Improved Performance Guarantees for Bandwidth Minimization Heuristics (draft)

Uriel Feige and Robert Krauthgamer

November 30, 1998

1 Introduction

Given an undirected graph $G = (V, E)$ on n vertices, a *linear arrangement* (also called *ordering*, *numbering* or *layout*) of the graph vertices is a one-to-one mapping of the vertices to the integer line $\tau : V \mapsto \{0, 1, 2, \ldots, n-1\}$. The *bandwidth* of a linear arrangement τ is the dilation of longest edge in the arrangement, i.e. $\max\{|\tau(u) - \tau(v)| : (u, v) \in E\}$.

The problem of minimizing the bandwidth of a graph is to find a linear arrangement for it with the minimum possible bandwidth.

One motivation for the bandwidth problem is its equivalence to minimizing the bandwidth of a sparse symmetric square matrix M. Performance of sparse matrix manipulations such as Gaussian elimination may be improved when all the nonzero entries of the matrix lie in a narrow band around the main diagonal, so it is desirable to reorder the rows and columns of the matrix so as to minimize the width of this band. When the (symmetric) nonzero elements of the matrix are viewed as vertex adjacencies, then the reordering problem is exactly the minimum bandwidth problem of this graph.

The bandwidth problem is known to be NP-hard [Pap76], even for certain families of trees [GGJK78, Mon86].

Recent results bound the approximability of the bandwidth problem. Feige [Fei97] gives a nearly linear time randomized algorithm that finds linear arrangement with bandwidth within ratio of $O(\log^{9/2} n)$ of the optimum. Blache, Karpinski and Wirtgen [BKW98] show that the bandwidth problem on trees cannot be approximated within ratio of $4/3$, unless $P = NP$. For general graphs they show an inapproximability ratio of $3/2$.

A known lower bound on the bandwidth is the *local density* bound. Let $\Gamma_j(v)$ be the set of vertices at distance at most j from v in the graph. Then the bandwidth of the graph is at least $D = \max_{j,v}\{(|\Gamma_j(v)|-1)/2j\}$. A gap of $\Omega(\log n)$ between the local density and the bandwidth can be demonstrated on trees [Chv80] and on expander graphs. On the other hand, Feige [Fei97] shows thru his algorithm that this gap is $O(\log^{9/2} n)$.

In practice, heuristics for minimizing the bandwidth appear to work rather well (see, for example, the survey [CCDG82]). One of the most succesful heuristic algorithms was introduced by Cuthill and McKee [CM69] in 1969. Their heuristic is based on a level structure of the graph, which is a partitioning the vertex set V into sets $L_0, L_1, L_2, \ldots, L_q$ (called *levels*) in which two endpoints of every edge are either in the same level or in two consecutive levels. Of particular interest is the *level structure rooted at vertex* u, which can be efficiently obtained by a Breadth First Search from u. This level structure is defined as $L_i = \{v \in V : d(u,v) = i\}$, where $d(u,v)$ is the distance between u and v in the graph G.

Arranging the vertices level by level (with an arbitrary ordering within each level), gives bandwidth which is governed by the size of the largest level, $L_{max} = \max_i |L_i|$. It is easy to see that the resulting bandwidth is bounded by L_{max} from below and by $2L_{max} - 1$ from above. Several variations of this idea appear in the literature, and are usually referred to as *level algorithms* (see [CCDG82] and the references therein). Some justification to the practical success of level algorithms is given by Turner [Tur86] using average case analysis, and we follow his framework.

1.1 Average Case Analysis

Since the bandwidth problem is NP-hard to solve in the worst case, one may consider analyzing the behavior of an algorithm in the average case. A common approach is the *random model*, in which some probability distribution on the input instances is assumed in order to prove that the algorithm succeeds *almost surely*, i.e. the probability that the algorithm succeeds (taken over the distribution of the inputs) approaches 1 as $n \to \infty$. See, for example [Kuc77, Bop87, Tur88, DF89, FS92, AK97].

In graph problems, the assumed input distribution is usually related to natural random graph models adjusted to the particular problem studied. Many of these distributions are derived from $G_{n,p}$, the random graph on n vertices, where each pair of vertices is joined by an edge with probability p (called the edge probability), independently of all other events.

Semi-Random Models

As a means to smoothen transition between worst case and average case analysis, Blum and Spencer [BS95] suggested to analyze an algorithm in a *semi-random model*, where the inputs distribution is partially controlled by an adversary. They studied graph k-coloring and suggested several semi-random variants. In one of their variants, an adversary chooses for each pair of vertices whether he wants to join them by an edge or not, and its choice is flipped with probability p.

Feige and Kilian [FK97] study the *monotone adversary model*, which is as strong as the strongest models of Blum and Spencer [BS95]. In the monotone adversary model, the graph is first drawn randomly (as in the random model) and then an adversary is allowed to alter the random decisions in a monotonic fashion (e.g. allowed to add edges but not to remove edges).

The Semi-Random Model for the Bandwidth Problem

Turner [Tur86] analyzed a random model denoted $G_{n,B,p}$for the bandwidth problem. In this model, an n vertices graph with bandwidth at most $B = B(n)$ and a probability $p = p(n)$ for edges is randomly generated as follows. Assume the graph vertices are labeled $\{0, 1, \ldots, n-1\}$, and call a pair of vertices u, v *potential neighbors* if $|u-v| \leq B$. Each pair of potential neighbors is joined by an edge with probability p, independently of all other pairs. The graph is then given without the vertices labels (or equivalently, the vertices are permuted arbitrarily).

The random model can be easily extended to a monotone adversary model. First, a graph is drawn randomly from $G_{n,B,p}$. After the random edges are generated, an adversary may only add edges between pairs of potential neighbors. As before, the graph is given without the vertices labels. Note that the resulting graph has bandwidth $\leq B$.

1.2 Previous Results

Turner [Tur86] studied the performance of level algorithms on the random model $G_{n,B,p}$ with fixed edge probability p and $B >> \ln n$. He shows that the level structure rooted at 0 and the optimal bandwidth of the graph are close to each other, almost surely. (one may assume that vertex 0 is known to the algorithm, e.g. by exhaustively trying all possible vertices, although Turner [Tur86] analyzes the average case behavior of several faster heuristics.) He concludes that the following properties hold almost surely for a graph drawn from $G_{n,B,p}$.

1. The bandwidth of the graph is $\geq B - O(\log n)$.
2. Any level algorithm starting at vertex 0 outputs an arrangement with bandwidth of ratio 2 of the optimum.
3. If $B < n/4$ then a certain variant outputs an arrangement with bandwidth at most $B + O(\log n)$.

1.3 Our Results

We analyze the (monotone adversary) semi-random model $G_{n,B,p}$. The first part of our work presents an algorithm which extends the analysis of Turner [Tur86] to smaller edge probabilities p. In the second part we use the local density bound to improve Turner's analysis [Tur86] of the probable bandwidth of the graph drawn from $G_{n,B,p}$.

The Algorithm

Turner [Tur86] gives an algorithm to compute an almost optimal linear arrangement for fixed edge probability p. The core of our algorithm and analysis are extensions of his results, but we consider a wider range of graphs by allowing lower edge densities and less regular structure. We require that the expected number of neighbors is (at least) logarithmic (in n), and consider $p = o(1)$ accordingly (instead of constant p). Allowing a monotone adversary significantly relaxes the regular structure of the input graphs. We remark that the analysis of Turner [Tur86] also extends to the monotone adversary model, but this is more important when the edge probability is smaller. Theorem 1 shows that our algorithm almost surely outputs a linear arrangement of bandwidth at most $B(1+\epsilon)$ on a semi-random input graph.

Section 2 describes our algorithm, which combines the linear arrangements of two level structures originating at opposite ends of the graph. In each of the two level structures the ordering within the levels is based on a certain neighborhood property, generalizing the algorithm of Turner [Tur86]. The idea of using two level structures originating at opposite ends appears in a variant suggested by Gibbs, Poole and Stockmeyer [GPS76], which show experimentally that this method improves over level algorithms (see the survey [CCDG82]).

Lower Bound on the Optimal Bandwidth

It is conjectured in [Tur86] that the bandwidth of a graph drawn randomly from $G_{n,B,p}$ is exactly B (almost surely), for fixed edge probability p and when $\ln n << B << n$. He showed a lower bound of $B - O(\log n)$ (almost surely) on the bandwidth by analyzing the behaviour of B consecutive vertices (e.g. $1, 2, \ldots, B$). An upper bound of B is immediate.

Section 3 presents an improved lower bound through average case analysis of the local density bound. Essentially, the bound follows from the version corresponding to all the graph vertices, i.e. $(n-1)/diameter(G)$. Theorem 2 showes that for $B << \sqrt{n/\ln n}$ the bandwidth of a random graph is almost surely at least $B - \lceil (1+\varepsilon)/p \rceil$, where $\varepsilon > 0$ is an arbitrary constant. For fixed p this turns out to be a lower bound of $B - O(1)$.

1.4 Notation

All probability measures refer solely to the probability distribution on the input graphs $G_{n,B,p}$, since the algorithms involved are deterministic. The probability for an edge to exist is $p = c \ln n/B$, where c is a large enough constant. The term *almost surely* will denote probability of $1 - o(1)$. The term *with high probability* will denote probability $> 1 - n^{-c'}$ for some constant c', which can be made arbitrary large, depending on c. The relation between c and c' will usually be straightforward and we will not insist on specifying it exactly. Note that when each of the events $A_1, \ldots, A_n$ happens with high probability, also $A_1 \cap \ldots \cap A_n$ does, by the union bound on the complementary events.

The notion of distance between vertices $v, w \in V = \{0, 1, 2, \ldots, n-1\}$ appears here in two different contexts. To avoid confusion, when $|u - w| = t$ we say that the *difference* between u and w is t or that they are t *apart* of each other.

The *distance in the graph* G between v and w is the length of the shortest path in the graph between v and w, denoted $d(v, w)$. Note that the distance in the graph coincides with the *level* of w from v, denoted $level_v(w)$. Clearly, the level of v from w is the same as that of w from v.

Let the j-*neighborhood* of v, denoted $\Gamma_j(v)$, be the set of vertices at distance at most j from v in the graph, so $\Gamma_j(v) = \{w : d(v, w) \leq j\} = \{w : level_v(w) \leq j\}$.

Throughout the paper, we use Chernoff bounds to argue that the sum X of Bernouli random variables is concentrated around its expectation EX

(e.g. a binomially distributed random variable such as the degree of a vertex in $G_{n,B,p}$). The probability for large deviation from the expectation is bounded by $Pr[X < (1-\delta)EX] < \exp(-\delta^2 EX/2)$, for any $\delta > 0$ (see for example [MR95, Chapter 4] and the references therein). We mostly use it for expectations $EX = \Omega(\ln n)$.

2 A Heuristic and Its Analysis

2.1 Overview of Algorithm and Its Analysis

Section 2 presents an algorithm for the random model $G_{n,B,p}$ allowing the edge probability to decrease as a function of n. We require $p \geq c \ln n/B$, where c is a large enough constant, so the expected number of neighbors of a vertex is $2Bp \geq 2c \ln n$. Assume that $\ln^2 n \leq B \leq n/\ln n$, and let $\varepsilon > 0$ be some fixed constant, such that $c \geq c_\varepsilon$. Theorem 1 states that the linear arrangement computed by the algorithm has bandwidth which is within ratio of $1+\varepsilon$ from the optimum, almost surely. Our analysis, as well as Turner's [Tur86], extend to the monotone adversary model. @@@

We now sketch the algorithm and the related average case analysis of $G_{n,B,p}$. The goal of the analysis is to show that the original labels of the vertices can be approximated almost surely using efficiently computable combinatorial properties.

Section 2.2 discusses the probable level structure rooted at vertex 0 of a random graph from $G_{n,B,p}$. This relates to an arbitrary level algorithm. We investigate which vertices tend to be in the j-th level, in terms of their original labels. For example, would the j-th level tend to consist of the vertices whose original label was between $(j-1)B$ and jB ?

We frequently refer to the original labels of the vertices, so let us define some terminology for it. Let $u, w \in V = \{0, 1, 2, \ldots, n-1\}$ be two vertices (i.e. according to their original labels). We say that u is *left* of w if $u < w$, and *right* of w if $u > w$. The *interval* $[u, w]$ is the set of vertices $\{x : v \leq x \leq w\}$, and we similarly define the half open interval $(u, w]$, etc.

Using these notions, Section 2.2 shows that level j is almost surely roughly an interval of size $(1+\varepsilon)B$, when j is large enough (i.e. far enough from 0). It is roughly an interval in the sense that some of the vertices in ε subintervals around the endpoints of the interval, might not belong to the j-th level. The exceptional vertices near the interval's ndpoints belong to the successive levels ($j-1$ and $j+1$, respectively). Specifically, level j is contained in a $(1+\varepsilon)B$ interval, and contains a $(1-\varepsilon)B$ interval, as formalized in Lemma 6. Note that the exact endpoints of the interval are not specified.

There is a particular level for which we provide a formulation of the interval as $[(j-1)B, jB]$. This is given in Corollary **??** for level $j = 2\gamma+1$, where $\gamma = \Theta(\ln B)$ appears naturally in the analysis. The particular result regarding $2\gamma+1$ is used for proving that all levels beyond it, are close to an interval.

Section 2.3 refines the general level algorithm by a particular ordering inside each level, based on a combinatorial property which tends to approximate the order of the original labels. We count the number of *markers* for each vertex v, which are (certain) vertices that lie on a shortest path between 0 and v. The complexity of computing the set of markers of a vertex v is dominated by the complexity of computing a BFS upto a limited number of levels. Corollary 8 shows that with this ordering all vertices far enough from 0 do not deviate more than εB from their original label.

Section 2.4 gives the final algorithm whose performance ratio is, almost surely, $1+\varepsilon$. Assume that vertices 0 and $n-1$ are given. Then two orderings can be obtained using the algorithm of Section 2.3. In the first one we start from 0, and in the second one we start from vertex $n-1$. Note that no vertex deviates much of its original label in at least one of the two orderings.

The two orderings are cut and pasted into one ordering, in which all vertices are placed close to their original label, almost surely. The bandwidth of this ordering is, almost surely, at most $(1+\varepsilon)B$. It is straightforward that the same arguments apply not only to the pair of vertices 0 and $n-1$, but to any pair of vertices v_1, v_2 which are in the top and bottom $n/4$ of the vertices, respectively. It is not hard to see that in this model one can get away without checking n^2 pairs. Run a BFS from an arbitrary vertex, and let v_1 be a vertex from the last level. Apply another BFS from v_1, and take a vertex v_2 from the last level. Almost surely the pair v_1, v_2 is suitable to serve as a basis for the algorithm (instead of 0 and $n-1$).

2.2 The Roughly Intervals Behaviour

As a first step we relate to arbitrary level algorithm and ignore the ordering within each level. A probable correspondance between the level of vertices and their labels would have been useful for designing an algorithm which efficiently generates a linear arrangement which is close to the original one.

For simplicity, assume that vertex 0 is given in advance, and the algorithm computes the level structure rooted at 0. Clearly, all vertices of the first level have labels within the interval $[1, B]$. It is expected that pB vertices of the interval $[1, B]$ will appear in the first level. Since $pB = c \ln n << B$, the information retrieved from the first level is only partial. Hence, we will have to settle for a weaker correspondance between the levels and the original labels.

We use expansion properties that are likely to occur in random graphs, to show that starting with a far enough level (which is $\Theta(\ln B)$ basically), with high probability all levels are "close" to intervals of length $\approx B$. We say "close" to an interval to say that each such level is an interval except maybe some vertices near the interval endpoints. In these areas there is a "transition" from the interval of one level to the interval of the next level.

We state the results with respect to the level structure rooted at 0, although the same arguments hold for any other vertex in any direction (right or left). We therefore use Γ_j as a shorthand for $\Gamma_j(0)$.

We first establish the existence of some $\lambda = \Theta(\ln B)$, such that with high probability the set of vertices at level λ from 0 is the interval $[(\lambda - 1 - \varepsilon)B, \lambda B]$, up to some of the εB vertices at the ends of the interval. We then use this result to conclude about the probable structure of the levels beyond λ.

We consider a series of particular disjoint intervals I_j, where interval I_{j+1} is to the right of I_j, with a certain gap between them (so the intervals are not consecutive). We would like each interval I_j to contain many vertices of level j (from 0), so the gaps are chosen such that each of the vertices of I_j is a potential neighbor of each of the vertices of I_{j+1}. Specifically, we choose gaps as large as possible while keeping every pair of vertices from I_j and I_{j+1} to be at most B apart from each other.

We think of $I_0 = \{0\}$, and need the first interval to be of length εB, so let $I_1 = [(1-\varepsilon)B, B]$. The rest of the intervals I_j ($j \geq 2$) are of length $B/\ln n$. We locate I_j such that its rightmost vertex is exactly B to the right of the leftmost vertex of I_{j-1}. It follows that the difference between the leftmost vertices of I_{j-1} and I_j is $(1 - \frac{1}{\ln n})B$.

Denote by L_j the vertices of interval I_j which are at level $\leq j$ (from vertex 0), so $L_j = I_j \cap \Gamma_j$. Later on we will show that for the small values of j which we consider, the vertices in L_j are exactly at level j and not less.

The following two lemmas show that with high probability the size of L_j grows rapidly, using arguments of expansion in the relevant random subgraph. We chose the size of the first interval I_1 such that the size of L_1 is with high probability logarithmic (in n). We then use this to show that the size of successive L_j's increases by a constant factor, until for some value γ, L_γ occupies at least 1/4 of the interval I_γ.

Lemma 1 *For some constant c_1 (depending on c and ε), $|L_1| \geq c_1 \ln n$ with high probability.*

Proof. Since all of I_1 are potential neighbors of 0, $|L_1|$ is random variable with the binomial distribution and expectation $E[|L_1|] = p \cdot |I_1| = \frac{c \ln n}{B} \cdot \varepsilon B =$

$\varepsilon c \ln n$. For large enough constant c (depending on ε) we get by the Chernoff bound that with high probability L_1 contains at least $c_1 \ln n$ vertices. $\square$

Lemma 2 *Assume $j \geq 2$ and $\ln n \leq |L_{j-1}| \leq B/c \ln n$. Then with high probability $|L_j| \geq \frac{1}{4}c|L_{j-1}|$.*

Proof. Consider a vertex $w \in I_j$. If w is connected by an edge to some vertex in L_{j-1}, then it is surely in L_j. All vertices in L_{j-1} are potential neighbors of w, so by the inclusion-exclusion bound: $Pr[w \in L_j] \geq |L_{j-1}|p - \frac{1}{2}|L_{j-1}|^2p^2$. Since $|L_{j-1}| \leq B/c \ln n = 1/p$, we get that $Pr[w \in L_j] \geq \frac{1}{2}|L_{j-1}|p$.

Hence, $E|L_j| \geq |I_j| \cdot \frac{1}{2}|L_{j-1}|p = \frac{1}{2}c|L_{j-1}|$. By our assumption that $|L_{j-1}| \geq \ln n$, the expectation is high enough to apply the Chernoff bound, and get that with high probability $|L_j| \geq \frac{1}{4}c|L_{j-1}|$. $\square$

Lemma 3 *For $\gamma \approx \frac{\ln(B/\ln^2 n)}{\ln c'}$ such that with high probability $|L_\gamma| \geq B/4 \ln n$.*

Proof. We can start with Lemma 1 and then apply iteratively Lemma 2 to get constant factor expansions of L_j until we get $|L_j| \geq B/4 \ln n$ (the number of iterations is $j \leq n$ so it still holds with high probability). When $|L_j|$ becomes too large, we can use only a subset of it of size $B/c \ln n$. In any case, we can fix γ to be a large enough j which suffices for $|L_\gamma| \geq B/4 \ln n$, so we require

$$|L_\gamma| \overset{whp}{\geq} |L_1|(c/4)^{\gamma-1} \overset{whp}{\geq} c_1 \ln n (c/4)^{\gamma-1} \geq B/4 \ln n$$

$$\gamma = \frac{\ln(B/4c_1 \ln^2 n)}{\ln(c/4)} + \Theta(1) = \frac{\ln B - 2\ln\ln n}{\ln(c/4)} + \Theta(1)$$

$\square$

So for $\gamma \approx \frac{\ln(B/\ln^2 n)}{\ln c'}$, with high probability at least 1/4 of the vertices of the interval I_γ are at level at most γ from 0. At this stage the constant factor expansion does not hold any more, so in order to increase the density of vertices at the "right" level inside our interval we use symmetry considerations. In addition to having $I_1, \ldots, I_\gamma$ starting at 0, we construct similarly $I'_1, \ldots, I'_\gamma$ starting from vertex v towards 0 (see Figure 1). When the vertices of I_γ and I'_γ are potential neighbors (i.e. at most B apart), with high probability there will be a path of length $\leq 2\gamma + 1$ between 0 and v.

Lemma 4 *For any vertex $v \leq (2\gamma+1-\varepsilon)B$, with high probability level$(v) \leq 2\gamma + 1$.*

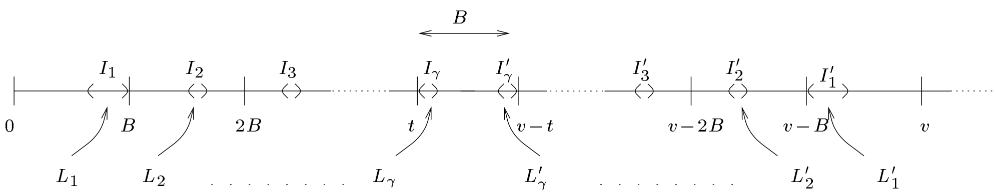

Figure 1: Symmetric Construction of L_γ and L'_γ

Proof. Assume first $v = (2\gamma + 1 - \varepsilon)B$, and let $\varepsilon = 4\varepsilon'$. Let us denote by t the leftmost vertex in I_γ, and so the rightmost vertex in I'_γ is $v-t$. Since the leftmost vertex of I_1 is $(1-\varepsilon)B$, and the difference between leftmost vertices of succesive levels is $(1-\frac{1}{\ln n})B$, we get that $t = (1-\varepsilon')B+(\gamma-1)(1-\frac{1}{\ln n})B = (\gamma - \varepsilon' - \frac{\gamma-1}{\ln n})B$. Since $\gamma \leq \frac{\ln n}{\ln(c/4)}$, for large enough c we have that $\frac{\gamma}{\ln n} \leq \varepsilon$, and then $t \geq (\gamma - 2\varepsilon')B$. The difference between the leftmost vertex in I_γ and the rightmost vertex in I'_γ is $(v-t)-t \leq (2\gamma+1-4\varepsilon')B - 2(\gamma-2\varepsilon')B = B$. Hence, each vertex of I_γ is a potential neighbor of each vertex of I'_γ.

Construct also $L'_1, \ldots, L'_\gamma$ in a similar way to $L_1, \ldots, L_\gamma$, so that L'_j is the set of vertices in the interval I'_j which are at level $\leq j$ from v. By Lemma 3 both L_γ, L'_γ are very dense in their intervals I_γ, I'_γ, respectively, so with high probability there is at least one edge between them:

$$Pr \begin{bmatrix} \text{no edge between} \\ L_\gamma \text{ and } L'_\gamma \end{bmatrix} < (1-p)^{(B/4\ln n)^2} \leq e^{-cB/16\ln n} \leq n^{-c'}$$

(recall we assumed that $B \geq \ln^2 n$). It follows that with high probability there is a path between 0 and v of length $2\gamma + 1$.

When $v < (2\gamma + 1 - \varepsilon)B$ we can construct the intervals $I_1, \ldots, I_\gamma$ and $I'_1, \ldots, I'_\gamma$ accordingly (i.e. closer to each other, but still disjoint), and with high probability they contain a path between 0 and v of length $\leq 2\gamma + 1$. Hence, with high probability $level(v) \leq 2\gamma + 1$, as claimed. □

Lemma 5 *With high probability all vertices of the interval* $[2\gamma B + 1, (2\gamma + 1 - \varepsilon)B]$ *are at level* $2\gamma + 1$ *from 0.*

Proof. On the one hand, for every vertex v in the interval, $v > 2\gamma B$. Every edge connects vertices whose difference is at most B, at least $2\gamma + 1$ edges are needed for a path from 0 to v, so $level(v) \geq 2\gamma + 1$.

On the other hand, by Lemma 4 with high probability each vertex in the interval is at level $\leq 2\gamma+1$ from 0. Hence, with high probability all the interval vertices are at level $2\gamma+1$. $\square$

The choice of 0 as the starting vertex was arbitrary, and the same results hold for any other vertex (applied separately to the right and left directions). Moreover, with high probability the conclusion of Lemma 5 holds for all the vertices simultaneously. This allows us to extend the characterization of the level of vertices to all vertices beyond $2\gamma B$. Our goal would be to show that with high probability the set of vertices at any level beyond $2\gamma+1$ is roughly an interval of size $\approx B$.

Let r_j be the rightmost vertex whose level from 0 is j (define $r_0 = 0$). Then r_j is at most B to the right of r_{j-1}. With high probability r_j is at least $(1-\varepsilon)B$ higher than r_{j-1}, because the probability that r_{j-1} itself will not have any edge which advances at least than $(1-\varepsilon)B$ is $(1-p)^{\varepsilon B} \leq e^{-\varepsilon p B} = n^{-c'}$.

Lemma 6 *With high probability for all $i > 2\gamma+1$ the i-th level is roughly an interval of length $\approx B$, i.e. there exists some r such that $[r+\varepsilon B, r+B] \subseteq \Gamma_i \setminus \Gamma_{i-1} \subseteq [r, r+(1+\varepsilon)B]$*

Proof. Let $j = i - 2\gamma - 1$, and recall that r_j is the rightmost vertex at level j from 0. We show that $r = r_j + 2\gamma B + 1$ satisfies the claim for level $i = j + 2\gamma + 1$, (see Figure 2).

From Lemma 4 with r_j as the starting vertex, with high probability all vertices in $[r_j, r_j + (2\gamma+1-\varepsilon)B]$ are at level $\leq 2\gamma+1$ from r_j, and in particular, at level $\leq j+2\gamma+1$ from 0. Since r_j is the rightmost vertex at level j, any vertex which is to the right of $r_j + 2\gamma B$ is at level $\geq j+2\gamma+1$ from 0. Therefore, with high probability all vertices in $[r_j + 2\gamma B + 1, r_j + (2\gamma+1+\varepsilon)B]$ are exactly at level $j+2\gamma+1 = i$, which proves the first part, i.e. that the i-th level contains an interval of size $(1-\varepsilon)B$.

Apply the argument of Lemma 4 for each of $r_0, r_1, \ldots, r_{j-1}$. The difference between successive applications is at most B, so together they cover all the vertices between 0 and $r_{j-1} + (2\gamma+1-\varepsilon)B$. The difference between r_j and r_{j-1} is at most B, so the covered area includes all vertices between 0 and $r_j + (2\gamma-\varepsilon)B$, which are thus with high probability at level $\leq (j-1)+2\gamma+1 < i$ from 0. On the other hand, any vertex to the right of $r_j + (2\gamma+1)B$ is at level $\geq j+2\gamma+2 > i$. Hence, with high probability all vertices in the i-th level are in the interval $[r_j + (2\gamma-\varepsilon)B, r_j + (2\gamma+1)B]$, satsifying the second part. $\square$

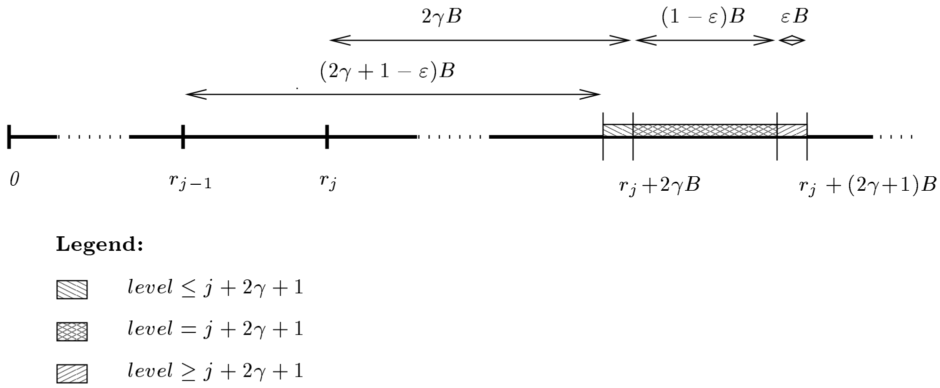

Figure 2: Exact level

In fact, Lemma 6 shows that with high probability each level $i \geq 2\gamma + 1$ is contained in an interval of size $(1+\varepsilon)B$ and contain an interval of size $(1-\varepsilon)B$. Let us denote by l_i the leftmost vertex of level i (and r_i is the rightmost vertex, as before), then $[l_i, r_i]$ is an upper bound on the set of vertices of level i, and with high probability $|r_i - l_i| < (1+\varepsilon)B$. Moreover, with high probability there is some interval $[l'_i, r'_i]$ which consists solely of vertices at level i, and its length is $\geq (1-\varepsilon)B$. The resulting level decomposition of the graph and the transition between the levels is shown in Figure 3. Note that the transition area between levels i and $i+1$ is a mixture of vertices of levels i and $i+1$. Moreover, with high probability $l_{i+1} = r'_i + 1$ and $l'_{i+1} = r_i + 1$.

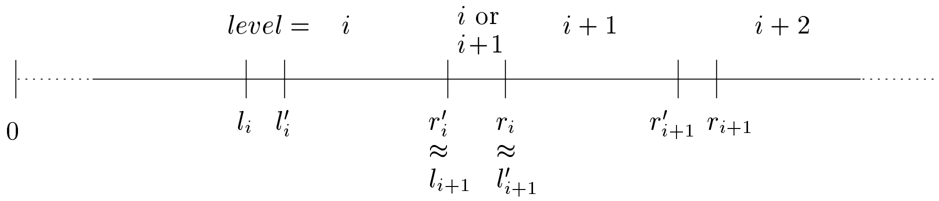

Figure 3: Transition between levels

2.3 Ordering within a Level

Following Turner [Tur86], with high probability the "correct" location of each vertex can be approximated by counting neighbors.

Definition 1 *Let x be a vertex at level $i > 2\gamma + 1$ (from 0). A* marker *of vertex x is any vertex v which is at level $i - 2\gamma - 1$ from 0 and at level $2\gamma + 1$ from x (i.e. $level_0(v) = i - 2\gamma - 1$ and $level_x(v) = 2\gamma + 1$).*

Note that v is a marker of x if and only if v is the vertex on the $(2\gamma + 1)$ place on some shortest path from v to 0.

We determine the order of vertices within a level by counting the number of markers they have, which can be computed efficiently. We show that if two vertices of the same level are at least εB apart, then with high probability the markers counting criterion will arrange them in the correct order.

Lemma 7 *Let x, y be a pair of vertices which are at least εB apart from each other ($x + \varepsilon B < y$). If x, y are both at the same level which is $\geq 5\gamma$, then with high probability x has more markers then y.*

Proof. Fix $\varepsilon = 4\varepsilon'$, so $x + 4\varepsilon' B < y$. Assume x and y are both at level $j + 2\gamma + 1$, so $j > 2\gamma + 1$. Markers for x and y must be at level j from 0, so denote by $[l_j, r_j]$ the shortest interval which bounds from above the vertices of level j from 0, and let $[l'_j, r'_j]$ be some interval which consists solely of vertices at level j from 0. Clearly $[l'_j, r'_j] \subset [l_j, r_j]$ as in Figure 4. Corollary **??** guarantees that with high probability there are some l'_j, r'_j such that $|l_j - l'_j| < \varepsilon' B$ and $|r_j - r'_j| < \varepsilon' B$, and also $|r'_j - l'_j| > (1\varepsilon')B$.

Similarly, we can bound the set of vertices of level $2\gamma + 1$ from y from above by the interval $[l_y, r_y]$ and from below by $[l'_y, r'_y]$. Clearly, $l_y < l'_y < r'_y < r_y$ (see Figure 4. Suppose $l_x < l'_x < r'_x < r_x$ be similar bounds on the vertices at level $2\gamma + 1$ from x.

The main argument of the proof is as follows. Any marker for y must be in the intersection of $[l_y, r_y]$ and $[l_j, r_j]$, and this gives us an *upper* bound on the number of markers for y. The number of markers for x can be *lower* bounded by the intersection of $[l'_x, r'_x]$ and $[l'_j, r'_j]$. Using the results from the previous section, we can derive some of this bounds more explicitly and compare the number of markers as follows.

We count the number of markers separately on three subintervals, denoted by 1,2 and 3 in Figure 4. In the first subinterval, with high probability every vertex which is a marker for y is also a marker for x, so we may ignore it in the count. The second subinterval contains vertices might be markers for y and not for x, but with high probability the size of this subinterval is at most $\varepsilon' B$. The third subinterval contains, with high probability, at least $2\varepsilon' B$ more markers for y than for x. Altogether, with high probability x has more markers than y.

In order to bound the size of these intervals, we establish some connections between them. For simplicity, we use r_{j-1} (the rightmost vertex at level $j-1$ from 0) instead of l'_j, (but note that with high probability $r_{j-1} = l'_j - 1$). See Figure 4 for illustration.

(a). We may assume that $[l'_x, r'_x]$ is the interval guaranteed from Lemma 5, i.e. $[l'_x, r'_x] = [x - 2t - B, x - 2t - \varepsilon' B]$. So with high probability any vertex of $[l'_x, r'_x]$ which is a marker for y is also a marker of x.

(b). l_y is at most $(2\gamma + 1)B \le 2t + (1 + \varepsilon')B$ to the left of y, by Lemma **??**. Combined with (a) we know that with high probability l'_x and l_y are at least $3\varepsilon' B$ apart.

(c). With high probability $r_{j-1} < l'_x$ or otherwise there would be a path from 0 through r_{j-1} to x, of length $\le (j-1) + (2\gamma + 1) < level_0(x)$.

(d). Note that $[r_{j-1}, r_j]$ is of length $\le B$ and $[l'_x, r'_x]$ is of length $(1 - \varepsilon')B$. By (c) with high probability r_{j-1} is to the left of l'_x and thus r_j cannot be more than $\varepsilon' B$ to the right of r'_x.

(e). r_j cannot be to the left of l_y or otherwise there is no path from 0 to y of length $j + 2\gamma + 1$. Since $|r_j - r'_j| < \varepsilon' B$ with high probability r'_j is at most $\varepsilon' B$ to the left of r_j and thus also at most $\varepsilon' B$ to the left of y_1.

We now turn to the count of markers. Markers for y cannot be to the left of l_y. The first interval is $[l_y, r'_x]$. From (a) we know that with high probability the vertices of this interval cannot be markers for y without being also markers for x, so we may ignore this interval in the count.

The second interval consists of the markers of y which are to the right of r'_x (if there are any such vertices). These vertices must be at level j and are thus bounded from the right by r_j. From (d) we know that there are at most $\varepsilon' B$ such vertices.

The rest of the vertices at level j are to the left of l_y, and thus are not markers for y. In particular, consider the interval between l'_x and l_y, which by (b) contains with high probability at least $3\varepsilon' B$ vertices. All the vertices

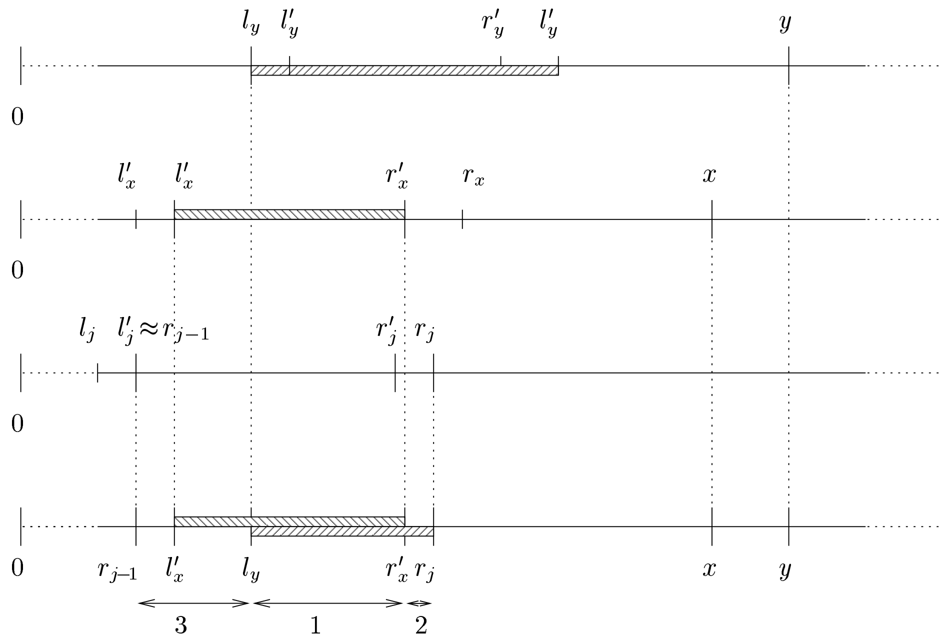

Figure 4: Ordering x, y by counting markers

of this interval are at level j, except at most $\varepsilon' B$ vertices by (e). Hence, it contains no markers for y and at least at least $2\varepsilon' B$ markers for x.

Overall, with high probability x will have at least $\varepsilon' B$ grandparents more than y, as claimed. $\square$

We call the variant of the level algorithm with counting markers the *markers level algorithm*, and denote by τ_s be the ordering output by the markers level algorithm started from vertex s. Lemma 7 states that with high probability τ_0 provides a linear arrangement in which all vertices beyond the first $5\gamma = O(\ln B)$ levels are at most εB far from their "correct" placement. For $B \leq n/\ln n$ this includes all vertices above $n/3$.

Corollary 8 *Let $B \leq n/\ln n$. Then with high probability all vertices beyond $n/3$ are placed by the markers level algorithm with starting vertex 0, at most εB far from their "original" location, i.e. $|v - \tau_0(v)| \leq \varepsilon B$.*

In fact, the same argument will work for every starting vertex s which is not too far from 0, say $s \in [0, n/10]$.

2.4 Cut and Paste

To obtain a linear arrangement which is good on all vertices, we apply the markers level algorithm from each of the endpoints 0 and $n-1$. This provides us with τ_0 and τ_{n-1}. We then compute the reverse ordering of τ_{n-1}, denoted by τ^R_{n-1}. The analysis of the previous section show that with high probability the upper 2/3 of τ_0 and the lower 2/3 of τ^R_{n-1} are close to the original ordering ("good").

The plan is to cut each arrangement about its middle and paste the two "good" halves into one complete linear arrangement, which will be the final output. We assume that the vertices 0 and $n-1$ are known to us in advance but the same argument holds for any pair of vertices which are close enough to the endpoints (say at most $n/10$), since we only use the upper 2/3 of each arrangement, which is "good" with high probability.

The computation of the final output τ_F is as follows. First copy the upper half of τ_0 to τ_F (i.e. on $[n/2, n-1]$ we have $\tau_F \equiv \tau_0$). The remaining $n/2$ vertices are then ordered by τ_F according to their order in τ^R_{n-1} (see Figure 5). We call this algorithm the *pasted level algorithm.*

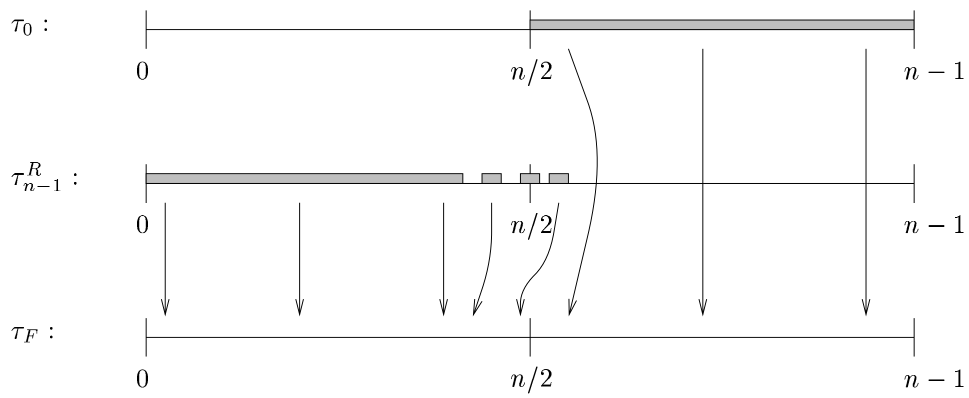

Figure 5: Pasting two linear arrangements

Lemma 9 *With high probability τ_F places all vertices at most εB from their original location.*

Proof. In a sense, the problem might be only in the "glued" area, which we can show to be small.

Fix $\varepsilon = 3\varepsilon'$. The location of vertices in the upper half of τ_F is determined by τ_0, and their proximity to the original location is guaranteed with high probability by Corollary 8.

Consider now the vertices which are placed by τ^R_{n-1} below $n/2 - 2\varepsilon' B$. With high probability none of them is placed by τ_0 in the upper half, or otherwise its original location is more than $\varepsilon' B$ far from its placement by one of τ_0 and τ^R_{n-1}. Hence, the vertices which are placed by τ^R_{n-1} below $n/2 - 2\varepsilon' B$ survive the first phase, and their location is determined by τ^R_{n-1}. Since they all remain for hte second phase, their location in τ_F is exactly their location in τ^R_{n-1}.

The remaining vertices are located by τ_F into an area of $2\varepsilon' B$ just below the glue point $n/2$. With high probability their original location is at most $n/2 + \varepsilon' B$ (or otherwise they would have been placed by τ_0 in the upper half) and at least $n/2 - 3\varepsilon' B$ (or otherwise they would have been placed by τ^R_{n-1} below $n/2 - 2\varepsilon' B$). Hence, they are at most $3\varepsilon' B$ far from their original location.

We conclude that with high probability all vertices are at most $3\varepsilon' B$ far from their original location, as claimed. □

Corollary 10 *For fixed $\varepsilon > 0$ with high probability the full algorithm outputs an arrangement τ_F which has bandwidth at most $(1 + \varepsilon)B$.*

Proof. Fix $\varepsilon' = 2\varepsilon'$. With high probability no vertex is misplaced by τ_F more than $\varepsilon' B$ from its original location, and thus all edges will be at most $2\varepsilon' B$ longer. Since all edges were originally at most B long, their length in τ_F is at most $(1 + 2\varepsilon')B$. □

Theorem 1 *Let $\varepsilon > 0$ be fixed. Then for $\ln^2 n \leq B < n/\ln n$ and $p \geq c \ln n/B$ for large enough constant $c > c_\varepsilon$. Then with high probability over the distribution of graphs $G_{n,B,p}$, the pasted level algorithm has performance ratio $1 + \varepsilon$.*

3 Local Density Lower Bound

In this section we present a bound on the probable bandwidth of a graph drawn from $G_{n,B,p}$.

We first state a bound on the probability of large deviation from its expectation of the sum of m independent geometric variables. It can be proved from the Chernoff bound on the deviation of binomial random variables.

Lemma 11 *Let X be the sum of m independent random variables, each with geometric distribution with success probability p. Then for any $\varepsilon > 0$, $Pr[X \geq (1+\varepsilon)EX] \leq \exp(-\frac{1}{2}\frac{\varepsilon^2}{1+\varepsilon}m)$.*

In the spirit of Lemma 4, let (κ, ε) be such that with high probability the κ-neighborhood of every vertex v in G contains all vertices of difference $\leq (\kappa - \delta)B$ from v. According to our definition of high probability, we may assume that $v = 0$. Note that Lemma 4 shows that $\kappa = 2\gamma + 1$ and $\varepsilon > 0$ fixed satisfy the above requirement.

The following technical Lemma proves a probable lower bound on the bandwidth of a graph in $G_{n,B,p}$. The lower bound assumes the existence of (κ, ε), as defined above. In the sequel, the choice of κ and ε by Lemma 4 provides a lower bound on the bandwidth when $B \leq \sqrt{n}$. When p is fixed, an easier choice of (κ, ε) provides a lower bound for $B << n/\log n$.

Lemma 12 *Let G be a graph drawn from $G_{n,B,p}$, for $B << n/\log n$, and $\delta > 0$ fixed. Let (κ, ε) be as defined above. Then with high probability the optimal bandwidth G is $\geq B - (1+\delta)/p - \varepsilon B^2/n$.*

Proof. Let B^* be the optimal bandwidth of G, and let $diam(G)$ be the diameter of G. A known version of the local density bound states that $B^* \geq (n-1)/diam(G)$. In order to use this lower bound, it suffices to show a probable upper bound on the diameter of G. We show that with high probability the distance in the graph between vertices 0 and $n-1$ is bounded by a specified value. The same argument applies to all pairs of vertices, so with high probability this value bounds the diameter of the graph.

Recall Lemma 4 which guarantees that with high probability the κ-neighborhood of 0 contains all vertices between 0 and $(\kappa - \varepsilon)B$, for some fixed $\varepsilon > 0$. It suffices to bound the length of a path from $n-1$ into the λ-neighborhood of 0 (see Figure 6). Indeed, we show that with high probability there is a path of length $\beta = \frac{n}{B-(1+\delta)/p}$ from $n-1$ to some vertex $y \leq (\kappa - \varepsilon)B$. Altogether, with high probability there is a path of length $\kappa + \beta$ from $n-1$ to 0.

Consider a "greedy" path of length β from $n-1$ towards 0, taking each time the edge with the largest possible difference towards 0. At each such step, the edge of difference B exists with probability p, and we manage to advance B towards 0. If the edge does not exist, we try to advance on the edge of difference $B-1$ which also succeeds with probability p, and so on.

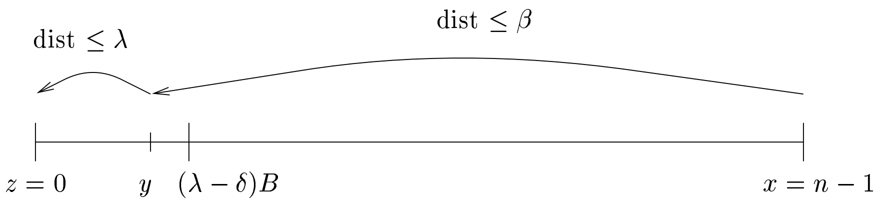

Figure 6: Diameter of $G_{n,B,p}$ is $\leq \beta + \lambda$

Therefore the advance at step i can be described by $B - X_i$, where X_i is a geometric random variable $X_i \sim G(p)$ (we assume that there exists at least one edge to advance on, which indeed happens with high probability). Note that the X_i are mutually independent since they correspond to advancing edges of different vertices). Denote $X = \sum X_i$, then the advance in β steps is $\sum_{i=1}^{\beta}(B - X_i) = \beta B - X$.

Clearly $EX = \beta/p$ and by Lemma 11, $Pr[X > (1+\delta)\beta/p] < e^{\delta'\beta}$ for $\delta' = -\frac{1}{2}\frac{\delta^2}{1+\delta}$. We shall soon see that $\beta >> \ln n$, so with high probability X is at most $(1+\delta)$ times its expectation and the advance in β steps is at least $\beta B - (1+\delta)\beta/p$. Recall that we want the advance to reach the κ neighborhood of 0, so we require that the advance in β steps is at least $n - (\kappa - \varepsilon)B$. We want $\beta(B - (1+\delta)/p) \geq n - (\kappa - \varepsilon)B$, so we choose $\beta \geq \frac{n-(\kappa-\varepsilon)B}{B-(1+\delta)/p}$. Note that indeed $\beta \geq n/2B >> \ln n$ as claimed. Altogether there is a path between 0 and $n-1$ of length $\leq \beta + \kappa$.

With high probability we bound the distance of all pairs of vertices simultaneously, so $diam(G) \leq \beta + \kappa$. This gives a lower bound on the graph bandwidth (with some elementary algebra, assuming $\varepsilon B > \kappa(1+\delta)/p$)

$$\begin{aligned} B^* \geq \frac{n}{\beta+\kappa} &= B - \frac{\varepsilon B^2 - B\kappa(1+\delta)/p + n(1+\delta)/p}{n + \varepsilon B - \kappa(1+\delta)/p} \\ &\geq B - \frac{\varepsilon B^2 + n(1+\delta)/p}{n} \geq B - \frac{(1+\delta)}{p} - \frac{\varepsilon B^2}{n} \end{aligned}$$

Since $B \leq \sqrt{n}$ and $\delta, \varepsilon > 0$ are chosen arbitrarily small, the bandwidth of the graph is with high probability at least $B - (1+\delta)/p$. $\square$

Theorem 2 *If $B \leq \sqrt{n}$, then with high probability the optimal bandwidth of $G \in G_{n,B,p}$ is $\geq B - (1+\delta)/p$.*

For fixed edge probability $p > 0$ we achieve tighter results by using better κ, ε.

Theorem 3 *Let p be fixed, and $B << n/\log n$. Then with high probability the optimal bandwidth of $G \in G_{n,B,p}$ is $\geq B - O(1)$.*

Proof. When p is fixed, we show that $\kappa = 2$ and $\varepsilon = O(\log B/B)$ suffice. So by Lemma 12, $B^* \geq B - (1+\delta)/p - O(B \log B/n) \geq B - O(1)$.

Indeed, for any vertex $v \leq (2-\varepsilon)B$, there are εB possible paths of lengths 2 from 0 to v. The paths are disjoint so the probability that none of them exists in G is $(1-p)^{2\cdot\varepsilon B} < e^{-2\varepsilon\cdot pB} = n^{-c'}$. Hence, with high probability v is contained in the 2-neighborhood of 0, and with high probability also all vertices $\leq (2-\varepsilon)B$. □

References

[AK97] Noga Alon and Nabil Kahale. A spectral technique for coloring random 3-colorable graphs. *SIAM Journal on Computing*, 26(6):1733–1748, December 1997.

[BKW98] Gunter Blache, Marek Karpinski, and Jürgen Wirtgen. On approximation intractability of the bandwidth problem. Manuscript, March 1998.

[Bop87] Ravi B. Boppana. Eigenvalues and graph bisection: An average-case analysis. In *28th Annual Symposium on Foundations of Computer Science*, pages 280–285, Los Angeles, California, 12–14 October 1987. IEEE.

[BS95] Avrim Blum and Joel Spencer. Coloring random and semi-random k-colorable graphs. *Journal of Algorithms*, 19(2):204–234, September 1995.

[CCDG82] P. Z. Chinn, J. Chvátalová, A. K. Dewdney, and N. E. Gibbs. The bandwidth problem for graphs and matrices—a survey. *J. Graph Theory*, 6(3):223–254, 1982.

[Chv80] J. Chvátalová. *On the bandwidth problem for graphs*. PhD thesis, Department of Combinatorics and Optimization, University of Waterloo, 1980.

[CM69] E. Cuthill and J. McKee. Reducing the bandwidth of sparse symmetric matrices. In *24th National Conference Proceedings*, pages 157–172, 1969.

[DF89] M. E. Dyer and A. M. Frieze. The solution of some random NP-hard problems in polynomial expected time. *Journal of Algorithms*, 10(4):451–489, December 1989.

[Fei97] Uriel Feige. Approximating the bandwidth via volume respecting embeddings. Manuscript, September 1997.

[FK97] Uriel Feige and Joe Kilian. Heuristics for graph coloring with improved performance guarantees. Manuscript, November 1997.

[FS92] Martin Fürer and C. R. Subramanian. Coloring random graphs. In *Algorithm theory—SWAT '92 (Helsinki, 1992)*, volume 621 of *Lecture Notes in Computer Science*, pages 284–291, Berlin, 1992. Springer.

[GGJK78] M. R. Garey, R. L. Graham, D. S. Johnson, and D. E. Knuth. Complexity results for bandwidth minimization. *SIAM J. Appl. Math.*, 34(3):477–495, 1978.

[GPS76] Norman E. Gibbs, Jr. Poole, William G., and Paul K. Stockmeyer. An algorithm for reducing the bandwidth and profile of a sparse matrix. *SIAM J. Numer. Anal.*, 13(2):236–250, 1976.

[Kuc77] L. Kucera. Expected behaviour of graph colouring algorithms. In *Fundamentals of Computation Theory*, volume 56 of *Lecture Notes in Computer Science*, pages 447–451. Springer-Verlag, Berlin/New-York, 1977.

[Mon86] Burkhard Monien. The bandwidth minimization problem for caterpillars with hair length 3 is NP-complete. *SIAM J. Algebraic Discrete Methods*, 7(4):505–512, 1986.

[MR95] Rajeev Motwani and Prabhakar Raghavan. *Randomized Algorithms*. Cambridge University Press, 1995.

[Pap76] Christos H. Papadimitriou. The NP-completeness of the bandwidth minimization problem. *Computing*, 16(3):263–270, 1976.

[Tur86] Jonathan S. Turner. On the probable performance of heuristics for bandwidth minimization. *SIAM Journal on Computing*, 15(2):561–580, May 1986.

[Tur88] Jonathan S. Turner. Almost all k-colorable graphs are easy to color. *Journal of Algorithms*, 9(1):63–82, March 1988.