Cops, Robbers, and Threatening Skeletons: Padded Decomposition for Minor-Free Graphs

Ittai Abraham* Cyril Gavoille† Anupam Gupta‡ Ofer Neiman§ Kunal Talwar¶

November 13, 2013

Abstract

We prove that any graph excluding K_r as a minor has can be partitioned into clusters of diameter at most Δ while removing at most $O(r/\Delta)$ fraction of the edges. This improves over the results of Fakcharoenphol and Talwar, who building on the work of Klein, Plotkin and Rao gave a partitioning that required to remove $O(r^2/\Delta)$ fraction of the edges.

Our result is obtained by a new approach to relate the topological properties (excluding a minor) of a graph to its geometric properties (the induced shortest path metric). Specifically, we show that techniques used by Andreae in his investigation of the cops-and-robbers game on excluded-minor graphs can be used to construct padded decompositions of the metrics induced by such graphs. In particular, we get probabilistic partitions with padding parameter $O(r)$ and strong-diameter partitions with padding parameter $O(r^2)$ for K_r-free graphs, padding $O(k)$ for graphs with treewidth k, and padding $O(\log g)$ for graphs with genus g.

*Microsoft Research SVC. Email: `ittaia@microsoft.com`.

†LaBRI - University of Bordeaux, Bordeaux, France. Email: `gavoille@labri.fr`.

‡Computer Science Department, Carnegie Mellon University, Pittsburgh, PA. Email: `anupamg@cs.cmu.edu`. Research was partly supported by NSF awards CCF-1016799 and CCF-1319811, and by a grant from the CMU-Microsoft Center for Computational Thinking.

§Department of Computer Science, Ben-Gurion University of the Negev, Beer-Sheva, Israel. Email: neimano@cs.bgu.ac.il. Supported in part by ISF grant No. (523/12) and by the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement n°303809.

¶Microsoft Research SVC. Email: `kunal@microsoft.com`.

1 Introduction

This paper considers the problem of constructing random partitioning schemes for minor-free graphs. Loosely speaking, the goal is to find a partition of the graph vertices so that each part (called a cluster) has small diameter, and the probability of any local neighborhood being cut (and not lying within some cluster) is small. There is a natural tradeoff between these two parameters (the diameter, and the probability of being cut). Such random partitions have found numerous applications in algorithm design, including: flow/cut gaps, metric embeddings, and recently as core primitives for several near linear time algorithms. Therefore improving the parameters of the partitions is a research program of considerable interest.

While tight parameters for such partitions are known in several settings, for the case of graphs that exclude some given graph H as a minor, the problem of finding the optimal tradeoff remains open. Progress was made in the seminal work of Klein, Plotkin and Rao [KPR93], and improved by Fakcharoenphol and Talwar [FT03]. Despite attracting the attention of several researchers (see, e.g., [Lee13]), the KPR framework remained the only known approach to this problem for over 20 years.

In this paper we make progress on this question and improve known parameters. Equally importantly, we also introduce techniques and structural insights that we hope will be useful for further improvements on this and related problems. In particular, we observe that a result of Andreae [And86] can be reinterpreted as a *structure theorem* for graphs excluding a fixed minor.[1] His result constructively gives us a *cop decomposition* of a graph, which is a lot like a tree decomposition except that instead of r vertices per bag, we have r shortest-paths (in some induced subgraph) in each bag. While such a cop decomposition may give less insight into the graph structure than the deep work of Robertson and Seymour [RS03], it has the benefit of a significantly better dependence on r. We extend this cop decomposition framework to produce probabilistic partitions, and we believe that this high level approach may be useful in getting better algorithms for other problems involving excluded-minor graphs.

To formally state our results, let us recall some notation. For an undirected weighted graph $G = (V, E)$ and a subset $C \subseteq V$, denote by $G[C]$ the induced subgraph on C. Let d_G denote the shortest path metric on G. The *(weak) diameter* of a set $S \subseteq V$ is $\max_{x,y \in S} d_G(x, y)$, whereas the *strong diameter* of the set S is $\max_{x,y\in S} d_{G[S]}(x, y)$—note that the latter distance is being measured in the induced subgraph.

Definition 1 (Δ-bounded partitions)**.** *A partition $P = \{C_1, \dots, C_t\}$ is* Δ-bounded *if for all i, the weak-diameter* $\mathrm{diam}(C_i) \leq \Delta$*. Partition P is* strong-diameter Δ-bounded *if the strong diameter* $\mathrm{diam}(G[C_i]) \leq \Delta$ *for all i.*

Given a partition $P = \{C_1, \dots, C_t\}$ of V, let $P(z)$ denote the unique cluster containing z.

Definition 2. *A distribution $\mathcal{P}$ over Δ-bounded partitions is* (β, δ)-padded *if for any $z \in V$ and any $0 \leq \gamma \leq \delta$,*

$$\Pr[B(z, \gamma\Delta) \subseteq P(z)] \geq 2^{-\beta\gamma} \ .$$

We call $\mathcal{P}$ β-padded *if it is* (β, δ)-padded *where δ is a universal constant that does not depend on β, and efficient if it can be sampled in polynomial time.*

Our main result is the following.

[1] For the rest of the discussion, we assume we are excluding a fixed graph of size r, say K_r.

Theorem 3. *Every K_r-minor-free graph G admits an efficient $O(r)$-padded partition scheme.*

It has long been known that for arbitrary graphs the best possible padding parameter is $\Theta(\log |V|)$. For special cases better bounds are known, e.g., for metrics of doubling constant λ, the padding parameter is $\Theta(\log \lambda)$ [GKL03]. For graphs that can be drawn on a surface of genus g, ideas developed in a recent sequence of papers [IS07, BLS10, Sid10] have culminated in the optimal padding parameter of $\Theta(\log g)$ [LS10].

The first bounds for K_r-minor-free graphs were due to the influential work of Klein, Plotkin, and Rao [KPR93], who gave $(O(r^3), 1/r)$-padded partition scheme. Fakcharoenphol and Talwar [FT03] improved this to an $(O(r^2), 1/r)$-padded partition scheme. In this work, we improve the padding parameter from $O(r^2)$ to $O(r)$; moreover, we provide padding guarantees to larger balls—the previous guarantees give padding only for balls of diameter $< O(\Delta/r)$, compared to $O(\Delta)$ for our result. The partitioning scheme in [KPR93] was motivated by bounding the maximum-multicommodity-flow/sparsest-cut gap for K_r-free graphs. Subsequently, it found applications to metric embeddings [Rao99, Rab03] with its natural connections to edge-cut problems [Mat02] and also to vertex-cut problems [FHL08], to bounding higher eigenvalues and higher-order Cheeger inequalities for graphs [BLR10, KLPT09, LGT12], to metric extension problems and approximation algorithms [CKR05, AFH$^+$04, LN05], and others. The quantitative improvements given by our results thus give improvement in all these settings.

Theorem 3 above gives us a weak-diameter guarantee. However, our techniques are versatile, and can be extended to give strong-diameter partitions—in particular, we obtain the following results.

Theorem 4. *Let $G = (V, E)$ be an undirected weighted graph.*

1. *If G is a K_r-free graph then it admits an efficient $(O(r^2), O(1/r^2))$-padded strong-diameter decomposition scheme.*

2. *If G is a treewidth r graph then G admits an efficient $(O(r), O(1/r))$-padded strong-diameter decomposition scheme.*

3. *If G is a genus g graph then G admits an efficient $O(\log g)$-padded strong-diameter decomposition scheme.*

The first result in Theorem 4 is an exponential improvement over the strong-diameter partitions of [AGMW10]. The third result strengthens the result of [LS10] by providing the same asymptotic padding guarantees while ensuring that clusters have a strong-diameter.

1.1 Discussion of Techniques

How does one prove a property for a graph that does not contain a K_r minor? One approach relies on the deep results of Robertson and Seymour that turn this "negative" property (namely, that of not having a certain minor) into a "positive" constructive one, by giving a complete structural characterization of how such graphs can be built from simple building blocks by applying simple rules to them. This structure theorem allows one to prove properties of excluded-minor graphs by structural induction on this construction procedure. Unfortunately, this approach typically inherits the rather bad dependence on r from the Robertson-Seymour structure theorem [RS03]. Nevertheless, this approach has been highly successful and used to prove several results for excluded-minor graphs.

The other approach is to work more directly and design an algorithm establishing the property we want to prove, with the guarantee that the failure of the algorithm constructs K_r as a minor. This approach is usually problem-specific but usually leads to better dependence on r. Examples of this approach include the work of Andreae [And86] for the cops-and-robbers game, results of Alon, Seymour and Thomas [AST90] on separators, and the aforementioned work of Klein, Plotkin and Rao [KPR93]. This is the approach we take.

Let us now give a high-level description of some of the ideas and techniques used to prove Theorem 3 and Theorem 4.

The Bounded Threatener Program. A natural approach to obtain Δ-bounded β-padded probabilistic partitions is to find a set of "suitable" centers S, and iteratively build balls around the points in S with radii drawn from a truncated exponential distribution in the range $[\Delta/4, \Delta/2]$ with rate β. The memoryless property of the exponential distribution ensures that balls of radius $\approx \Delta/\beta$ avoid being cut with constant probability, conditioned on the exponential distribution not being truncated. To handle the truncation, we need to bound the number of centers at distance at most $(1/2 + 1/\beta)\Delta$ from some vertex z. We will call such centers the *threateners* of z. If the number of threateners is bounded by $2^{O(\beta)}$ then a trivial union bound implies that with constant probability none of them will reach diameter $(1/2 - 1/\beta)\Delta$ and intersect the ball $B(z, \Delta/\beta)$. A contribution of this work is in extending the bounded threatener program and showing how a bound on the *expected* number of threateners suffices.

Cop Decompositions. Andreae [And86] considered the following game, a set of cops plays against a robber. At each round the robber can move across one edge and then each one of the cops can move across one edge. The cops win if they land on the same vertex as the robber. Andreae showed that if G is K_r-free then $O(r^2)$ cops have a winning strategy. The cop strategy is simple: each cop can control one shortest path and together they try to build a K_r minor. The shortest paths controlled by the cops induce a set of supernodes (disjoint connected subsets) and edges containing a minor that is a subgraph of K_r. At each round one can fix a center for a new supernode and use free cops to connect the center to all previous supernodes via shortest paths. The new center and each new shortest path is fully contained in the component containing the robber that is induced by removing the supernodes from G (hence this new path is disjoint from all previous supernodes). We can interpret Andreae's result as constructing a *cop-decomposition* of width r, which is similar to a tree-decomposition; the difference being that in a tree-decomposition each bag B contains at most *width*+1 vertices while in a cop-decomposition each bag B contains a tree with at most *width*+1 leaves and each root-leaf path $p \in B$ is a shortest path in the subgraph induced by the bag B and its subtree in the cop-decomposition.

From Cop Decompositions to Padded Partitions via Skeletons. The cop decomposition induces a partition of the vertices of the graph into trees. Note that the number of vertices in each tree in a cop decomposition may be large, and depend on n. Why are these trees useful? Since each tree contains at most r shortest paths in the induced subgraph, one can choose a "net" of centers along each path so that each node in the graph is threatened by $O(r)$ centers from any one tree. Hence it now suffices to bound the number of trees that get close enough to a vertex z so that some centers from this tree may threaten z. (We call such a tree a *"threatening skeleton"* for z.) As mentioned above, we do not bound the worst-case number of such threatening skeletons; we prove it suffices to bound their expected number.

Bounding the Expected Number of Threateners. How to bound the expected number of threatening skeletons for some node $z \in V$? We need a notion of progress. The cop-decomposition ensures that in any given moment there are at most r trees (a.k.a. threatening skeletons) that z can see on the boundaries of its component, where each tree consists of at most r paths. We observe the following property of the distances from z to these trees: if constructing a new tree T_{new} in the induced subgraph containing z causes some current tree T_{curr} to become farther from z (or even to be disconnected from z) because it cuts off some short path from z to T_{curr}, the distance from z to T_{new} is strictly less than the distance from z to T_{curr}. Indeed, if this distance were to miraculously decrease (deterministically) by Δ/k then one can prove a bound of $O\binom{r+k}{k}$ on the number of threateners. But why should such a large decrease happen? It doesn't, but *we force this to happen.* We change the above construction and build a "buffer" of some random radius around each skeleton we build. Note that the supernodes did not have to be trees in the above arguments, and hence "fattening" them by growing buffers around the trees would not change any of the preceding arguments. Now by choosing the buffer radius from a truncated exponential with rate $O(r)$, we may naïvely hope to decrease the distance by Δ/r with constant probability (assuming no truncation). The proof is much more subtle, and requires to overcome the truncation of the buffer. We use a potential function with delicately chosen parameters, such that for each new tree, this potential increases in expectation by $\approx r/2^r$. The potential starts at 0 and once it reaches r, it means that z is at distance 0 from some buffered tree and will not be threatened again. Finally, the optional stopping theorem helps us bound the expected number of threateners by $\approx 2^r$.

Bounding Expected Drift in Potential. In order to bound the number of threateners for z, the potential function we use is a sum of exponentials $\sum_{\text{buffers}B} e^{-\alpha\, d(z,B)}$ for some parameter α; the sum is over those buffered trees that the node z can see. The main challenge is that in the worst case, one new buffered tree can cause all the other current buffered trees to be disconnected from the component containing z, hence losing r summands of the potential. To overcome this we need to guarantee that the expected gain from the new tree is $O(r)$ times more than the expected loss of any single current tree, which is one of the technical cores of the analysis. We note that obtaining any deterministic bound on the number of threateners using a cop decomposition remains an open question.

1.2 Other Related Work

The ideas of either finding a "good" decomposition or else building a K_r-minor used by [KPR93, And86] also appear in "shallow-minor theorems" of Alon, Seymour, and Thomas [AST90], Plotkin, Rao, and Smith [PRS94], and others. The parameters and run-times of these constructions have been considerably improved, see the paper of Wulff-Nilsen [WN11] and the references therein.

Busch, LaFortune, and Tirthapura [BLT07] first suggested the idea of decomposing a graph into paths and building balls around these paths; they considered this in the context of strong-diameter covers. They give the best constants for covers of planar graphs; for K_r-free graphs, they give $O(1)$-padding and $O(\log |V| \cdot f(r))$-overlap, where $f(r)$ depends on the Robertson-Seymour structure theorem.

In contrast to the weak-diameter partitions of [KPR93, FT03], the previously best *strong-diameter* partitions are due to [AGMW10], who guarantee strong diameter Δ and probability of an edge $\{u, v\}$ being separated is $O(6^r r^2 \cdot \frac{d(u,v)}{\Delta})$. [AGMW10] also present *sparse covers* with strong-diameter Δ, padding of $O(r^2)$ and overlap of $2^{O(r)} r!$.

The papers [IS07, BLS10, Sid10] give algorithms to probabilistically embed genus-g graphs into planar graphs with $2^{O(g)}$, $O(g^2)$ and $O(\log g)$ distortion respectively. The ideas developed in this line of work lead to an asymptotically optimal padding parameter of $O(\log g)$ for genus-g graphs [LS10].

For general graphs, the decomposition schemes in, e.g., [Awe85, LS93, Bar96, CKR05, FRT04] give asymptotically optimal $O(\log |V|)$ padding. The best result known for treewidth-r graphs was the same as for K_r-free graphs, i.e., $O(r^2)$-padding partitions.

1.3 Organization of the Paper

After a few preliminary definitions, we provide in Section 3 a bound on the expected number of threateners for a wide range of partition algorithms. We then show how to use this to bound the padding probability. Our main result Theorem 3 is proved is Section 4. The three assertions of Theorem 4 are then proven in Sections 5, 6 and 7.

2 Definitions and Notation

Graphs. We assume familiarity with graph-theoretic notions; see, e.g., [Die00] for background. Here are some definitions we will use. Given a graph $G = (V, E)$, a ball around $v \in V$ of radius $r \geq 0$ is $B_G(v, r) = \{u \in V \mid d_G(v, u) \leq r\}$, and similarly for a subset $A \subseteq V$, $B(A, r) = \{u \in V \mid d_G(A, u) \leq r\}$. Also let $N(A) = \{u \in V \mid \exists v \in A,\ \{u, v\} \in E\}$. For subsets $A, B \subseteq V$ define a relation $\sim$ where $A \sim B$ iff $A \cap N(B) \neq \emptyset$, that is, iff there is an edge between A and B.

A *minor* of G is a graph G' obtained by deleting and contracting edges. Equivalently, G' is a minor of G if there exists a map $f : V(G) \to V(G')$ such that (a) for each $u' \in V(G')$ the "supernode" $f^{-1}(u')$ is connected in G, and (b) for every edge $\{u', v'\} \in E(G')$, there is at least one edge between $f^{-1}(u')$ and $f^{-1}(v')$ in $E(G)$. A graph G is *H-free* (or *excludes an H-minor*) if G does not contain a subgraph isomorphic to H as a minor. As is well-known, planar graphs are exactly the graphs excluding $K_{3,3}$ and K_5 as minors. In fact, Robertson and Seymour proved that every graph family closed under taking minors is characterized by a set of excluded minors.

Many one-way implications are also known: if we can show that a class $\mathscr{G}$ of graphs is closed under taking minors, and $H \notin \mathscr{G}$, then $\mathscr{G}$ contains only H-free-graphs. Hence, graphs with treewidth at most r are K_{r+2}-free (since treewidth of a clique is one smaller than its size, and the treewidth of a graph does not increase under edge deletions and contractions); graphs with genus g exclude K_r as a minor for some $r = \Theta(g^2)$, since the genus of K_r is $\Theta(g^2)$.

Truncated Exponential Distributions. We will extensively use the following probability distribution over positive reals. The $[\theta_1, \theta_2]$-*truncated exponential distribution* with parameter b is denoted by $\mathrm{Texp}_{[\theta_1,\theta_2]}(b)$, and has the probability density function:

$$f_{texp;b;\theta_1,\theta_2}(y) := \frac{b\, e^{-b\cdot y}}{e^{-b\cdot\theta_1} - e^{-b\cdot\theta_2}} \qquad \text{for } y \in [\theta_1, \theta_2]. \tag{2.1}$$

For the $[0, 1]$-*truncated exponential distribution* we drop the subscripts and denote it by $\mathrm{Texp}(b)$; the density function is

$$f_{texp;b}(y) := \frac{b\, e^{-b\cdot y}}{1 - e^{-b}} \qquad \text{for } y \in [0, 1]. \tag{2.2}$$

Note that if $Y \sim \mathrm{Texp}(b)$ then $u \cdot Y \sim \mathrm{Texp}_{[0,u]}(b/u)$.

3 Analysis

Our algorithms induce an iterative process that creates “skeletons” (e.g., trees, paths, or vertices) and remove their neighborhoods (a buffer), defined according to some truncated exponential distribution, from the graph. Once we have these skeletons, our algorithms define a second iterative process that creates clusters from the skeletons[2].

Let us abstract out the properties needed from our first and second processes.

Definition 5. *[Skeleton-Process] Given a graph G, parameters $0 \leq l < u \leq 1$ and $b > 0$, any process which generates a sequence of graphs $G = G_0, G_1, \ldots,$ skeletons $A_0, A_1, \ldots$ and vertex sets $K_0, K_1, \ldots,$ that satisfies the following property is a* skeleton-process*:*

- *For any $i \geq 0$ we are given some $A_i \subseteq V(G_i)$, and define $K_i = B_{G_i}(A_i, R_i\Delta)$, where $R_i \sim \mathrm{Texp}_{[l,u]}(b/(u-l))$.*

The process is threatening *if the graph sequence satisfies $G_{i+1} = G_i \setminus K_i$, and the process is* cutting *if the graph sequence satisfies $G_{i+1} \supseteq G_0 \setminus (\cup_{j\leq i} K_j)$.*

The first process is a threatening process which creates buffers around the trees of the cop-decomposition. The second process is a cutting process that creates the actual clusters centered at net-points of the trees. For the strong-diameter results, we will have a single process that satisfies both definitions.

3.1 Analysis of the threatening process: Bounding the expected threats

A crucial property of all of our algorithms is that any point z can “see” at most s buffers (the K_i sets) at any time, for some parameter s (in the weak-diameter partition we will have $s = r$). By this we mean that for any connected component C in one of the remaining graphs (after some buffers were removed), there are at most s buffers that are connected to C by an edge. This property will enable us to prove that any vertex z is expected to be “threatened” by a small number of skeletons, that is, we expect a few skeletons that are sufficiently close to cut a certain ball around z.

Consider a threatening skeleton-process with parameters $l = 0$, $u \in [0, 1]$ and $b = 2s$. We prove a bound on the expected number of threateners for a ball around any vertex $z \in V(G)$ with padding parameter $\gamma > 0$. Let $\mathcal{J}_z = \{A_i \mid d_{G_i}(z, A_i) \leq (u+\gamma)\Delta\}$ be the set of vertex sets whose subset K_i may intersect $B_z = B_G(z, \gamma\Delta)$. Observe that once $z \in K_t$ for some integer t then it is removed from the graph, and $\mathcal{J}_z$ cannot increase anymore. For a connected component $C_i \in G_i$ let $\mathcal{K}_{|C_i} = \{K_j \mid j < i \wedge C_i \sim K_j\}$. (Recall that $A \sim B$ if there exists an edge from a node in A to some node in B.)

Lemma 6. *Suppose that in a threatening skeleton-process we have the property that for every $i \in \mathbb{N}$ and every connected component $C_i \in G_i$, we are guaranteed that $|\mathcal{K}_{|C_i}| \leq s$, then*

$$\mathbb{E}[|\mathcal{J}_z|] \leq 3e^{(2s+1)\cdot(1+\gamma/u)} .$$

We defer the proof to Appendix A.

[2] For our strong-diameter results we have a single process that extracts skeletons and clusters in one sweep.

3.2 Analysis of the cutting process: Bounding the probability of cutting a ball

In this section we give a bound on the probability that a ball is cut by a cutting skeleton-process, which depends on the *expected* number of threateners.

Consider a cutting skeleton-process as in Definition 5 with parameters $0 \leq l < u \leq 1$, $b > 0$. Fix $z \in V(G)$, a parameter $\gamma > 0$ and set $B_z = B_G(z, \gamma\Delta)$. Let $\mathcal{T}_z = \{A_i \mid d_{G_i}(z, A_i) \leq (u+\gamma)\Delta\}$ be the set of vertex sets whose subset K_i may intersect B_z. Let $N := |\mathcal{T}_z|$ be a random variable with $\tau = \mathbb{E}[N]$. We say that B_z is *cut* by the skeleton-process if it intersects more than a single K_i.

Lemma 7. *For $\delta = e^{-2b\gamma/(u-l)}$, the probability that $B_z = B_G(z, \gamma\Delta)$ is cut by a cutting skeleton-process with the property that $\tau = \mathbb{E}[|\mathcal{T}_z|]$, is at most*

$$(1-\delta)\left(1 + \frac{\tau}{e^b - 1}\right) .$$

We defer the proof to Appendix B.

4 A Weak-Diameter Partition

In this section, we show how to construct a weak-diameter partition for K_{r+1}-free graphs which is $O(r)$-padded (with constant $\delta = 1/40$). The ideas here will later extend to the case of strong-diameter partitions with a weaker $(O(r^2), O(1/r^2))$-padding.

4.1 The Algorithm

At a high level, the algorithm works as follows: in each step, pick a connected component of the remaining graph, and find (in a specific way) a shortest-path tree T in this component. Delete a random neighborhood of T from the graph, and recurse on each connected component of the graph, if any. We then construct a net of points on each tree, and from these net points grow "balls" of random radius to form the small-diameter regions of the partition. A key property to ensure the padding guarantee is that each node is expected to be close to few of these paths. We show that this property holds, otherwise we can construct a K_{r+1}-minor in G.

More specifically, the algorithm maintains a set of trees T_i and supernodes S_i that will be used in the construction, each tree and supernode have a "center" vertex associated with them. Let us describe a generic i-th iteration of the algorithm. Let $\mathcal{S}$ be the set containing all the supernodes created so far, initially this will be empty. Let C be a connected component in the graph $G_i = G \setminus (\cup\mathcal{S})$, where $\cup\mathcal{S}$ is the set of all vertices lying in the supernodes in $\mathcal{S}$, initially this will be the entire graph. Let $\mathcal{S}_{|C} = \{S \in \mathcal{S} : S \sim C\}$ be the set of supernodes that have a neighbor in component C. Say $\mathcal{S}_{|C} = \{S'_1, S'_2, \ldots, S'_k\}$, and consider the nodes $F_j = N(S'_j) \cap C$ for each supernode, which are vertices in C neighbors of these "adjacent" supernodes. (These F_j's may intersect.) We pick an arbitrary vertex u_i from C and build a tree T_i rooted at u_i, which is comprised of shortest paths from u_i to each of the sets F_j. Define the next supernode

$$S_i := B_{G_i}(T_i, R_i\Delta),$$

where $R_i \sim \mathrm{Texp}_{[0,1/8]}(16r)$. (Recall the definition of the truncated exponential distribution from (2.1).)

In order to create the random partition, choose a $\Delta/8$-net N_i over T_i, and enumerate $N_i = \{v_1, \ldots, v_{|N_i|}\}$. For each $1 \leq j \leq |N_i|$, create a cluster $B_{G_i}(v_j, \alpha_j\Delta) \cap U_j$ (where U_j is the set of points which have no cluster yet), where each $\alpha_j \sim \text{Texp}_{[1/4,1/2]}(20r)$. This completes the description of the algorithm; it is also given as Algorithm 1 and 2.

Algorithm 1 `Weak-Random-Partition`(G,Δ,r)

Let $G_0 \leftarrow G$, $i \leftarrow 0$.
Let $\mathcal{S} \leftarrow \emptyset$.
Let $\mathcal{T} \leftarrow \emptyset$.
while G_i is non-empty **do**
  Let C_i be a connected component of G_i.
  Pick $u_i \in C_i$. Let T_i be a tree rooted at u_i that consists of shortest paths (in G_i) from u_i to the closest vertex of $N(S)$ for each supernode $S \in \mathcal{S}_{|C_i}$.
  Let R_i be a random variable drawn independently from the distribution $\text{Texp}_{[0,1/8]}(16r)$.
  Let $S_i \leftarrow B_{G_i}(T_i, R_i\Delta)$ be a neighborhood of T_i.
  Add S_i to $\mathcal{S}$.
  Add T_i to $\mathcal{T}$.
  $G_{i+1} \leftarrow G_i \setminus S_i$.
  $i \leftarrow i + 1$.
end while
return `Create-Balls`$(G,\mathcal{T},\Delta,r)$.

Algorithm 2 `Create-Balls`$(G,\mathcal{T},\Delta,r)$

$P = \emptyset$.
for $i = 1, \ldots, |\mathcal{T}|$ **do**
  Let $N_i = \{v_1, \ldots, v_{|N_i|}\}$ be a $\Delta/8$-net of T_i.
  for $j = 1, \ldots, |N_i|$ **do**
    Let α_j be a random variable drawn independently from the distribution $\text{Texp}_{[1/4,1/2]}(20r)$
    Add $B_{G_i}(v_j, \alpha_j\Delta) \setminus \cup P$ as a cluster to the partition P.
  end for
end for
return P.

4.2 The Analysis

The following invariant holds for each time step i:

Invariant 1. *For every $i \geq 0$, every connected component C of G_i satisfies that if $S, S' \in \mathcal{S}_{|C}$ then $S \sim S'$.*

Proof. The proof is by induction; the base case is trivial as there are no supernodes in $\mathcal{S}_{|C}$. Now by induction, assume that the invariant holds in G_i. Let T_i and S_i be the tree and supernode constructed in step i in the component C_i. Let C be some connected component of G_{i+1}, and $S, S' \in \mathcal{S}_{|C}$. If $C \cap C_i = \emptyset$ then C is a component of G_i as well; moreover, as $S_i \subseteq C_i$ it must be that $S_i \nsim C$ so neither of S, S' can be S_i, and hence we can use the induction hypothesis to infer that $S \sim S'$. On the other hand, suppose that $C \subseteq C_i$. There are two cases: if $S_i \notin \{S, S'\}$ we have

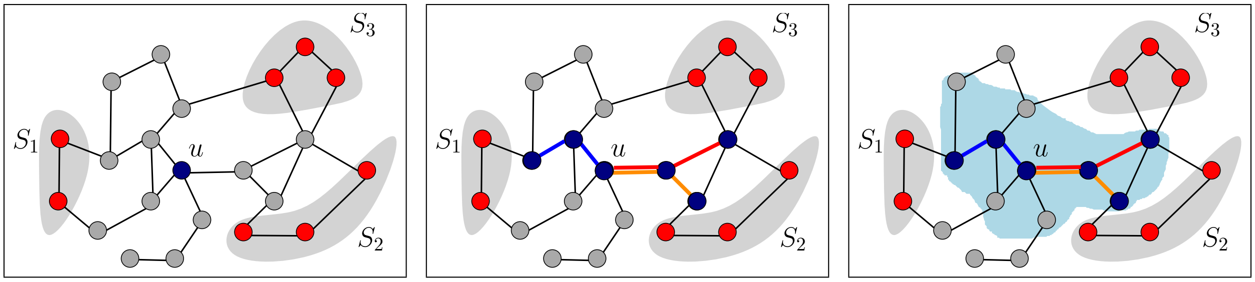

Figure 4.1: *An iteration of the algorithm. On the left, there are three supernodes S_1, S_2, S_3 neighboring the current component with u as a root. In the middle, we have a tree T_4 comprised of three shortest path from u. On the right, the new supernode S_4 which is a 1-neighborhood of T_4 (observe that this neighborhood is taken in the connected component containing u).*

$S \sim S'$ by the induction hypothesis on C_i. On the other hand, suppose $S_i = S$ (w.l.o.g.). Recall that T_i was chosen so that it contains a neighbor of every supernode in $\mathcal{S}_{|C_i}$ and $T_i \subseteq S_i$, we have that $S_i \sim S'$. ☐

Invariant 1 implies that for each connected component C, contracting the supernodes of $\mathcal{S}_{|C}$ yields a $K_{|\mathcal{S}_{|C}|}$ minor, so we obtain the following corollary.

Corollary 8. *If G excludes K_{r+1} as a minor, then for every time step i, the connected component C_i has $|\mathcal{S}_{|C_i}| \leq r$. In particular, the tree T_i is made up of at most r shortest paths in G_i.*

Claim 9. *The algorithm above generates a Δ-bounded partition of G.*

Proof. First we prove that we generate a partition. Indeed, we delete supernodes from the graph, and recurse on the remaining components, so we need to show that vertices within the supernodes are contained in some cluster. Consider a vertex x in supernode S_i. By definition, $d_{G_i}(x, T_i) \leq \Delta/8$. Since N_i is a $\Delta/8$-net in T_i, some net point $v_j \in N_i$ satisfies $d_{G_i}(x, v_j) \leq \Delta/4$. And since $\alpha_j \geq 1/4$, the ball $B_{G_i}(v_j, \alpha_j\Delta)$ contains x. Hence each node within the deleted supernode is contained in some cluster, and we get a partition of G. Moreover, each cluster is a ball of radius at most $\alpha_j\Delta \leq \Delta/2$ (and hence diameter at most Δ) in G_i. Finally, distances in G_i are no smaller than those in G. ☐

Lemma 10. *For $r \geq 4$, and any $\gamma \leq 1/40$, the probability that a ball B_z of radius $\gamma\Delta$ is cut by the above process is*

$$\Pr[B_z \text{ cut}] \leq 1 - e^{-80r\gamma}.$$

Proof. First observe that the process defined in Algorithm 1. is a threatening skeleton-process, with the sequence of graphs $G_0, G_1, \ldots$ as defined in the algorithm and with $A_i = T_i$, $K_i = S_i$, $l = 0, u = 1/8$, $s = r$ and $b = 2s$. Recall that $B_z = B_G(z, \gamma\Delta)$ and $\mathcal{J}_z = \{T_i \mid d_{G_i}(z, T_i) \leq (u+\gamma)\Delta\}$. By Invariant 1 we get that for all $i \in \mathbb{N}$, $|\mathcal{S}_{|C_i}| \leq r$, so by Lemma 6 (using that $\gamma \leq 1/40$),

$$\mathbb{E}[|\mathcal{J}_z|] \leq 3e^{(2r+1)\cdot(1+\gamma/u)} \leq 10e^{5r/2}\ . \tag{4.3}$$

For each i such that $T_i \in \mathcal{J}_z$, let $U_i = \{v \in N_i \mid d_{G_i}(v, z) \leq (1/2+\gamma)\Delta\}$ be the net points in N_i that are sufficiently close to threaten B_z, and denote $\mathcal{T}_z = \cup_{i|T_i \in \mathcal{J}_z} U_i$. By Corollary 8, T_i is comprised of at most r shortest paths, and we claim that on each shortest path there can be at most 10 points that are in U_i. This is because the distance between any two consecutive net points on a path is at least $\Delta/8$, and if there are $q > 10$ points, because this is a shortest path, the distance from the first point to the last is at least $(q-1)\cdot\Delta/8 > (1+2\gamma)\Delta$. The triangle inequality implies that it can't be that both are within $(1/2+\gamma)\Delta$ from z. We conclude that for all i (with $T_i \in \mathcal{J}_z$) we have $|U_i| \le 10r$, thus by (4.3)

$$\tau := \mathbb{E}[|\mathcal{T}_z|] \le 10r \cdot 10e^{5r/2} = 100r \cdot e^{5r/2} \ . \tag{4.4}$$

Next, we show that our `Create-Balls` algorithm generates a cutting skeleton-process. Simply take the sequence $G_0, \dots, G_0, G_1, \dots, G_1, G_2, \dots,$ where each G_i is taken $|N_i|$ times. Then the skeleton sets A are in fact singletons: for each i we will take $|N_i|$ sets - the points of N_i, to be these singletons. The parameters for the exponential distribution are $l = 1/4$, $u = 1/2$ and $b = 5r$. To see the cutting property of Definition 5, note that once we move from the graph G_i to G_{i+1}, G_{i+1} will contain all the points yet uncovered by clusters, because we already observed in Claim 9 that once all the points of N_i create a cluster, the supernode S_i is completely covered (recall $G_{i+1} = G_i \setminus S_i$). Finally, applying Lemma 7, we obtain that the probability that B_z is cut is at most

$$(1-e^{-2b\gamma/(u-l)})\left(1+\frac{\tau}{e^b-1}\right) \ = (1-e^{-40r\gamma})\left(1+\frac{100r\cdot e^{5r/2}}{e^{5r}-1}\right) \ .$$

The expression $\frac{100r\cdot e^{5r/2}}{e^{5r}-1} \le e^{-r}$ for $r \ge 4$, and this completes the proof as

$$(1-e^{-40r\gamma})\cdot(1+e^{-r}) \le (1-e^{-40r\gamma})\cdot(1+e^{-40r\gamma}) = 1-e^{-80r\gamma} \ ,$$

using that $\gamma \le 1/40$. $\square$

5 A Strong Diameter Partition

In the previous section, we saw how to get a weak-diameter partition for minor-free graphs. In this section, we give a strong-diameter guarantee with a slightly weaker padding parameter of $(O(r^2), O(1/r^2))$ instead of $O(r)$. However, this is still an exponential improvement over the best previous padding for such strong-diameter partitions of minor-free graphs.

5.1 The Algorithm

The algorithm for strong-diameter partitions is similar in spirit to that of Section 4.1 for weak-diameter partitions, but there are some crucial differences that we highlight here.

At a high level, the algorithm works as follows: in each step, pick a connected component of the remaining graph, and find (in a specific way) a shortest path P in this component. Delete a random neighborhood of P from the graph, and recurse on each connected component of the graph, if any. Each such random neighborhood is decomposed into small diameter regions using cones centered at some of P's points. A key property to ensure the padding guarantee is that each node is expected to be close to few of these paths. We show that this property holds, otherwise we can construct a K_{r+1}-minor in G.

The algorithm again maintains a set of paths (instead of trees), and associated supernodes that will be used in the construction. These will be denoted as P_{ij} and S_i respectively, and supernode S_i will consist of the union of neighborhoods of the paths P_{ij}. The main difference from the weak-diameter construction is that instead of building a shortest-path tree all at once, we build a "tree" one path at a time, and remove a neighborhood of the path from the graph before constructing the subsequent paths.

Let us describe the i-th iteration of the algorithm. Let $\mathcal{S} \subseteq V$ be the set containing all the supernodes created so far. Let C_i be a connected component in the graph $G_i = G \setminus (\cup \mathcal{S})$. Let $\mathcal{S}_{|C_i} = \{S \in \mathcal{S} : S \sim C_i\}$ be the set of supernodes that have a neighbor in component C_i. We pick an arbitrary vertex u_i from C_i and build a supernode S_i. Again, the intuition behind the construction is that we wish for the new supernode to "touch" every supernode $S \in \mathcal{S}_{|C_i}$ (i.e., $S_i \sim S$). However, this is done slightly differently from Section 4.1, one path at a time. At the first iteration ($j = 1$) we create a shortest path P_{ij} from u_i to some supernode $S \in \mathcal{S}_{|C_i}$, and remove a random neighborhood S_{ij} from the graph to obtain $G_{i(j+1)}$. This neighborhood S_{ij} is defined as all the vertices within distance $R_{ij} \cdot \Delta$ of P_{ij} (in the current component C_{ij}), where $R_{ij} \sim \mathrm{Texp}_{[0,1/4]}(8(r^2 + r))$. We increase the iteration counter j and continue in this manner on every connected component of G_{ij} that is contained in C_i, until the new supernode $S_i = \cup_j S_{ij}$ touches every supernode $S \in \mathcal{S}_{|C}$ for every connected component $C \subseteq C_i$ in the remaining graph G_{ij}.

Finally, each such neighborhood S_{ij} is partitioned to "cones". Each cone B, centered at some (yet uncovered) point $c \in P_{ij}$, consists of the (yet uncovered) points in S_{ij} whose distance to c is not "much larger" than their distance to P_{ij}. The notion of being "much larger" is determined by a random variable α drawn independently and uniformly from $[\Delta/8, \Delta/4]$. The algorithms are formally presented as Algorithms 3 and 4 respectively. Observe that the subroutine `Create-Cones` is invoked in line 13 of `Strong-Random-Partition`.

Algorithm 3 `Strong-Random-Partition`(G,Δ,r)

1: Let $G_0 \leftarrow G$, $i \leftarrow 0$.
2: Let $\mathcal{S} \leftarrow \emptyset$.
3: Let $\mathcal{C} \leftarrow \emptyset$.
4: **while** G_i is non-empty **do**
5:   Select a connected component C_i of G_i, and pick $u_i \in C_i$.
6:   Let $W = \{u_i\}$.
7:   Let $j = 1$ and $G_{ij} = G_i \setminus W$.
8:   **while** there exist a connected component C_{ij} in G_{ij} and a supernode $S \in \mathcal{S}_{|C_{ij}}$ such that $C_{ij} \sim S$ and $C_{ij} \sim W$ but $W \nsim S$ **do**
9:     Choose $u \in N(W) \cap C_{ij}$.
10:     Let P_{ij} be a shortest path (in G_{ij}) from u to $N(S)$.
11:     Let R_{ij} be a random variable drawn independently from the distribution $\mathrm{Texp}_{[0,1/4]}(8(r^2 + r))$.
12:     Let $S_{ij} \leftarrow B_{G_{ij}}(P_{ij}, R_{ij}\Delta)$ be a neighborhood of P_{ij}.
13:     `Create-Cones`(S_{ij},P_{ij}).
14:     $W \leftarrow W \cup S_{ij}$.
15:     $G_{i(j+1)} \leftarrow G_{ij} \setminus S_{ij}$.
16:     $j \leftarrow j + 1$.
17:   **end while**
18:   Set $S_i = W$, and add S_i to $\mathcal{S}$.
19:   $G_{i+1} \leftarrow G_i \setminus S_i$.
20:   $i \leftarrow i + 1$.
21: **end while**

Algorithm 4 `Create-Cones`(S,P)

while $P \neq \emptyset$ **do**
    Choose $c \in P$.
    Choose $\alpha \in [1/8, 1/4]$ uniformly at random.
    Let $B = \{u \in S \mid d_S(u,c) - d_S(u,P) \leq \alpha\Delta\}$. Add B to $\mathcal{C}$.
    Set $S \leftarrow S \setminus B$.
    Set $P \leftarrow P \setminus B$.
end while

5.2 The Analysis

We begin by arguing that the algorithm creates a partition $\mathcal{C}$ with strong diameter Δ. The following properties will be useful.

Proposition 11. *For any S and P obtained during the run of the algorithm* `Create-Cones`*:*

- *If $u, v \in S$ are such that a shortest path from u to P contains v, and $v \in B$ for a cone B, then also $u \in B$.*
- *If $u, v \in S$ are such that a shortest path from u to c contains v, and $u \in B$ for a cone B centered at c, then also $v \in B$.*

Proof. Let $c \in P$ be the center of the cone B. We begin by proving the first item: Since $v \in B$ we have that $d_S(v,c) - d_S(v,P) \leq \alpha\Delta$. Since v is on the shortest path from u to P, $d_S(u,P) = d_S(u,v) + d_S(v,P)$ and thus

$$d_S(u,c) - d_S(u,P) \leq (d_S(u,v) + d_S(v,c)) - (d_S(u,v) + d_S(v,P)) = d_S(v,c) - d_S(v,P) \leq \alpha\Delta \ ,$$

which implies that $u \in B$.

The second item is proved in a similar manner: Since $u \in B$ we have that $d_S(u,c) - d_S(u,P) \leq \alpha\Delta$. Since v is on the shortest path from u to c, $d_S(v,c) = d_S(u,c) - d_S(u,v)$ and thus

$$d_S(v,c) - d_S(v,P) \leq (d_S(u,c) - d_S(u,v)) - (d_S(u,P) - d_S(u,v)) = d_S(u,c) - d_S(u,P) \leq \alpha\Delta \ ,$$

which implies that $v \in B$. □

Lemma 12. *Each cone B created in the algorithm has* $\mathrm{diam}(G[B]) \leq \Delta$.

Proof. Recall that each neighborhood S of a shortest path P contains points within distance at most $\Delta/4$ from P. Let S be the remaining part after some cones have been created, and P is the remaining path. The first property in Proposition 11 implies that the shortest path from any $u \in S$ to P is fully contained in S, and thus

$$d_S(u,P) \leq \Delta/4 \ . \tag{5.5}$$

Consider a certain cone B centered at $c \in P$, and by definition of B, for each $u \in B$,

$$d_S(u,c) \leq \alpha\Delta + d_S(u,P) \overset{(5.5)}{\leq} \Delta/4 + \Delta/4 = \Delta/2 \ . \tag{5.6}$$

By the second property of Proposition 11, if $u \in B$ then surely any $v \in S$ on the shortest path from u to c will also be in B, so $d_B(u, c) \leq \Delta/2$ as well, and thus $\mathrm{diam}(G[B]) \leq \Delta$.

Finally, it remains to see that `cut-cones` is indeed a partition of S (i.e. that it covers S), and this can be verified by the first property of Proposition 11. If for $u \in S$ there is a shortest path from u to P ending at $v \in P$, then whenever v is covered by a cone, u must be covered as well (the algorithm does not stop until $P = \emptyset$). ☐

For a time step i, we say that W is the *working supernode*, and at the end of this step it will become the supernode S_i. Note that W induces a connected subgraph, because we always choose a node u in $N(W)$ to be a start of the next path. We denote by $G_{i0} = G_i$. The following invariant holds for each time step i:

Invariant 2. *For every $i, j \geq 0$, every connected component C of G_{ij} satisfies that if $S, S' \in \mathcal{S}_{|C}$ then $S \sim S'$.*

Proof. Assume inductively that the invariant holds until time step i at iteration j. First consider the case $j > 0$, then as G_{ij} is obtained from $G_{i(j-1)}$ by removing some vertices, and the set of supernodes remains unchanged, the invariant will still hold: Every connected component C of G_{ij} is a subset of a connected component D of $G_{i(j-1)}$, in particular $\mathcal{S}_{|C} \subseteq \mathcal{S}_{|D}$, and so any pair of supernodes $S, S' \in \mathcal{S}_{|C}$ is also in $\mathcal{S}_{|D}$ and thus $S \sim S'$.

For the case $j = 0$, a new supernode S_{i-1} was just introduced, but the termination condition of line 9. guarantees that for any connected component C in G_i, any supernode $S \in \mathcal{S}_{|C}$ must have $S \sim S_{i-1}$. ☐

Corollary 13. *If G excludes K_{r+1} as a minor, then for every time step i and iteration j, the connected component C_{ij} has $|\mathcal{S}_{|C_{ij}}| \leq r$. Moreover, fix some $z \in V$. If $P_{i1}, \ldots, P_{il}$ are the shortest paths chosen while creating S_i in the components containing z, then $l \leq r$.*

Proof. If $|\mathcal{S}_{|C_{ij}}| = q$, then using Invariant 2, contracting each supernode in $\mathcal{S}_{|C_{ij}}$ will yield a K_q minor, so it must be that $q \leq r$. To see the second part of the assertion, note that each P_{ij} will connect the component containing z with some supernode $S \in \mathcal{S}_{|C_{ij}}$, so that $S_{ij} \sim S$. Finally, as $|\mathcal{S}_{|C_{ij}}| \leq r$, there can be at most r such paths. ☐

Lemma 14. *For $\gamma \leq 1/r^2$, the probability that a ball B_z of radius $\gamma\Delta$ is cut by the above process is*

$$\Pr[B_z \text{ cut}] \leq O(\gamma r^2) \ .$$

Proof. First observe that our algorithm is a threatening skeleton-process with parameters $l = 0$, $u = 1/4$, $s = r^2 + r$, $b = 2s$ and the G_i (respectively A_i, K_i) are the G_{ij} (resp. P_{ij}, S_{ij}) ordered lexicographically. By Invariant 2 we get that for all $i, j \in \mathbb{N}$, $|\mathcal{S}_{|C_{ij}}| \leq r$. By Corollary 13, each of these supernodes $S \in \mathcal{S}_{|C_{ij}}$ can have at most r paths that were built in a component containing C_{ij}, so it may contribute at most r to the number of sets in $\mathcal{K}_{|C_{ij}}$, to a total of r^2. We must also add in the (at most) r paths of the current working supernode, to obtain that $|\mathcal{K}_{|C_{ij}}| \leq s$. Recall that $\mathcal{T}_z = \{P_{ij} \mid d_{G_{ij}}(P_{ij}, z) \leq (u + \gamma)\Delta\}$, and let $\tau = \mathbb{E}[|\mathcal{T}_z|]$. With this we may apply Lemma 6 to infer that

$$\tau \leq 3e^{(2s+1)\cdot(1+\gamma/u)} \ .$$

Next, we show that our process is also a cutting skeleton-process, with the graph sequence G_{ij} and the skeletons are the P_{ij}, ordered lexicographically. The parameters are the same as before:

$l = 0$, $u = 1/4$ and $b = 2s$ (this is the exact same process, after all). The condition that the graph sequence contains every uncovered point is trivial by definition of G_{ij}. By Lemma 7 we obtain that the probability that B_z is cut is at most

$$(1 - e^{-2b\gamma/(u-l)}) \left(1 + \frac{\tau}{e^b - 1}\right) \leq (1 - e^{-20r^2\gamma}) \cdot (1 + 9e^{10r^2\gamma}) = O(\gamma r^2) \ , \tag{5.7}$$

where the last equality follows as $\gamma \leq 1/r^2$. In what follows we bound the probability of event $\mathcal{E}_{\text{cone}}$, which is the event that the ball B_z is cut in the `cut-cones` procedure, while conditioning that it was not cut while creating the S_{ij}. Let $S = S_{ij}$ be the set that contains B_z, which was built around the path $P = P_{ij}$. Let $c_1, \dots, c_k$ be the centers chosen in `cut-cones`(S,P). We claim that there can be at most 9 of them that may cut B_z. To see this, observe that each cone contains a ball of radius at least $\Delta/8$, and since P is a shortest path, in any set of 10 centers there are two centers c_g, c_h such that $d_S(c_g, c_h) \geq 9\Delta/8 > 2(1/2 + \gamma)\Delta$. By the triangle inequality it must be that at least one of them is more than $(1/2 + \gamma)\Delta$ away from z. Finally, by Lemma 12 any cone centered at c may only contain points at distance at most $\Delta/2$ from c (see (5.6)), so it may not be the first to cut B_z. As α is chosen uniformly from an interval of size $\Delta/8$, the probability that a ball of radius $\gamma\Delta$ will be cut is at most $2\gamma\Delta/(\Delta/8) = 16\gamma$. By a simple union bound,

$$\Pr[\mathcal{E}_{\text{cone}} \mid B_z \subseteq S] < 144\gamma \ ,$$

which is dominated by (5.7), thus the final bound is

$$\Pr[B_z \text{ is cut}] \leq O(\gamma r^2) \ .$$

□

6 Bounded Treewidth Graphs

Since graphs of treewidth r are K_{r+2}-free, the result of Section 4 already implies a (weak diameter) probabilistic partition which is $O(r)$-padded. The purpose of this section is to show a *strong diameter* $(O(r), O(1/r))$-padded partition for graphs of bounded treewidth. We will use the same framework as the previous sections, and exploit the special structure of bounded treewidth graphs.

Definition 15. *A graph $G = (V, E)$ has treewidth r if there exists a collection of sets $I = \{X_1, \dots, X_k\}$ with each $X_i \subseteq V$, and a tree $T = (I, F)$, such that the following conditions hold:*

- $\cup_{i \in [k]} X_i = V$*,*
- *For all $i \in [k]$, $|X_i| \leq r + 1$,*
- *For all $\{u, v\} \in E$, there exists $i \in [k]$ such that $u, v \in X_i$,*
- *For all $u \in V$, the tree nodes containing u form a connected subtree of T.*

Corollary 16. *Let U be a bag in the tree decomposition $T = (I, F)$ of $G = (V, E)$. Then if $U_1, U_2 \in I$ lie in different connected components of $T \setminus \{U\}$, and $x_1 \in U_1 \setminus U$, $x_2 \in U_2 \setminus U$, then x_1, x_2 are in different connected components of $G \setminus U$.*

6.1 The Algorithm

Let $G = (V, E)$ be a graph of treewidth $r - 1$, and let T be its tree decomposition, where T has an arbitrary root R. The height of a tree node U, $h(U)$, is its distance in T from the root R. For a vertex $u \in V$ let $h(v)$ denote the minimal height of a tree node U containing u, and denote by $b(u) = U$ the node achieving this minimum. Order the vertices of the graph $(v_1, \dots v_n)$ such that for all $1 \le i < j \le n$, $h(v_i) \le h(v_j)$. In the i-th iteration of the algorithm we will have a graph G_i (initially $G_1 = G$), and if $v_i \in G_i$ we shall create a cluster $S_i = B_{G_i}(v_i, R_i\Delta)$, where $R_i \sim \mathrm{Texp}_{[0,1/2]}(8r)$. Then set $G_{i+1} = G_i \setminus S_i$ and continue. If $v_i \notin G_i$ then we do nothing in this iteration.

Algorithm 5 `Treewidth-Partition`(G,Δ,r)

Let $G_1 \leftarrow G$.
Let $\mathcal{P} \leftarrow \emptyset$.
for $i = 1, \dots n$ **do**
  if $v_i \in G_i$ **then**
    Let $R_i \sim \mathrm{Texp}_{[0,1/2]}(8r)$.
    Let $S_i = B_{G_i}(v_i, R_i\Delta)$.
    Set $G_{i+1} \leftarrow G_i \setminus S_i$.
  else
    Set $G_{i+1} \leftarrow G_i$.
  end if
end for

6.2 The Analysis

Fix some $z \in V$, $\gamma = O(1/r)$ and $B_z = B_G(z, \gamma\Delta)$. Let $U = b(z) \in I$ be the tree node containing z such that $h(z) = h(U)$. The first observation is that when analyzing the probability that B_z is cut, we may restrict our attention to vertices $v \in V$ whose $b(v)$ lies on the path from R to U in T. The reason is that if $b(v_i)$ is not on this path, then if $C \in I$ is the least common ancestor of U and $b(v_i)$ in T, we claim that G_i does not contain any vertex from C. To see this, note that by the choice of ordering all vertices in C appear before v_i, and thus either created a cluster or were removed from the graph. By Corollary 16 z and v_i are in different component of G_i, so S_i cannot be the first to cut B_z.

Consider then the process restricted to the vertices contained in bags on the path from R to U (we may assume w.l.o.g that these appear first in the ordering). For any $i \in [n]$, denote by C_i the connected component in G_i that contains z, and let $\mathcal{S}_{|C_i} = \{S_j \mid S_j \sim C_i\}$.

Claim 17. *For any $i \in [n]$, $|\mathcal{S}_{|C_i}| \le 2r$.*

Proof. Let $R = U_1, \dots, U_k = U$ be the sequence of bags from the root to U in the tree decomposition. For any $j \in [k]$, let $i_j \in [n]$ be the minimal such that $U_j \cap V(G_{i_j}) = \emptyset$. We prove that $|\mathcal{S}_{|C_{i_j}}| \le r$, by noting that there are at most r supernodes that can intersect U_j (as $|U_j| \le r$). If a supernode S_h does not intersect U_j, then since this supernode is not centered at some vertex of $U_{j'}$ for $j' > j$ (using the ordering and the minimality of i_j), then by Corollary 16 there is no path from z to $N(S_h)$ in G_{i_j}. Since there are at most r new supernodes created between time i_j to i_{j+1} (as each bag is covered after at most r clusters are formed), the claim follows. $\square$

Observe that the algorithm generates a threatening skeleton-process with the sequence $G_1, \ldots,$ the skeletons are $A_i = \{v_i\}$, $K_i = S_i$, $l = 0$, $u = 1/2$, $s = 2r$ and $b = 4r$. Let $\mathcal{J}_z = \{v_i \mid d_{G_i}(z, v_i) \le (u+\gamma)\Delta\}$. By Claim 17 we may apply Lemma 6 and obtain that

$$\tau \le 3e^{(4r+1)\cdot(1+\gamma/u)} . \tag{6.8}$$

Finally, as our process can also be made to be a cutting skeleton-process, as long as we omit the steps in which $v_i \notin G_i$ (note that the next i for which $v_i \in G_i$ may depend on previous random choices of R_j for $j < i$, but this is allowed), and with $l = 0$, $u = 1/2$ and $b = 4r$. Applying Lemma 7, we obtain that the probability that B_z is cut is at most

$$(1 - e^{-2b\gamma})\left(1 + \frac{\tau}{e^b - 1}\right) \le (1 - e^{-8r\gamma}) \cdot 9e^{8r\gamma} = O(\gamma r),$$

using that $\gamma \le 1/r$.

7 Bounded Genus Graphs

For $g \ge 0$, a graph $G = (V, E)$ has genus at most g if it can be drawn on the surface of a sphere with g "handles" without any edges crossing. The following result is folklore (see, e.g., [IS07])

Lemma 18. *If G is a genus g graph, there exists a cycle A comprised of two shortest paths emanating at a common root, such that $G \setminus A$ has genus at most $g - 1$.*

This fits nicely in the bounded threateners program: Our algorithm will iteratively take such a cycle A, create a random buffer S around it, and recurse on the connected components of $G \setminus S$. The base case is when the component is planar, then we may apply our strong-diameter padding algorithm. Formally, in iteration i take a connected component C_i in G_i, if C_i is not planar, find a cycle A_i as in Lemma 18. Let $S_i = B_{G_i}(A_i, R_i\Delta)$ where $R_i \sim \mathrm{Texp}_{[0,1/4]}(8 \log g)$, set $G_{i+1} = G_i \setminus S_i$. Each S_i is partitioned to clusters by iteratively taking cones centered at some of the points of A_i. If C_i is planar, invoke the decomposition scheme of Section 5.

We now turn to analyzing the algorithm. The fact that the resulting partition is strong-diameter Δ-bounded follows from the fact that `Strong-Random-Partition` generates strong-diameter Δ-bounded clusters, and by Lemma 12, the cones are also strong-diameter Δ-bounded (the proof of that lemma never used that P is a shortest path, we only need that any point in S_i is within distance $\Delta/4$ from A_i).

Fix some $z \in V$, $\gamma \le \delta$ for sufficiently small constant δ (which is independent of g), and set $B_z = B_G(z, \gamma\Delta)$.

Lemma 19. *The probability that the ball B_z is cut by the above process is*

$$\Pr[B_z \text{ cut}] \le 1 - e^{-O(\gamma \log g)} .$$

Proof. Let $\mathcal{E}_{\text{genus}}$ be the event that B_z is first cut by some set S_i. Divide the event $\neg\mathcal{E}_{\text{genus}}$ into $\mathcal{F}_{\text{cone}} = \{\exists i,\ B_z \subseteq S_i\}$ and $\mathcal{F}_{\text{planar}} = \{\exists i,\ B_z \subseteq C_i \wedge C_i \text{ is planar}\}$. Let $\mathcal{E}_{\text{cone}}$ be the event that $\mathcal{F}_{\text{cone}}$ holds and also B_z is first cut by a cone in the `Create-Cones`(S_i, A_i), and finally let $\mathcal{E}_{\text{planar}}$ be the event that $\mathcal{F}_{\text{planar}}$ holds and also B_z is cut while calling `Strong-Random-Partition` on a planar component containing B_z. We will bound each of the $\mathcal{E}$ events separately.

Algorithm 6 `Genus-Partition`(G,Δ,g)

Let $G_0 \leftarrow G$, $i = 0$.
Let $P \leftarrow \emptyset$.
while G_i is non-empty **do**
  Let C_i be a connected component of G_i.
  if C_i is planar **then**
    Let P_i be a partition obtained by invoking `Strong-Random-Partition`($C_i, \Delta, 5$). Add the clusters of P_i to P.
    Set $G_{i+1} \leftarrow G_i \setminus \cup P_i$.
  else
    Let A_i be cycle as in Lemma 18.
    Let $R_i \sim \mathrm{Texp}_{[0,1/4]}(8 \log g)$.
    Add $S_i = B_{G_i}(A_i, R_i\Delta)$ as a cluster to P.
    `Create-Cones`(S_i, A_i). Add the resulting clusters to P.
    Set $G_{i+1} \leftarrow G_i \setminus S_i$.
  end if
  $i \leftarrow i + 1$.
end while

Assume w.l.o.g that non-planar components are chosen first, then the process until time T (where all components are planar) is a cutting skeleton-process, with the graph sequence $G_1, \dots,$ the skeletons A_i and $K_i = S_i$, the parameters are $l = 0$, $u = 1/2$ and $b = 2\log g$. Let $\mathcal{T}_z = \{A_i \mid i \in [T],\ d_{G_i}(A_i, z) \le (1/2+\gamma)\Delta\}$. Note that by Lemma 18 there can be at most g iterations (on components containing z) in which z lies in a non-planar component, so $|\mathcal{T}_z| \le g$. By Lemma 7

$$\Pr[\mathcal{E}_{\text{genus}}] \le (1 - e^{-8\gamma \log g}) \cdot (1 + g/(e^{2\log g} - 1)) \le 1 - e^{-16\gamma \log g} ,$$

using that $\gamma \le 1/20$, say. If $\Pr[\neg\mathcal{E}_{\text{genus}}] = p$, then $p \ge e^{-16\gamma \log g}$ and if $p_{\text{cone}} = \Pr[\mathcal{F}_{\text{cone}}]$ and $p_{\text{planar}} = \Pr[\mathcal{F}_{\text{planar}}]$ then

$$p = p_{\text{cone}} + p_{\text{planar}} \ . \tag{7.9}$$

By the first assertion of Theorem 4, there is a large constant C such that

$$\Pr[\mathcal{E}_{\text{planar}}] = p_{\text{planar}} \cdot O(\gamma) = p_{\text{planar}}(1 - e^{-C\gamma}) ,$$

since γ is sufficiently small.

Finally, we bound the probability of event $\mathcal{E}_{\text{cone}}$. Conditioning on $B_z \subseteq S_i$ for some i, we use a similar argument as in the proof of Lemma 14, here we claim that there can be at most 18 centers whose cone may intersect B_z. This is because if there are more, at least 10 of them lie on one of the two shortest path A_i is comprised of, and using the argument appearing in the proof of Lemma 14, it cannot be that all of them threaten B_z. Since α is chosen uniformly from an interval of length $\Delta/8$, the probability that any cone cuts B_z is at most $2\gamma\Delta/(\Delta/8)$, thus by a union bound, using that C is large enough,

$$\Pr[\mathcal{E}_{\text{cone}}] = p_{\text{cone}} \cdot O(\gamma) = p_{\text{cone}}(1 - e^{-C\gamma}) \ .$$

Combining the three bounds, we obtain that the probability that B_z is cut is at most

$$
\begin{aligned}
\Pr[\mathcal{E}_{\text{genus}}] + \Pr[\mathcal{E}_{\text{cone}}] + \Pr[\mathcal{E}_{\text{planar}}] &\leq 1 - p + p_{\text{cone}}(1 - e^{-C\gamma}) + p_{\text{planar}}(1 - e^{-C\gamma}) \\
&\overset{(7.9)}{=} 1 - p \cdot e^{-C\gamma} \\
&\leq 1 - e^{-16\gamma \log g} \cdot e^{-C\gamma} \\
&= 1 - e^{-O(\gamma \log g)} .
\end{aligned}
$$

□

We note that this result can also be extended to generate a sparse cover with strong diameter Δ, such that for each vertex, its ball of radius $O(\Delta)$ is contained in some cluster and each node belongs to at most $O(\log g)$ clusters.

Acknowledgments

We are grateful to Alex Andoni and Daniel Berend for fruitful discussions. A. Gupta thanks Microsoft Research SVC for their kind hospitality.

References

[AFH+04] Aaron Archer, Jittat Fakcharoenphol, Chris Harrelson, Robert Krauthgamer, Kunal Talwar, and Éva Tardos. Approximate classification via earthmover metrics. In *Proceedings of the 15th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1079–1087, New York, 2004. ACM.

[AGMW10] Ittai Abraham, Cyril Gavoille, Dahlia Malkhi, and Udi Wieder. Strong-diameter decompositions of minor free graphs. *Theory Comput. Syst.*, 47(4):837–855, 2010.

[And86] Thomas Andreae. On a pursuit game played on graphs for which a minor is excluded. *J. Combin. Theory Ser. B*, 41(1):37–47, 1986.

[AST90] Noga Alon, Paul Seymour, and Robin Thomas. A separator theorem for nonplanar graphs. *J. Amer. Math. Soc.*, 3(4):801–808, 1990.

[Awe85] Baruch Awerbuch. Complexity of network synchronization. *J. ACM*, 32(4):804–823, October 1985.

[Bar96] Y. Bartal. Probabilistic approximation of metric spaces and its algorithmic applications. In *Proceedings of the 37th Annual Symposium on Foundations of Computer Science*, FOCS '96, pages 184–, Washington, DC, USA, 1996. IEEE Computer Society.

[BLR10] Punyashloka Biswal, James R. Lee, and Satish Rao. Eigenvalue bounds, spectral partitioning, and metrical deformations via flows. *J. ACM*, 57(3), 2010.

[BLS10] Glencora Borradaile, James R. Lee, and Anastasios Sidiropoulos. Randomly removing g handles at once. *Comput. Geom.*, 43(8):655–662, 2010.

[BLT07] Costas Busch, Ryan LaFortune, and Srikanta Tirthapura. Improved sparse covers for graphs excluding a fixed minor. In *Proceedings of the twenty-sixth annual ACM symposium on Principles of distributed computing*, PODC '07, pages 61–70, New York, NY, USA, 2007. ACM.

[CKR05] Gruia Calinescu, Howard Karloff, and Yuval Rabani. Approximation algorithms for the 0-extension problem. *SIAM J. Comput.*, 34(2):358–372, 2004/05.

[Die00] Reinhard Diestel. *Graph theory*, volume 173 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 2000.

[FHL08] Uriel Feige, MohammadTaghi Hajiaghayi, and James R. Lee. Improved approximation algorithms for minimum weight vertex separators. *SIAM J. Comput.*, 38(2):629–657, 2008.

[FRT04] Jittat Fakcharoenphol, Satish Rao, and Kunal Talwar. A tight bound on approximating arbitrary metrics by tree metrics. *J. Comput. System Sci.*, 69(3):485–497, 2004.

[FT03] Jittat Fakcharoenphol and Kunal Talwar. An improved decomposition theorem for graphs excluding a fixed minor. *RANDOM-APPROX*, pages 36–46, 2003.

[GKL03] Anupam Gupta, Robert Krauthgamer, and James R. Lee. Bounded geometries, fractals, and low–distortion embeddings. In *FOCS*, pages 534–543, 2003.

[GS01] Geoffrey R. Grimmett and David R. Stirzaker. *Probability and random processes*. Oxford University Press, New York, third edition, 2001.

[IS07] Piotr Indyk and Anastasios Sidiropoulos. Probabilistic embeddings of bounded genus graphs into planar graphs. In *Symposium on Computational Geometry*, pages 204–209, 2007.

[KLPT09] Jonathan A. Kelner, James R. Lee, Gregory N. Price, and Shang-Hua Teng. Higher eigenvalues of graphs. In *FOCS*, pages 735–744, 2009.

[KPR93] Philip N. Klein, Serge A. Plotkin, and Satish Rao. Excluded minors, network decomposition, and multicommodity flow. In *STOC*, pages 682–690, 1993.

[Lee13] James R. Lee. Open question recap, February 2013. http://tcsmath.wordpress.com/2013/02/25/open-question-recap/.

[LGT12] James R. Lee, Shayan Oveis Gharan, and Luca Trevisan. Multi-way spectral partitioning and higher-order cheeger inequalities. In *STOC*, pages 1117–1130, 2012.

[LN05] James R. Lee and Assaf Naor. Extending Lipschitz functions via random metric partitions. *Invent. Math.*, 160(1):59–95, 2005.

[LS93] Nathan Linial and Michael Saks. Low diameter graph decompositions. *Combinatorica*, 13(4):441–454, 1993. (Preliminary version in *2nd SODA*, 1991).

[LS10] James R. Lee and Anastasios Sidiropoulos. Genus and the geometry of the cut graph. In *SODA*, pages 193–201, 2010.

[Mat02] Jiří Matoušek. *Lectures on discrete geometry*, volume 212 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2002.

[PRS94] Serge Plotkin, Satish Rao, and Warren D. Smith. Shallow excluded minors and improved graph decompositions. In *Proceedings of the Fifth Annual ACM-SIAM Symposium on Discrete Algorithms (Arlington, VA, 1994)*, pages 462–470, New York, 1994. ACM.

[Rab03] Yuri Rabinovich. On average distortion of embedding metrics into the line and into ℓ_1. In *Proceedings of the thirty-fifth ACM symposium on Theory of computing*, pages 456–462. ACM Press, 2003.

[Rao99] Satish B. Rao. Small distortion and volume preserving embeddings for planar and Euclidean metrics. In *SOCG*, pages 300–306, 1999.

[RS03] Neil Robertson and Paul D. Seymour. Graph minors. XVI. excluding a non-planar graph. *Journal of Combinatorial Theory, Series B*, 89(1):43 – 76, 2003.

[Sid10] Anastasios Sidiropoulos. Optimal stochastic planarization. In *FOCS*, pages 163–170, 2010.

[WN11] Christian Wulff-Nilsen. Separator theorems for minor-free and shallow minor-free graphs with applications. In *FOCS*, pages 37–46, 2011.

A Proof of Lemma 6

Fix any $i \in \mathbb{N}$. W.l.o.g., we may assume that the process always picks the set A_i in the connected component C_i of G_i that contains z. Let $\mathbf{x} = \mathbf{x}(i)$ be a vector of the "normalized distances" from z to $\mathcal{K}_{|C_i}$. More formally, if $\mathcal{K}_{|C_i} = \{K_{i_1}, \dots K_{i_l}\}$ (with $l \leq s$ by the assumption of the lemma), then for $j \in [l]$ define $x_j := \frac{d_{G_i \cup K_{i_j}}(z, A_{i_j}) - R_{i_j}\Delta}{u\Delta}$. Intuitively, x_j should have been the distance from z to K_{i_j}, normalized by $u\Delta$. Observe that $d_{G_i \cup K_{i_j}}(z, K_{i_j}) \geq d_{G_i \cup K_{i_j}}(z, A_{i_j}) - R_{i_j}\Delta$.

Define the potential function for the vector $\mathbf{x} := (x_1, \dots, x_l)$ as

$$\Phi(\mathbf{x}) = \sum_{j=1}^{l} e^{-(2s+1)\cdot x_j} \ . \tag{A.10}$$

We would like to analyze the change to $\mathbf{x}$ over time. Assume w.l.o.g that $x_1 \leq \cdots \leq x_l$. Let $h := \frac{d_{G_i}(z, A_i)}{u\Delta} \geq 0$ be the normalized distance of z from the set A_i, and let $y = h - R_i/u$. Observe that if $x_j \leq y$ then the shortest path from K_{i_j} to z is completely disjoint from K_i; indeed, if there was an intersection, then the distance from z to K_i would be smaller, a contradiction. We get that if j^* is the maximal index such that $x_{j^*} \leq y$, then the first j^* entries of $\mathbf{x}$ will not change. The new set K_i will always be in $\mathcal{K}_{|C_{i+1}}$ (recall that C_{i+1} is the component containing z in $G_{i+1} = G_i \setminus K_i$), so we have that the $j^* + 1$ entry in $\mathbf{x}(i+1)$ will be $x_{j^*+1} = y$. For $j^* < j \leq l$, it could be the case that K_i intersects the shortest path from K_{i_j} to z, in which case the distance may increase or K_{i_j} can even be disconnected from z. Note that if $l = s$, then it must be that at least one K_{i_j} is disconnected from z, because we assume that $|\mathcal{K}_{|C_{i+1}}| \leq s$.

Next we attempt to bound the expected change to the potential function Φ in any single step. To this end, it suffices to consider the worst scenario, in which all the K_{i_j} for $j^* < j \leq l$ become disconnected from z by K_i (in such a case the potential decreases the most). To this end, define the "filtered subsequence" $\mathbf{x} \downarrow y$ to be the sequence obtained by dropping all the coordinates of $\mathbf{x}$ which are strictly larger than y, and adding in y. [3]

Lemma 20. *Let Φ be a function defined as in (A.10). Fix any non-decreasing sequence $\mathbf{x}$ of length at most s, and a random variable $Y \sim \mathrm{Texp}(2s)$ as in (2.2). Then for any $h \geq 0$:*

$$\mathbb{E}_Y[\Phi(\mathbf{x} \downarrow (h - Y)) - \Phi(\mathbf{x})] \geq \frac{s(e-2) \cdot e^{-(2s+1)h}}{1 - e^{-2s}}.$$

Proof of Lemma 20. Let $b = 2s$ and $a = -(b+1)$. The increase of the potential due to the new coordinate $y = h - Y$ is $e^{a\cdot(h-Y)}$, so the expected gain is

$$\mathbb{E}[e^{a(h-Y)}] = e^{ah} \cdot \int_0^1 e^{-aw} f_{texp;b}(w)\, dw = e^{ah} \cdot \int_0^1 \frac{b}{1 - e^{-b}}\, e^{-(a+b)w}\, dw = \frac{b(e-1) \cdot e^{ah}}{1 - e^{-b}}. \tag{A.11}$$

Next we analyze the loss in Φ for the coordinates x_j that are dropped. Recall that a coordinate x_j is dropped exactly when $x_j > h - Y$. Since $Y \in [0,1]$ the only interesting case is when $x_j = h - \gamma$ for some $\gamma \in [0,1]$, which is dropped when $Y > \gamma$, so the maximum loss is

$$\max_{\gamma \in [0,1]} e^{a(h-\gamma)} \Pr[Y > \gamma] \quad = \quad e^{ah} \cdot \max_{\gamma \in [0,1]} e^{-a\gamma} \int_\gamma^1 f_{texp;b}(w)\, dw$$

[3] E.g., $(-0.4, -0.3, 0.7, 5, 6.9) \downarrow 1.42 = (-0.4, -0.3, 0.7, 1.42)$.

$$
\begin{aligned}
&= e^{ah} \cdot \max_{\gamma \in [0,1]} e^{-a\gamma} \frac{e^{-b\gamma} - e^{-b}}{1 - e^{-b}} \\
&= \frac{e^{ah}}{1 - e^{-b}} \cdot \max_{\gamma \in [0,1]} e^{\gamma} - e^{-a\gamma - b} \\
&\leq \frac{e^{ah}}{1 - e^{-b}} \cdot e \ .
\end{aligned}
$$

Since $\mathbf{x}$ has only s coordinates, the total expected loss incurred on Φ is thus at most

$$
\frac{s \cdot e \cdot e^{ah}}{1 - e^{-b}} \ . \tag{A.12}
$$

Finally using (A.11) and (A.12) we see that

$$
\begin{aligned}
\mathbb{E}[\Phi(\mathbf{x} \downarrow (h - Y)) - \Phi(\mathbf{x})] &\geq \frac{b(e-1) \cdot e^{ah}}{1 - e^{-b}} - \frac{s \cdot e^{ah+1}}{1 - e^{-b}} \\
&= \frac{(e(b-s) - b) \cdot e^{-(b+1)h}}{1 - e^{-b}}.
\end{aligned}
$$

□

Applying Lemma 20 with the vector $\mathbf{x}$ (whose length is indeed at most s) and $Y = R_i/u$, noting that $Y \sim \text{Texp}(2s)$ by a property of the exponential distribution mentioned in Section 2, we get that

$$
\mathbb{E}_Y[\Phi(\mathbf{x} \downarrow y) - \Phi(\mathbf{x})] \geq \frac{(e-2)s \cdot e^{-(2s+1)h}}{1 - e^{-2s}} \geq 2s \cdot e^{-(2s+1)h}/3 \ . \tag{A.13}
$$

It will be more convenient to analyze a slightly different potential function

$$
\Phi'(\mathbf{x}) = \begin{cases} 2s & \exists j, \ x_j \leq 0 \\ \Phi(\mathbf{x}) & \text{otherwise} \end{cases}
$$

We claim that (A.13) still holds for Φ'. This is because it is the same as Φ as long as $\mathbf{x} > 0$, and if $\mathbf{x}(i+1)$ is the first to have a non-positive coordinate, then $\Phi'(\mathbf{x}(i)) = \Phi(\mathbf{x}(i)) < s$ and $\Phi'(\mathbf{x}(i+1)) = 2s$, but since $h \geq 0$ the claimed expected increase in (A.13) is never more than s.

Denote by $\Phi_i = \Phi'(\mathbf{x}(i))$. Recall that for every $i \in \mathcal{J}_z$ we have that $d_{G_i}(z, A_i) \leq (u+\gamma)\Delta$ and thus $h = d_{G_i}(z, A_i)/(u\Delta) \leq 1 + \gamma/u$. Observe that the expectation of (A.13) is taken only over the current choice of Y, and since Y is chosen independently we can condition on any other event that depends on previous steps, and obtain the same bound. In particular, for $i \in \mathcal{J}_z$,

$$
\mathbb{E}[\Phi_{i+1} \mid \Phi_i] \geq \Phi_i + 2s \cdot e^{-(2s+1)h}/3 \geq \Phi_i + 2s \cdot e^{-(2s+1)\cdot(1+\gamma/u)}/3 \ , \tag{A.14}
$$

and define $\zeta = 2s \cdot e^{-(2s+1)\cdot(1+\gamma/u)}/3$. Also note that the bound of (A.13) is always positive, so even if $i \notin \mathcal{J}_z$ we still have

$$
\mathbb{E}[\Phi_{i+1} \mid \Phi_i] \geq \Phi_i \ . \tag{A.15}
$$

For $t \in \mathbb{N}$ let $j_t = |\{i \in \mathcal{J}_z \mid i \leq t\}|$ be the number of time steps until t in which z is threatened. We claim that the process $X_0, X_1, \ldots$ where $X_t = \Phi_t - \zeta \cdot j_t$, is a submartingale. To prove this consider two cases: If $t+1 \in \mathcal{J}_z$ then $j_{t+1} = j_t + 1$, and by (A.14) we get $\mathbb{E}[\Phi_{t+1} \mid \Phi_t, j_t] \geq \Phi_t + \zeta$, and so

$$\mathbb{E}[X_{t+1} \mid X_1, \ldots X_t] = \mathbb{E}[\Phi_{t+1} - \zeta \cdot j_{t+1} \mid \Phi_t, j_t] \geq \Phi_t + \zeta - \zeta \cdot j_{t+1} = \Phi_t + \zeta - \zeta \cdot (j_t + 1) = \Phi_t - \zeta \cdot j_t \ .$$

If it is the case that $t+1 \notin \mathcal{J}_z$, then $j_{t+1} = j_t$ and

$$\mathbb{E}[\Phi_{t+1} - \zeta \cdot j_{t+1} \mid \Phi_t, j_t] \geq \Phi_t - \zeta \cdot j_t \ .$$

The stopping time of a (sub)martingale $X_0, X_1, \ldots$ is a random variable τ that has support in $\mathbb{N}$, and such that the event $\tau = t$ depends only on $X_0, \ldots, X_t$. Define τ as the first time in which $\Phi_\tau = 2s$. Observe that if t is the time where $z \in K_t$, then it must be that $d_{G_t}(z, A_t) \leq R_t\Delta$, and so we get a non-positive coordinate in $\mathbf{x}(t)$ which implies that $\tau = t$. Since the stopping time is bounded by $|V|$ (there can be at most $|V|$ rounds, because at least one vertex is removed every round), we can apply Doob's optional stopping time Theorem [GS01, Section 12.5] and obtain that

$$\mathbb{E}[\Phi_\tau] - \zeta \cdot \mathbb{E}[j_\tau] = \mathbb{E}[X_\tau] \geq \mathbb{E}[X_0] = 0 \ .$$

Finally, as $\Phi_\tau = 2s$, we obtain that

$$\mathbb{E}[|\mathcal{J}_z|] = \mathbb{E}[j_\tau] \leq 2s/\zeta = 3e^{(2s+1)\cdot(1+\gamma/u)} \ .$$

This completes the proof.

B Proof of Lemma 7

Let us introduce some more notation and properties before proving this lemma. Define the following events:

$$\begin{aligned}
\mathcal{C}_i &= \{B_z \cap K_i \notin \{\emptyset, B_z\}\} && \text{“}B_z \text{ cut in round } i \text{ (by } A_i\text{)”},\\
\mathcal{F}_i &= \{B_z \cap K_i = \emptyset\} && \text{“}i \text{ was a } \texttt{no-op} \text{ round”},\\
\mathcal{E}_i &= \left\{\mathcal{C}_i \wedge \bigwedge_{j<i} \mathcal{F}_j\right\} && \text{“}B_z \text{ first cut in round } i\text{”}.
\end{aligned}$$

Denote by $\mathcal{F}_{(<i)}$ the event $\bigwedge_{j<i} \mathcal{F}_j$, so that $\mathcal{E}_i = (\mathcal{C}_i \wedge \mathcal{F}_{(<i)})$. Denote by $\overline{\mathcal{F}}_i$ the complement of $\mathcal{F}_i$. Observe that $\mathcal{C}_i$ (respectively $\overline{\mathcal{F}}_i$) implies that $i \in \mathcal{T}_z$, so

$$\Pr[\mathcal{C}_i] = \Pr[\mathcal{C}_i \wedge i \in \mathcal{T}_z] = \Pr[i \in \mathcal{T}_z] \cdot \Pr[\mathcal{C}_i \mid i \in \mathcal{T}_z] \ , \tag{B.16}$$

$$\Pr[\overline{\mathcal{F}}_i] = \Pr[i \in \mathcal{T}_z] \cdot \Pr[\overline{\mathcal{F}}_i \mid i \in \mathcal{T}_z] \ . \tag{B.17}$$

and the same holds also when conditioning on any other event. We have the following claim:

Claim 21. *For each $i \in \mathbb{N}$,*

$$\Pr[\mathcal{C}_i \mid \mathcal{F}_{(<i)}, i \in \mathcal{T}_z] \leq (1-\delta) \cdot \left(\Pr[\overline{\mathcal{F}}_i \mid \mathcal{F}_{(<i)}, i \in \mathcal{T}_z] + \frac{1}{e^b - 1}\right) \ .$$

Proof. Fix any graph G_i and any set $A_i \subseteq V(G_i)$ that agree with the conditioning on $\mathcal{F}_0, \ldots \mathcal{F}_{i-1}$ and so that $i \in \mathcal{T}_z$. Denote by $\rho = d_{G_i}(A_i, B_z)$, $\bar{b} = b/(u-l)$, and let $m = \max\{l, \rho\}$. Recall that R_i is chosen independently, so

$$\begin{aligned}
\Pr[\overline{\mathcal{F}}_i \mid \mathcal{F}_0, \ldots, \mathcal{F}_{i-1}, i \in \mathcal{T}_z, A_i] &= \int_m^u \frac{\bar{b}e^{-\bar{b}y}}{e^{-\bar{b}l} - e^{-\bar{b}u}} dy \\
&= \frac{e^{-\bar{b}m} - e^{-\bar{b}u}}{e^{-\bar{b}l} - e^{-\bar{b}u}} .
\end{aligned}$$

Since $\mathcal{F}_0, \ldots, \mathcal{F}_{i-1}$ occurred and $G_i \supseteq G_0 \setminus (\cup_{j<i} K_j)$, we have that $B_z \subseteq G_i$. Now if $R_i \geq \rho + 2\gamma$ then by the triangle inequality $B_z \subseteq K_i$, and the ball is "saved". This bounds the cut probability thus:

$$\begin{aligned}
\Pr[\mathcal{C}_i \mid \mathcal{F}_{(<i)}, i \in \mathcal{T}_z, A_i] &\leq \int_m^{\rho+2\gamma} \frac{\bar{b}e^{-\bar{b}y}}{e^{-\bar{b}l} - e^{-\bar{b}u}} dy \\
&\leq \frac{e^{-\bar{b}m} - e^{-\bar{b}(m+2\gamma)}}{e^{-\bar{b}l} - e^{-\bar{b}u}} \\
&= \frac{e^{-\bar{b}m}(1-\delta)}{e^{-\bar{b}l} - e^{-\bar{b}u}} \\
&= (1-\delta) \cdot \Pr[\overline{\mathcal{F}}_i \mid \mathcal{F}_{(<i)}, i \in \mathcal{T}_z, A_i] + (1-\delta) \frac{e^{-\bar{b}u}}{e^{-\bar{b}l} - e^{-\bar{b}u}} \\
&= (1-\delta) \cdot \left(\Pr[\overline{\mathcal{F}}_i \mid \mathcal{F}_{(<i)}, i \in \mathcal{T}_z, A_i] + \frac{1}{e^b - 1} \right) .
\end{aligned}$$

Finally, because the bound holds for any A_i, it holds without conditioning on it. □

Proof of Lemma 7. Observe that for each $i \in [N]$, the events $\left\{\overline{\mathcal{F}}_i \wedge \mathcal{F}_{(<i)}\right\}$ are pairwise disjoint (this is the event that B_z is either cut or contained in K_i for the first time), thus by the law of total probability,

$$\sum_{i \in \mathbb{N}} \Pr\left[\overline{\mathcal{F}}_i \wedge \mathcal{F}_{(<i)}\right] \leq 1 . \tag{B.18}$$

Also, by linearity of expectation

$$\tau = \sum_{i \in \mathbb{N}} \Pr[i \in \mathcal{T}_z] . \tag{B.19}$$

To bound the probability of the ball being cut, we start off with the trivial union bound:

$$\begin{aligned}
\Pr\left[\bigcup_{i \in \mathbb{N}} \mathcal{E}_i\right] &\leq \sum_i \Pr[\mathcal{E}_i] = \sum_i \Pr\left[\mathcal{C}_i \wedge \mathcal{F}_{(<i)}\right] \\
&= \sum_i \Pr[\mathcal{C}_i \mid \mathcal{F}_{(<i)}] \cdot \Pr[\mathcal{F}_{(<i)}] \\
&\overset{(B.16)}{=} \sum_i \Pr[\mathcal{C}_i \mid \mathcal{F}_{(<i)}, i \in \mathcal{T}_z] \cdot \Pr[i \in \mathcal{T}_z \mid \mathcal{F}_{(<i)}] \cdot \Pr[\mathcal{F}_{(<i)}] \\
&\overset{\text{Claim 21}}{\leq} \sum_i (1-\delta) \left(\Pr[\overline{\mathcal{F}}_i \mid \mathcal{F}_{(<i)}, i \in \mathcal{T}_z] + \frac{1}{e^b - 1} \right) \cdot \Pr[i \in \mathcal{T}_z \mid \mathcal{F}_{(<i)}] \cdot \Pr[\mathcal{F}_{(<i)}]
\end{aligned}$$

$$
\begin{aligned}
&\stackrel{(B.17)}{=} (1-\delta)\cdot\sum_i \Pr\left[\overline{\mathcal{F}}_i \wedge \mathcal{F}_{(<i)}\right] + \sum_i \Pr\left[i \in \mathcal{T}_z \wedge \mathcal{F}_{(<i)}\right]\cdot\frac{1-\delta}{e^b-1} \\
&\stackrel{(B.18)}{\leq} (1-\delta) + \frac{1-\delta}{e^b-1}\cdot\sum_i \Pr[i \in \mathcal{T}_z] \\
&\stackrel{(B.19)}{=} (1-\delta)\left(1+\frac{\tau}{e^b-1}\right) .
\end{aligned}
$$

This completes the proof. □