neighborhoods in G'. $\square$

Using the earlier observations, we have the following corollary of Theorem 10.

COROLLARY 1. *Let $\alpha > -\frac{1}{3}$, and define $\theta_\alpha \stackrel{\text{def}}{=} \frac{1}{4} + \frac{3}{4}\alpha$,*

$$a_\alpha \stackrel{\text{def}}{=} \frac{3+2\alpha+3\alpha^2}{5-2\alpha-3\alpha^2}, \qquad \text{and} \qquad k_\alpha \stackrel{\text{def}}{=} \left\lceil \frac{-3-1/\alpha}{2} \right\rceil.$$

Let G be a graph with vector 3-coloring $\{v_i\}$ which satisfies odd cycle constraints. Then for any v_i, if the maximum degree in G is bounded by $|V_i^{\theta_\alpha}|^{\tau_i^\alpha}$, KMS-Aux$(G_i^{\theta_\alpha}, \{t_j^i | v_j \in V_i^{\theta_\alpha}\})$ returns an independent set of size $\Omega\left(|V_i^{\theta_\alpha}|^{\varphi(\alpha,\tau_i^\alpha)}\right)$, where

1. *for $\alpha \geq 0$, $\varphi(\alpha,\tau) = 1$, and*
2. *for $-\frac{1}{3} < \alpha < 0$,*

$$\varphi(\alpha,\tau) = 1 - \min\left\{\frac{1}{k_\alpha}, \frac{1-a_\alpha}{1+a_\alpha}\cdot\frac{\tau}{1+c_{k_\alpha}^{a_\alpha}(\tau)} - o(1)\right\}.$$

(Here we set $c_k^a(\tau) = 0$ when otherwise undefined.)

Now, we present our final rounding algorithm for strict vector 3-coloring.

KMS$''(G, \{v_i\})$

- Run KMS$'(G, \{v_i\})$, and for every v_i and $\theta > 0$, run KMS-Aux$(G_i^\theta, \{t_j^i | v_j \in V_i^\theta\})$. Return the largest of these independent sets.

Using the notation from Corollary 1, we state the main result regarding the performance of KMS".

THEOREM 11. *For every $\tau > \frac{15}{32}$ there exists $c_1'(\tau) > 0$ such that for all $c < min\{\frac{1}{8}, c_1'(\tau)\}$, if the input to KMS$''$ has degree of inefficiency $\leq c$ and maximum degree n^τ, then more than half the vertices must be good. Moreover, c_1' is defined as follows:*

$$c_1'(\tau) \stackrel{\text{def}}{=} \sup\left\{c \,\middle|\, \forall\alpha \in \left[-\sqrt{\frac{c}{1+c}}, \sqrt{\frac{c}{1+c}}\right] \lambda_c(\alpha)^2 \cdot \varphi\left(\alpha, \frac{1+c}{\lambda_c(\alpha)^2}\right) > \frac{1+c}{\tau} - \frac{1}{3}\right\}.$$

PROOF. By our choice of c, there is some fixed constant $\rho = \rho(c)$ such that for any α satisfying $\alpha^2 \leq c/(1+c)$ we have

$$\lambda_c(\alpha)^2 \cdot \varphi\left(\alpha, \frac{1+c}{\lambda_c(\alpha)^2}\right) > \frac{1+c}{\tau} - \frac{1}{3} + \rho. \qquad (14)$$

Suppose, first, that $|V_i^{\theta_\alpha}| \leq (N(\varepsilon)n)^{1/\varphi(\alpha,\tau_1(c,\alpha))}$ for every α satisfying $\alpha^2 \leq c/(1+c)$ and every vertex v_i, where $\tau_1(c,\alpha) = (1+c)/\lambda_c(\alpha)^2$. Assume, for the sake of contradiction that KMS$'(G, \{v_i\})$ does not return an independent set of expected size $\Omega(N(\varepsilon)n)$. Then, let $\{w_j\}$, $\{t_j\}$ and $\{\alpha_j\}$ be as in Theorem 7, and let $\hat{t}_j = t_j/\|t_j\|$. It is easy to see that for some α the probability of some $\hat{t}_j$ having projection $\geq \lambda_c(\alpha_j)\sqrt{3}\varepsilon$ is still $\Omega(1/\log n)$ when restricted to j for which $|\alpha_j - \alpha| \leq \frac{1}{\log n}$. This implies that $|V_i^{\theta_\alpha}| N(\sqrt{3}\varepsilon)^{\lambda_c(\alpha)^2 - o(1)} = \Omega(\frac{1}{\log n})$. But, together with the assumed bound on $|V_i^{\theta_\alpha}|$ and the inefficiency of ε, this contradicts (14).

Now, suppose on the contrary that for vertex v_i and some α satisfying $\alpha^2 \leq c/(1+c) + o(1)$ (note that $\alpha > -\frac{1}{3}$ by our choice of c), we have $|V_i^{\theta_\alpha}| > (N(\varepsilon)n)^{1/\varphi(\alpha,\tau_1(c,\alpha))}$. Then the maximum degree in G is bounded by $|V_i^{\theta_\alpha}|^{\tau_2}$, where $\tau_2 = \varphi(\alpha,\tau_1(c,\alpha)) \cdot (\frac{1}{\tau} - \frac{1}{3+3c} - o(1))^{-1}$. Hence, by Corollary 1, KMS-Aux$(G_i^{\theta_\alpha}, \{t_j^i | v_j \in V_i^{\theta_\alpha}\})$ returns an independent set of expected size

$$|V_i^{\theta_\alpha}|^{\varphi(\alpha,\tau_2)} > (N(\varepsilon)n)^{\varphi(\alpha,\tau_2)/\varphi(\alpha,\tau_1(c,\alpha))}.$$

Note that by (14) we have $\tau_1(c,\alpha) < \tau_2$, which in turn implies that $\varphi(\alpha,\tau_1(c,\alpha)) \leq \varphi(\alpha,\tau_2)$, which gives the desired bound above. $\square$

7. CONCLUSIONS

We have shown some improvement in the analysis of vector coloring. We strongly believe that there remains much room for improvement. One starting point is detailed in the conjecture of Appendix B.

8. REFERENCES

[1] S. Arora, S. Rao, and U. Vazirani. Expander flows, geometric embeddings, and graph partitioning. In *Proc. of 36th Annual ACM Symposium on Theory of Computing (STOC)*, pages 222–231, 2004.

[2] A. Blum. New approximations for graph coloring. *Journal of the ACM*, 41(3):470–516, 1994.

[3] A. Blum and D. Karger. An $\tilde{O}(n^{3/14})$-coloring algorithm for 3-colorable graphs. *Information Processing Letters*, 61(1):49–53, 1997.

[4] C. Borell. The Brunn-Minkowski inequality in gauss space. *Invent. Math.*, 30:207–216, 1975.

[5] S. Chawla, R. Krauthgamer, R. Kumar, Y. Rabani, and D. Sivakumar. On the hardness of approximating multicut and sparsest-cut. In *IEEE Conference on Computational Complexity*, pages 144–153, 2005.

[6] I. Dinur, E. Mossel, and O. Regev. Conditional hardness for approximate coloring. In *Proc. of 38th Annual ACM Symposium on Theory of Computing (STOC)*, 2006.

[7] U. Feige, M. Langberg, and G. Schechtman. Graphs with tiny vector chromatic numbers and huge chromatic numbers. *SIAM J. Computing*, 33(6):1338–1368, 2004.

[8] M. X. Goemans and D. P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM*, 42(6):1115–1145, 1995.

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

[10] J. Håstad. Some optimal inapproximability results. *Journal of the ACM*, 48(4):798–859, 2001.

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

[12] H. Karloff and U. Zwick. A 7/8-approximation algorithm for max 3SAT? In *Proc. of 38th IEEE FOCS*, pages 406–415, 1997.

[13] S. Khot, G. Kindler, E. Mossel, and R. O'Donnell. Optimal inapproximability results for max-cut and other 2-variable CSPs? In *Proc. of 45th IEEE FOCS*, pages 146–154, 2004.

[14] S. Khot and N. Vishnoi. Metrics, cuts, and unique games. In *Proc. of 46th IEEE FOCS*, pages 53–62, 2005.

[15] T. Leighton and S. Rao. Multicommodity max-flow min-cut theorems and their use in designing approximation algorithms. *Journal of the ACM*, 46(6):787–832, 1999.

[16] V. Sudakov and B. Tsirel'son. Extremal properties of half-spaces for spherically invariant measures. *Soviet Math*, 9:9–18, 1978. *Translated from Zap. Nauch. Sem. L.O.M.I 41:14–24, 1964*

[17] A. Wigderson. Improving the performance guarantee of approximate graph coloring. *Journal of the ACM*, 30(4):729–735, 1983.

APPENDIX

A. A MODIFIED BLUM KARGER ALGORITHM

In this section we give a summary of the technique of Blum and Karger [3], which relies on the coloring tools of Blum [2]. This is in fact a slight strengthening of the technique, and allows us to present the current result in the same framework as those of [2], [3]. Our main result is the following, whose proof we sketch later:

THEOREM 12. *Let $\mathcal{A}$ be a polynomial time algorithm that takes an n-vertex 3-colorable graph with maximum degree at most Δ as input, and makes progress towards a $f(n,\Delta)$-coloring (where f is monotonically increasing in n and Δ). Then there is a polynomial time algorithm which, for any n-vertex 3-colorable graph, finds an $\tilde{O}(\min_{1\le\Delta\le n}(f(n/2,\Delta)+(n/\Delta)^{3/5}))$ coloring.*

Remark 3. As a corollary of the proof technique of [3], one could show a variant of the above result, where the algorithm $\mathcal{A}$ requires only that the input graph be orientable in such a way that the *maximum out-degree* is at most Δ. Since our requirement of $\mathcal{A}$ is weaker, the above theorem is a slight strengthening of the approach in [3].

This immediately implies the following results of Blum [2] and Blum and Karger [3], respectively.

COROLLARY 2. *For 3-colorable graphs, using the greedy $\Delta+1$-coloring approach, one can find an $\tilde{O}(n^{3/8})$ coloring.*

COROLLARY 3. *For 3-colorable graphs, using the KMS guarantee of progress towards an $\tilde{O}(\Delta^{1/3})$-coloring in Theorem 1, one can find an $\tilde{O}(n^{3/14})$ coloring.*

In order to explain the approach, we use the notion of progress towards a coloring, as defined in [2].

Definition 6. For an n-vertex 3-colorable graph G, and monotonically increasing function $f:\mathbb{N}\to\mathbb{N}$, we define *progress towards a $f(n)$-coloring* as finding any one of the following objects:

Progress Type 1 An independent set of size $\Omega(n/f(n))$.

Progress Type 2 An independent set S having a neighborhood of size $|\cup_{v\in S}\Gamma(v)| = O(|S|\,f(n))$.

Progress Type 3 Two vertices that must have the same color in any legal 3-coloring of G.

We will need the following (slightly simplified) lemma from [2].

LEMMA 7. *Let $f(n)$ be any monotonically increasing function, then for an n-vertex 3-colorable graph G. In order to find an $\tilde{O}(f(n))$-coloring, it suffices to have an algorithm which makes progress towards a $f(n)$-coloring.*

The proof of Theorem 12 relies mainly on the following result, a slightly weaker version of which follows from the techniques outlined in [3]. We omit the proof from the current version.

THEOREM 13. *For any 3-colorable graph $G=(V,E)$, and $\hat{d}=d_{\text{avg}}$ there is a polynomial time algorithm to make progress towards an $\tilde{O}\left(\left(n/\hat{d}\right)^{3/5}\right)$-coloring.*

Using this theorem, we can now give a generalization of [2] and [3]:

PROOF OF THEOREM 12. By Lemma 7, it suffices to show that we can make progress towards the desired coloring. Let Δ_0 be the value of Δ minimizing $f(n,\Delta)+(n/\hat{d})^{3/5}$ (if it is not computable, we can try all values of Δ). If the average degree is at least Δ_0, then by Theorem 13, we can make progress towards an $\tilde{O}((n/\Delta_0)^{3/5})$-coloring. Otherwise, remove one by one vertices of degree $>\Delta_0$ (this eliminates at most $\frac{n}{2}$ vertices), obtaining a subgraph G' with maximum degree at most Δ_0. Apply algorithm $\mathcal{A}$ to G' to make progress towards an $O(f(n/2,\Delta_0))$-coloring. □

B. A NEW GEOMETRIC CONJECTURE AND IMPLICATIONS FOR GRAPH COLORING

We present a geometric conjecture which, if true, would yield the following result when used in the context of the KMS″ analysis:

THEOREM 14. *For 3-colorable graphs, one can find an $O(n^{0.1991})$ coloring in polynomial time.*

Recall Lemma 6, which uses measure concentration to show that an (η,δ) cover grows fast (in measure) as ε shrinks. This lemma can be loosely rephrased as follows: if η is tight for an $(\eta,\frac{1}{2})$ cover $\{v_i\}$ (i.e. $\Pr[\max_i(\zeta,v_i)\le\eta]=\frac{1}{2}$), then with very high probability the maximum projection $\max_i(\zeta,v_i)$ is very close to η. This phenomenon is known as measure concentration.

While measure concentration was used to obtain a tight analysis in [1], we believe that in our case it is suboptimal. Specifically, the theorem of measure concentration states that for any set A of measure $\frac{1}{2}$, the set of radius r around A (that is, $A_r=\{\zeta:\exists z\in A:\|\zeta-z\|\le r\}$) has measure at least $1-N(r)$. However, compare this to the case of an $(\varepsilon,\frac{1}{2})$ cover of mutually orthogonal unit vectors $\{v_i\}$ (this cover is at most $o(1)$-inefficient), and taking $r=\rho\varepsilon$, we see that the set $\{\zeta:\exists i:(\zeta,v_i)\ge(1-\rho)\varepsilon\}$ has measure $1-e^{-\Omega(N(\varepsilon)^{-\rho(2-\rho)})}$, which grows much faster.

Drawing on the intuition that measure is least concentrated around a set which is "closest" to a half-space, we conjecture that the worst configuration is a regular simplex of unit vectors, whose center is far from the origin.

CONJECTURE 1. *For any uniform (ε,δ)-cover $\{u_i\}_{i=1}^k$, if $\{v_i\}_{i=1}^k$ are mutually equidistant unit vectors which also form an (ε,δ)-cover, then for any $\rho\ge 0$*

$$\Pr[\max_i(\zeta,u_i)\le\varepsilon-\rho]\le\Pr[\max_i(\zeta,v_i)\le\varepsilon-\rho]$$

For efficient covers this would give the following corollary:

CONJECTURE 2. *For any constants $c>0$ and $0\le\rho<1$, if $\{u_i\}$ is an $(\varepsilon,\frac{1}{2})$-cover which is at most c-inefficient, then it is also a $((1-\rho)\varepsilon, 1-N(\varepsilon)^{\rho^2(1+1/c)-o(1)})$-cover.*

Using this in place of Lemma 6 would yield the stated improvement.