CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Dagstuhl Research Online Publication Server

On Strong Diameter Padded Decompositions

Arnold Filtser
Ben Gurion University of the Negev, Beersheva, Israel
arnold273@gmail.com

Abstract

Given a weighted graph $G = (V, E, w)$, a partition of V is Δ-bounded if the diameter of each cluster is bounded by Δ. A distribution over Δ-bounded partitions is a β-padded decomposition if every ball of radius $\gamma\Delta$ is contained in a single cluster with probability at least $e^{-\beta\cdot\gamma}$. The weak diameter of a cluster C is measured w.r.t. distances in G, while the strong diameter is measured w.r.t. distances in the induced graph $G[C]$. The decomposition is weak/strong according to the diameter guarantee.

Formerly, it was proven that K_r free graphs admit weak decompositions with padding parameter $O(r)$, while for strong decompositions only $O(r^2)$ padding parameter was known. Furthermore, for the case of a graph G, for which the induced shortest path metric d_G has doubling dimension ddim, a weak $O(\mathrm{ddim})$-padded decomposition was constructed, which is also known to be tight. For the case of strong diameter, nothing was known.

We construct strong $O(r)$-padded decompositions for K_r free graphs, matching the state of the art for weak decompositions. Similarly, for graphs with doubling dimension ddim we construct a strong $O(\mathrm{ddim})$-padded decomposition, which is also tight. We use this decomposition to construct $\left(O(\mathrm{ddim}), \tilde{O}(\mathrm{ddim})\right)$-sparse cover scheme for such graphs. Our new decompositions and cover have implications to approximating unique games, the construction of light and sparse spanners, and for path reporting distance oracles.

2012 ACM Subject Classification Theory of computation → Graph algorithms analysis; Theory of computation → Approximation algorithms analysis; Theory of computation → Random projections and metric embeddings

Keywords and phrases Padded decomposition, Strong Diameter, Sparse Cover, Doubling Dimension, Minor free graphs, Unique Games, Spanners, Distance Oracles

Digital Object Identifier 10.4230/LIPIcs.APPROX-RANDOM.2019.6

Category APPROX

Related Version `https://arxiv.org/abs/1906.09783`

Funding *Arnold Filtser*: Supported in part by ISF grant No. (1817/17) and by BSF grant No. 2015813.

Acknowledgements The author would like to thank Ofer Neiman for helpful discussions.

1 Introduction

Divide and conquer is a widely used algorithmic approach. In many distance related graph problems, it is often useful to randomly partition the vertices into clusters, such that small neighborhoods have high probability to be clustered together. Given a weighed graph $G = (V, E, w)$, a partitions is Δ-*bounded* if the diameter of every cluster is at most Δ. A distribution $\mathcal{D}$ over partitions is called a (β, δ, Δ)-*padded decomposition*, if every partition is Δ-bounded, and for every vertex $v \in V$ and $\gamma \in [0, \delta]$, the probability that the entire ball $B_G(v, \gamma\Delta)$ of radius $\gamma\Delta$ around v is clustered together, is at least $e^{-\beta\gamma}$. If G admits a (β, δ, Δ)-padded decomposition for every $\Delta > 0$, we say that G is (β, δ)-decomposable. If in addition $\delta = \Omega(1)$ is a universal constant, we say that G is β-decomposable. Among other

© Arnold Filtser;
licensed under Creative Commons License CC-BY

applications, padded decompositions have been used for multi-commodity flow [33], metric embeddings [41, 40, 34], edge and vertex cut problems [37, 24], routing [4], near linear SDD solvers [12], approximation algorithms [16], and many more.

The *weak* diameter of a cluster $C \subseteq V$ is the maximal distance between a pair of vertices in the cluster w.r.t. the shortest path metric in the entire graph G, i.e. $\max_{u,v\in C} d_G(u,v)$. The *strong* diameter is the maximal distance w.r.t. the shortest path metric in the induced graph $G[C]$, i.e. $\max_{u,v\in C} d_{G[C]}(u,v)$. Padded decomposition can be weak/strong according to the provided guarantee on the diameter of each cluster. It is considerably harder to construct padded decompositions with strong diameter. Nevertheless, strong diameter is more convenient to use, and some applications indeed require that (e.g. for routing, spanners etc.).

Previous results on padded decompositions are presented in Table 1. General n-vertex graphs are strongly $O(\log n)$-decomposable [10], which is also tight. In a seminal work, given a K_r free graph G, Klein, Plotkin and Rao [33] showed that G is weakly $O(r^3)$-decomposable. Fakcharoenphol and Talwar [23] improved the decomposability of K_r free graph to $O(r^2)$ (weak diameter). Finally, Abraham et al. [5] improved the decomposition parameter to $O(r)$, still with weak diameter. The first result on strong diameter for K_r free graphs is by Abraham et al. [6], who constructed decompositions with padding parameter exponential in r. In fact, they study a somewhat weaker notion of decomposition called separating decompositions (see Definition 16). Afterwards, in the same paper providing the state of the art for weak diameter, Abraham et al. [5] proved that K_r free graphs are strongly $(O(r^2),\Omega(\frac{1}{r^2}))$-decomposable. It was conjectured [5] that K_r free graphs are $O(\log r)$-decomposable. However, even improving strong diameter decompositions to match the state of the art of weak diameter remained elusive.

Another family of interest are graph with bounded doubling dimension[1]. Abraham, Bartal and Neiman [2] showed that a graph with doubling dimension ddim is weakly $O(\text{ddim})$-decomposable, generalizing a result from [29]. No prior strong diameter decomposition for this family is known.

A related notion to padded decompositions is *sparse cover*. A collection $\mathcal{C}$ of clusters is a (β,s,Δ)-sparse cover if it is strongly Δ-bounded, each ball of radius $\frac{\Delta}{\beta}$ is contained in some cluster, and each vertex belongs to at most s different clusters. A graph admits (β,s)-sparse cover scheme if it admits (β,s,Δ)-sparse cover for every $\Delta>0$. Awerbuch and Peleg [9] showed that for $k\in\mathbb{N}$, general n-vertex graphs admit a strong $(2k-1,2k\cdot n^{\frac{1}{k}})$-sparse cover scheme. For K_r free graphs, Abraham et al. [6] constructed $(O(r^2),2^r(r+1)!)$-sparse cover scheme. Busch, LaFortune and Tirthapura [15] constructed $(4,f(r)\cdot\log n)$-sparse cover scheme for K_r free graphs[2] .

For the case of graphs with doubling dimension ddim, Abraham et al. [4] constructed a $(2,4^{\text{ddim}})$-sparse cover scheme. No other tradeoff are known. In particular, if ddim is larger than $\log\log n$, the only way to get a sparse cover where each vertex belongs to $O(\log n)$ clusters is through [9], with only $O(\log n)$ padding.

1.1 Results and Organization

In our first result (Theorem 15 in Section 5), we prove that K_r free graphs are strongly $(O(r),\Omega(\frac{1}{r}))$-decomposable. Providing quadratic improvement compared to [5].

---

[1] A metric space (X,d) has doubling dimension ddim if every ball of radius $2r$ can be covered by 2^{ddim} balls of radius r. The doubling dimension of a graph is the doubling dimension of its induced shortest path metric.

[2] $f(r)$ is a function coming from the Robertson and Seymour structure theorem [42].

Table 1 Summery of all known and new padding decompositions for various graph families.

Family	Partition type	Padding	δ	Reference
Previous results				
General graphs	Strong	$O(\log n)$	$\Omega(1)$	[10]
Doubling	Weak	$O(\text{ddim})$	$\Omega(1)$	[29, 2]
K_r minor free	Weak	$O(r^3)$	$\Omega(1)$	[33]
K_r minor free	Weak	$O(r^2)$	$\Omega(1)$	[23]
K_r minor free	Weak	$O(r)$	$\Omega(1)$	[5]
K_r minor free	Strong	$\exp(r)$	$\exp(-r)$	[6] [3]
K_r minor free	Strong	$O(r^2)$	$\Omega(\frac{1}{r^2})$	[5]
Our results				
Doubling	Strong	$O(\text{ddim})$	$\Omega(1)$	Corollary 9
K_r minor free	Strong	$O(r)$	$\Omega(\frac{1}{r})$	Theorem 15

Our second result (Corollary 9 in Section 4) is the first strong diameter padded decompositions for doubling graphs, which is also asymptotically tight. Specifically, we prove that graphs with doubling dimension ddim are strongly $O(\text{ddim})$-decomposable.

Both of these padded decomposition constructions are based on a technical theorem (Theorem 4 in Section 3). Given a set of centers N, such that each vertex has a center at distance at most Δ and at most τ centers at distance at most 3Δ ($\forall v,\ |B_G(v,3\Delta)\cap N|\le\tau$), we construct a strong $(O(\log\tau),\Omega(1),4\Delta)$-padded decomposition. We also provide an alternative construction for the decomposition of Theorem 4 in Appendix B. All of our decompositions can be efficiently constructed in polynomial time. See Table 1 for a summery of results on padded decompositions.

Our third result (Theorem 10 in Section 4) is a sparse cover for doubling graphs. For every parameter $t\ge 1$, we construct an $\big(O(t),O(2^{\text{ddim}/t}\cdot\text{ddim}\cdot\log t)\big)$-sparse cover scheme. Note that for $t=1$ we (asymptotically) obtain the result of [6]. However, we also get the entire spectrum of padding parameters. In particular, for $t=\text{ddim}$ we get an $(O(\text{ddim}),O(\text{ddim}\cdot\log\text{ddim}))$-sparse cover scheme.

Next, we overview some of the previously known applications of strong diameter padded decomposition, and analyze the various improvements achieved using our results. Specifically:

1. Given an instance of the unique games problem where the input graph is K_r free, Alev and Lau [7] showed that if there exist an assignment that satisfies all but an ϵ-fraction of the edges, then there is an efficient algorithm that finds an assignment that satisfies all but an $O(r\cdot\sqrt{\epsilon})$-fraction. Using our padded decompositions for minor-free graphs, we can find an assignment that satisfies all but an $O(\sqrt{r\cdot\epsilon})$-fraction of the edges. See Section 6.1.
2. Using the framework of Filtser and Neiman [26], given an n vertex graph, with doubling dimension ddim, for every parameter $t>1$ we construct a graph-spanner with stretch $O(t)$, lightness $O(2^{\frac{\text{ddim}}{t}}\cdot t\cdot\log^2 n)$ and $O(n\cdot 2^{\frac{\text{ddim}}{t}}\cdot\log n\cdot\log\Lambda)$ edges [4]. The only previous spanner of this type appeared in [26], was based on weak diameter decompositions, had the same stretch and lightness, while having no bound whatsoever on the number of edges. See Section 6.2.

[3] In fact [6] studied separating decompositions instead of padded (see Definition 16).
[4] Lightness is the ratio between the weight of the spanner to the weight of the MST. $\Lambda=\max_{u,v\in V}d_G(u,v)/\min_{u,v\in V}d_G(u,v)$ is the aspect ratio.

3. Elkin, Neiman and Wulff-Nilsen [19] constructed a path reporting distance oracle for K_r free graphs with stretch $O(r^2)$, space $O(n \cdot \log \Lambda \cdot \log n)$ and query time $O(\log\log \Lambda)$. That is, on a query $\{u,v\}$ the distance oracle returns a $u-v$ path P of weight at most $O(r^2) \cdot d_G(u,v)$ in $O(|P| + \log\log \Lambda)$ time. Using our strong diameter padded decompositions we improve the stretch to $O(r)$, while keeping the other parameters intact. See Appendix A.
4. We further use the framework of [19] to create a path reporting distance oracle for graphs having doubling dimension ddim with stretch $O(\mathrm{ddim})$, space $O(n \cdot \mathrm{ddim} \log \Lambda)$ and query time $O(\log\log \Lambda)$. This is the first path reporting distance oracle for doubling graphs. The construction uses our sparse covers. See Appendix A.

1.2 Related Work

Other than padded decompositions, separating decompositions have been studied. Here, instead of analyzing the probability to cut a ball, we analyze the probability to cut an edge [8, 36, 16, 22]. Separating decompositions been used to minimize the number of inter-cluster edges in a partition. In particular, strong diameter version of such partitions were used for SDD solvers [12].

Miller et al. [38] constructed strong diameter partitions for general graphs, which they later used to construct spanners and hop-sets in parallel and distributed regimes (see also [18]). Hierarchical partitions with strong diameter had been studied and used for constructing distributions over spanning trees with small expected distortion [17, 1], Ramsey spanning trees [3] and for universal Steiner trees [14]. Another type of partitions studied is when we require only weak diameter, and in addition for each cluster to be connected [21, 25].

Padded decompositions were studied for additional graph families. Kamma and Krauthgamer [31] showed that treewidth r graphs are weakly $O(\log r + \log\log n)$-decomposable. Abraham et al. [5] showed that treewidth r graphs are strongly $O(\log r + \log\log n)$-decomposable and strongly $(O(r), \Omega(\frac{1}{r}))$-decomposable. [5] also showed that pathwidth r graphs are strongly $O(\log r)$- decomposable. Finally [5] proved that genus g graphs are strongly $O(\log g)$-decomposable, improving a previous weak diameter version of Lee and Sidiropoulos [35].

1.3 Technical Ideas

The basic approach for creating padded decompositions is by ball carving [10, 2]. That is, iteratively create clusters by taking a ball centered around some vertex, with radius drawn according to exponential distribution. The process halts when all the vertices are clustered. Intuitively, if every vertex might join the cluster associated with at most τ centers, the padding parameter is $O(\log \tau)$. We think of these centers as *threateners*. This approach worked very well for general graphs as the number of vertices is n. Similarly it also been used for doubling graphs, where the number of threateners is bounded by $2^{O(\mathrm{ddim})}$. However, in doubling graphs ball carving produces only weak diameter clustering.

Our main technical contribution is a proof that the intuition above holds for strong diameter as well. Specifically, we show that if there is a set N of centers such that each vertex has a center at distance at most Δ, and at most τ centers at distance 3Δ (these are the threateners), then the graph is strongly $(O(\log \tau), \Omega(1), 4\Delta)$-decomposable. We use the clustering approach of Miller et al. [38] with exponentially distributed starting times. In short, in [38] clustering each center x samples a starting time δ_x. Vertex v joins the cluster of the center x_i maximizing $\delta_x - d_G(x,v)$. This approach guaranteed to creates strong diameter clusters. The key observation is that if for every center $y \neq x_i$, $(\delta_{x_i} - d_G(x_i, v)) - (\delta_y - d_G(y, v)) \geq 2\gamma\Delta$, then the ball $B_G(v, \gamma\Delta)$ is fully contained in the cluster of x_i. Using truncated exponential distribution, we lower bound the probability of this event by $e^{-\gamma \cdot O(\log \tau)}$. It is the first time [38]-like algorithm is used to create padded decompositions.

In addition to the [38]-based algorithm, we also show a simpler algorithm, based on cone carving ([17]). The cone approach, although less involved, is inherently sequential and implies dependencies of each vertex on the entire center set. [38] algorithm can be efficiently implemented in distributed and parallel setting. Moreover, as each vertex depends only on centers in its local area, we are able to use the Lovász Local Lemma to create a sparse cover from padded decompositions.

Decompositions of K_r free graphs did not use ball carving directly. Rather, they tend to use the topological structure of the graph. We say that a cluster of G has an r-core with radius Δ if it contains at most r shortest paths (w.r.t. d_G) such that each vertex is at distance at most Δ from one of these paths. [5]'s strong decomposition for K_r free graphs is based on a partition into 1-core clusters, such that a ball with radius $\gamma\Delta$ is cut with probability at most $1 - e^{-O(\gamma r^2)}$. This partition is the reason for their $O(r^2)$ padding parameter. Although not stated explicitly, [5] also constructed a partition into r-core clusters, such that a ball with radius $\gamma\Delta$ is cut with probability at most $1 - e^{-O(\gamma r)}$. Apparently, [5] lacked an algorithm for partitioning r-clusters. Taking a union of Δ-nets from each shortest path to the center set N, it will follow that each vertex has at most $O(r)$ centers in its $O(\Delta)$ neighborhood. In particular, our theorem above implies a clustering of each r-core cluster into bounded diameter clusters. Our strong decomposition with parameter $O(r)$ follows.

2 Preliminaries

Graphs. We consider connected undirected graphs $G = (V, E)$ with edge weights $w : E \to \mathbb{R}_{\geq 0}$. We say that vertices v, u are neighbors if $\{v, u\} \in E$. Let d_G denote the shortest path metric in G. $B_G(v, r) = \{u \in V \mid d_G(v, u) \leq r\}$ is the ball of radius r around v. For a vertex $v \in V$ and a subset $A \subseteq V$, let $d_G(x, A) := \min_{a \in A} d_G(x, a)$, where $d_G(x, \emptyset) = \infty$. For a subset of vertices $A \subseteq V$, let $G[A]$ denote the induced graph on A, and let $G \setminus A := G[V \setminus A]$. The *diameter* of a graph G is $\mathrm{diam}(G) = \max_{v,u \in V} d_G(v, u)$, i.e. the maximal distance between a pair of vertices. Given a subset $A \subseteq V$, the *weak*-diameter of A is $\mathrm{diam}_G(A) = \max_{v,u \in A} d_G(v, u)$, i.e. the maximal distance between a pair of vertices in A, w.r.t. to d_G. The *strong*-diameter of A is $\mathrm{diam}(G[A])$, the diameter of the graph induced by A.

A graph H is a *minor* of a graph G if we can obtain H from G by edge deletions/contractions, and isolated vertex deletions. A graph family $\mathcal{G}$ is *H-minor-free* if no graph $G \in \mathcal{G}$ has H as a minor. Some examples of minor free graphs are planar graphs (K_5 and $K_{3,3}$ free), outer-planar graphs (K_4 and $K_{3,2}$ free), series-parallel graphs (K_4 free) and trees (K_3 free).

Doubling dimension. The doubling dimension of a metric space is a measure of its local "growth rate". A metric space (X, d) has doubling constant λ if for every $x \in X$ and radius $r > 0$, the ball $B(x, 2r)$ can be covered by λ balls of radius r. The doubling dimension is defined as $\mathrm{ddim} = \log_2 \lambda$. A d-dimensional ℓ_p space has $\mathrm{ddim} = \Theta(d)$, and every n point metric has $\mathrm{ddim} = O(\log n)$. We say that a weighted graph $G = (V, E, w)$ has doubling dimension ddim, if the corresponding shortest path metric (V, d_G) has doubling dimension ddim. The following lemma gives the standard packing property of doubling metrics (see, e.g., [29]).

▶ **Lemma 1** (Packing Property). *Let (X, d) be a metric space with doubling dimension* ddim*. If $S \subseteq X$ is a subset of points with minimum interpoint distance r that is contained in a ball of radius R, then $|S| \leq \left(\frac{2R}{r}\right)^{O(\mathrm{ddim})}$.*

Nets. A set $N \subseteq V$ is called a Δ-net, if for every vertex $v \in V$ there is a net point $x \in N$ at distance at most $d_G(v,x) \leq \Delta$, while every pair of net points $x, y \in N$, is farther than $d_G(x,y) > \Delta$. A Δ-net can be constructed efficiently in a greedy manner. In particular, by Lemma 1, given a Δ-net N in a graph of doubling dimension ddim, a ball of radius $R \geq \Delta$, will contain at most $\left(\frac{2R}{\Delta}\right)^{O(\mathrm{ddim})}$ net points.

Padded Decompositions and Sparse Covers. Consider a *partition* $\mathcal{P}$ of V into disjoint clusters. For $v \in V$, we denote by $P(v)$ the cluster $P \in \mathcal{P}$ that contains v. A partition $\mathcal{P}$ is strongly Δ-*bounded* (resp. weakly Δ-bounded) if the strong-diameter (resp. weak-diameter) of every $P \in \mathcal{P}$ is bounded by Δ. If the ball $B_G(v, \gamma\Delta)$ of radius $\gamma\Delta$ around a vertex v is fully contained in $P(v)$, we say that v is γ-*padded* by $\mathcal{P}$. Otherwise, if $B_G(v, \gamma\Delta) \not\subseteq P(v)$, we say that the ball is *cut* by the partition.

▶ **Definition 2** (Padded Decomposition). *Consider a weighted graph $G = (V, E, w)$. A distribution $\mathcal{D}$ over partitions of G is strongly (resp. weakly) (β, δ, Δ)-padded decomposition if every $\mathcal{P} \in \mathrm{supp}(\mathcal{D})$ is strongly (resp. weakly) Δ-bounded and for any $0 \leq \gamma \leq \delta$, and $z \in V$,*

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

We say that a graph G admits a strong (resp. weak) (β, δ)-padded decomposition scheme, if for every parameter $\Delta > 0$ it admits a strongly (resp. weakly) (β, δ, Δ)-padded decomposition that can be sampled in polynomial time.

A related notion to padded decompositions is sparse covers.

▶ **Definition 3** (Sparse Cover). *A collection of clusters $\mathcal{C} = \{C_1, ..., C_t\}$ is called a (β, s, Δ)-sparse cover if the following conditions hold.*
1. *Bounded diameter: The strong diameter of every $C_i \in \mathcal{C}$ is bounded by Δ.*
2. *Padding: For each $v \in V$, there exists a cluster $C_i \in \mathcal{C}$ such that $B_G(v, \frac{\Delta}{\beta}) \subseteq C_i$.*
3. *Overlap: For each $v \in V$, there are at most s clusters in $\mathcal{C}$ containing v.*

We say that a graph G admits a (β, s)-sparse cover scheme, if for every parameter $\Delta > 0$ it admits a (β, s, Δ)-sparse cover that can be constructed in expected polynomial time.

Truncated Exponential Distributions. To create padded decompositions, similarly to previous works, we will use truncated exponential distribution. That is, exponential distribution conditioned on the event that the outcome lays in a certain interval. The $[\theta_1, \theta_2]$-*truncated exponential distribution* with parameter λ is denoted by $\mathsf{Texp}_{[\theta_1,\theta_2]}(\lambda)$, and the density function is: $f(y) = \frac{\lambda\, e^{-\lambda \cdot y}}{e^{-\lambda \cdot \theta_1} - e^{-\lambda \cdot \theta_2}}$, for $y \in [\theta_1, \theta_2]$. For the $[0, 1]$-*truncated exponential distribution* we drop the subscripts and denote it by $\mathsf{Texp}(\lambda)$; the density function is $f(y) = \frac{\lambda\, e^{-\lambda \cdot y}}{1 - e^{-\lambda}}$.

3 Strongly Padded Decomposition

In this section we prove the main technical theorem of this paper.

▶ **Theorem 4.** *Let $G = (V, E, w)$ be a weighted graph and $\Delta > 0, \tau = \Omega(1)$ parameters. Suppose that we are given a set $N \subseteq V$ of center vertices such that for every $v \in V$:*
- Covering. *There is $x \in N$ such that $d_G(v, x) \leq \Delta$.*
- Packing. *There are at most τ vertices in N at distance 3Δ, i.e. $|B_G(v, 3\Delta) \cap N| \leq \tau$.*

Then G admits a strongly $\left(O(\ln\tau), \frac{1}{16}, 4\Delta\right)$-padded decomposition that can be efficiently sampled.

We start with description of the [38] algorithm (with some adaptations), and its properties. Later, in Section 3.2 we will prove Theorem 4. An alternative construction is given in Appendix B.

3.1 Clustering Algorithm Based on Starting Times

As we make some small adaptations, and the role of the clustering algorithm is essential, we provide full details. Let $\Delta > 0$ be some parameter and let $N \subseteq V$ be some set of centers such that for every $v \in V$, $d_G(v, N) \leq \Delta$. For each center $x \in N$, let $\delta_x \in [0, \Delta]$ be some parameter. The choice of $\{\delta_x\}_{x\in N}$ differs among different implementations of the algorithm. In our case we will sample δ_x using truncated exponential distribution. Each vertex v will join the cluster C_x of the center $x \in N$ for which the value $\delta_x - d_G(x, v)$ is maximized. Ties are broken in a consistent manner [5]. Note that it is possible that a center $x \in N$ will join the cluster of a different center $x' \in N$. An intuitive way to think about the clustering process is as follows: each center x wakes up at time $-\delta_x$ and begins to “spread” in a continuous manner. The spread of all centers done in the same unit tempo. A vertex v joins the cluster of the first center that reaches it.

▷ Claim 5. Every non-empty cluster C_x created by the algorithm has strong diameter at most 4Δ.

Proof. Consider a vertex $v \in C_x$. First we argue that $d_G(v, x) \leq 2\Delta$. This will already imply that C_x has weak diameter 4Δ. Let x_v be the closest center to v, then $d_G(v, x_v) \leq \Delta$. As v joined the cluster of x, it holds that $\delta_x - d_G(v, x) \geq \delta_{x_v} - d_G(v, x_v)$. In particular $d_G(v, x) \leq \delta_x + d_G(v, x_v) \leq 2\Delta$.

Let $\mathcal{I}$ be the shortest path in G from v to x. For every vertex $u \in \mathcal{I}$ and center $x' \in N$, it holds that

$$\begin{aligned}\delta(x) - d_G(u, x) = \delta(x) - (d_G(v, x) - d_G(v, u)) &\geq \delta(x') - d_G(v, x') + d_G(v, u)\\ &\geq \delta(x') - d_G(u, x') \ .\end{aligned}$$

We conclude that $\mathcal{I} \subseteq C_x$, in particular $d_{G[C_x]}(v, x) \leq 2\Delta$. The claim now follows. ◁

▷ Claim 6. Consider a vertex v, and let $x_1, x_2, \dots$ be an ordering of the centers w.r.t. $\delta(x_i) - d_G(v, x_i)$. That is $\delta(x_1) - d_G(v, x_1) \geq \delta(x_2) - d_G(v, x_2) \geq \dots$. Set $\Upsilon = (\delta(x_1) - d_G(v, x_1)) - (\delta(x_2) - d_G(v, x_2))$. Then for every vertex u such that $d_G(v, u) < \frac{\Upsilon}{2}$ it holds that $u \in C_{x_1}$.

Proof. For every center $x_i \neq x_1$ it holds that,

$$\delta(x_1) - d_G(u, x_1) > \delta(x_1) - d_G(v, x_1) - \frac{\Upsilon}{2} \geq \delta(x_i) - d_G(v, x_i) + \frac{\Upsilon}{2} > \delta(x_i) - d_G(u, x_i) \ .$$

In particular, $u \in C_{x_1}$. ◁

3.2 Proof of Theorem 4

For every center $x \in N$, we sample $\delta'_x \in [0, 1]$ according to $\mathsf{Texp}(\lambda)$ truncated exponential distribution with parameter $\lambda = 2 + 2\ln\tau$. Set $\delta_x = \delta'_x \cdot \Delta \in [0, \Delta]$. We execute the clustering algorithm from Section 3.1 with parameters $\{\delta_x\}_{x\in N}$ to get a partition $\mathcal{P}$.

[5] That is we have some order $x_1, x_2, \dots$. Among the centers x_i that minimize $\delta_{x_i} - d_G(x_i, v)$, v joins the cluster of the center with minimal index.

According to Claim 5, we created a distribution over strongly 4Δ-bounded partitions. Consider some vertex $v \in V$ and parameter $\gamma \le \frac{1}{4}$. We will argue that the ball $B = B_G(v, \gamma\Delta)$ is fully contained in $P(v)$ with probability at least $e^{-O(\gamma \log \tau)}$. Let N_v be the set of centers x for which there is non zero probability that C_x intersects B. Following the calculation in Claim 5, each vertex joins the cluster of a center at distance at most 2Δ. By triangle inequality, all the centers in N_v are at distance at most $(2+\gamma)\Delta \le 3\Delta$ from v. In particular $|N_v| \le \tau$.

Set $N_v = \{x_1, x_2, \dots\}$ ordered arbitrarily. Denote by $\mathcal{F}_i$ the event that v joins the cluster of x_i, i.e. $v \in C_{x_i}$. Denote by $\mathcal{C}_i$ the event that v joins the cluster of x_i, but not all of the vertices in B joined that cluster, that is $v \in C_{x_i} \cap B \ne B$. To prove the theorem, it is enough to show that $\Pr[\cup_i \mathcal{C}_i] \le 1 - e^{-O(\gamma\cdot\lambda)}$. Set $\alpha = e^{-2\gamma\cdot\lambda}$.

▷ Claim 7. For every i, $\Pr[\mathcal{C}_i] \le (1-\alpha)\left(\Pr[\mathcal{F}_i] + \frac{1}{e^{\lambda}-1}\right)$.

Proof. As the order in N_v is arbitrary, assume w.l.o.g. that $i = |N_v|$ and denote $x = x_{|N_v|}$, $\mathcal{C} = \mathcal{C}_i$, $\mathcal{F} = \mathcal{F}_i$, $\delta = \delta_{x_i}$ and $\delta' = \delta'_{x_i}$. Let $X \in [0,1]^{|N_v|-1}$ be the vector where the j'th coordinate equals δ'_{x_j}. Set $\rho_X = \frac{1}{\Delta}\cdot\left(d_G(x,v) + \max_{j<|N_v|}\left\{\delta_{x_j} - d_G(x_j,v)\right\}\right)$. Note that ρ_X is the minimal value of δ' such that if $\delta' > \rho_X$, that x has the maximal value $\delta_x - d_G(x,v)$, and therefor v will join the cluster of x. Note that it is possible that $\rho_X > 1$. Conditioning on the samples having values X, and assuming that $\rho_X \le 1$ it holds that

$$\Pr[\mathcal{F} \mid X] = \Pr[\delta' > \rho_X] = \int_{\rho_X}^{1} \frac{\lambda\cdot e^{-\lambda y}}{1-e^{-\lambda}}dy = \frac{e^{-\rho_X\cdot\lambda} - e^{-\lambda}}{1-e^{-\lambda}} .$$

If $\delta' > \rho_X + 2\gamma$ then $\delta - d_G(x,v) > \max_{j\ne i}\{\delta_{x_i} - d_G(x_i,v)\} + 2\gamma\Delta$. In particular, by Claim 6 the ball B will be contained in C_x. We conclude

$$\begin{aligned}\Pr[\mathcal{C} \mid X] &\le \Pr[\rho_X \le \delta' \le \rho_X + 2\gamma] \\ &= \int_{\rho_X}^{\max\{1,\rho_X+2\gamma\}} \frac{\lambda\cdot e^{-\lambda y}}{1-e^{-\lambda}}dy \\ &\le \frac{e^{-\rho_X\cdot\lambda} - e^{-(\rho_X+2\gamma)\cdot\lambda}}{1-e^{-\lambda}} \\ &= \left(1-e^{-2\gamma\cdot\lambda}\right)\cdot\frac{e^{-\rho_X\cdot\lambda}}{1-e^{-\lambda}} \\ &= (1-\alpha)\cdot\left(\Pr[\mathcal{F} \mid X] + \frac{1}{e^{\lambda}-1}\right) .\end{aligned}$$

Note that if $\rho_X > 1$ then $\Pr[\mathcal{C} \mid X] = 0 \le (1-\alpha)\cdot\left(\Pr[\mathcal{F} \mid X] + \frac{1}{e^{\lambda}-1}\right)$ as well. Denote by f the density function of the distribution over all possible values of X. Using the law of total probability, we can bound the probability that the cluster of x cuts B

$$\begin{aligned}\Pr[\mathcal{C}] &= \int_X \Pr[\mathcal{C} \mid X]\cdot f(X)\ dX \\ &\le (1-\alpha)\cdot\int_X \left(\Pr[\mathcal{F} \mid X] + \frac{1}{e^{\lambda}-1}\right)\cdot f(X)\ dX \\ &= (1-\alpha)\cdot\left(\Pr[\mathcal{F}] + \frac{1}{e^{\lambda}-1}\right)\end{aligned}$$ ◁

We bound the probability that the ball B is cut.

$$\begin{aligned}\Pr\left[\cup_i\mathcal{C}_i\right]=\sum_{i=1}^{|N_v|}\Pr\left[\mathcal{C}_i\right]&\le(1-\alpha)\cdot\sum_{i=1}^{|N_v|}\left(\Pr\left[\mathcal{F}_i\right]+\frac{1}{e^\lambda-1}\right)\\&\le\left(1-e^{-2\gamma\cdot\lambda}\right)\cdot\left(1+\frac{\tau}{e^\lambda-1}\right)\\&\le\left(1-e^{-2\gamma\cdot\lambda}\right)\cdot\left(1+e^{-2\gamma\cdot\lambda}\right)=1-e^{-4\gamma\cdot\lambda}\ ,\end{aligned}$$

where the last inequality follows as $e^{-2\gamma\lambda}=\frac{e^{-2\gamma\lambda}(e^\lambda-1)}{e^\lambda-1}\ge\frac{e^{-2\gamma\lambda}\cdot e^{\lambda-1}}{e^\lambda-1}\ge\frac{e^{\frac{\lambda}{2}-1}}{e^\lambda-1}=\frac{\tau}{e^\lambda-1}$.

▶ Remark 8. Actually we can prove a generalized version of Theorem 4. Suppose that there is a set N of centers such that each vertex $v\in V$ has at least one center at distance at most Δ and at most τ_v centers at distance 3Δ. Then for every parameter $\lambda=\Omega(1)$, there is a distribution over partitions with strong diameter 4Δ such that for every parameter $\gamma\in(0,\frac{1}{4})$, the ball around every vertex v of radius $\gamma\Delta$ is cut with probability at most $(1-e^{-2\gamma\lambda})(1+\frac{\tau_v}{e^\lambda-1})$.

4 Doubling Dimension

Our strongly padded decompositions for doubling graphs are a simple corollary of Theorem 4.

▶ **Corollary 9.** *Let $G=(V,E,w)$ be a weighted graph with doubling dimension* ddim*. Then G admits a strong $(O(\mathrm{ddim}),\Omega(1))$-padded decomposition scheme.*

Proof. Fix some $\Delta>0$. Let N be a Δ-net of X. According to Lemma 1, for every vertex v, the number of net points at distance 3Δ is bounded by $2^{O(\mathrm{ddim})}$. The corollary follows by Theorem 4. ◀

Next, we construct a sparse cover scheme.

▶ **Theorem 10.** *Let $G=(V,E,w)$ be a weighted graph with doubling dimension* ddim *and parameter $t=\Omega(1)$. Then G admits an $\left(O(t),O(2^{\mathrm{ddim}/t}\cdot\mathrm{ddim}\cdot\log t)\right)$-sparse cover scheme. In particular, there is an $(O(\mathrm{ddim}),O(\mathrm{ddim}\cdot\log\mathrm{ddim}))$-sparse cover scheme.*

Proof. Let $\Delta>0$ be the diameter parameter. Let $\alpha=\theta(1)$ be a constant to be determined later, set $\beta=\alpha\cdot t$. We will construct a $\left(\beta,O(2^{\mathrm{ddim}/t}\cdot\mathrm{ddim}\cdot\log t),4\Delta\right)$-sparse cover. As Δ is arbitrary, this will imply $\left(4\beta,O(2^{\mathrm{ddim}/t}\cdot\mathrm{ddim}\cdot\log t)\right)$-sparse cover scheme.

The sparse cover is constructed by sampling $O(2^{\mathrm{ddim}/t}\cdot\mathrm{ddim}\cdot\log t)$ independent partitions using Corollary 9 with diameter parameter Δ, and taking all the clusters from all the partitions to the cover. The sparsity and strong diameter properties are straightforward. To argue that each vertex is padded in some cluster we will use the constructive version of the Lovász Local Lemma by Moser and Tardos [39].

▶ **Lemma 11** (Constructive Lovász Local Lemma)**.** *Let $\mathcal{P}$ be a finite set of mutually independent random variables in a probability space. Let $\mathcal{A}$ be a set of events determined by these variables. For $A\in\mathcal{A}$ let $\Gamma(A)$ be a subset of $\mathcal{A}$ satisfying that A is independent from the collection of events $\mathcal{A}\setminus(\{A\}\cup\Gamma(A))$. If there exist an assignment of reals $x:\mathcal{A}\to(0,1)$ such that*

$$\forall A\in\mathcal{A}\ :\ \Pr[A]\le x(A)\Pi_{B\in\Gamma(A)}(1-x(B))\ ,$$

then there exists an assignment to the variables $\mathcal{P}$ not violating any of the events in $\mathcal{A}$. Moreover, there is an algorithm that finds such an assignment in expected time $\sum_{A\in\mathcal{A}}\frac{x(A)}{1-x(A)}\cdot$ poly $(|\mathcal{A}|+|\mathcal{P}|)$.

Formally, recall the construction of Theorem 4 used in Corollary 9. Let N be a Δ-net, that we will use as centers. Consider a vertex $v \in V$, and fix some sample of the starting times $\{\delta_x\}_{x\in N}$. Let x_v be the vertex maximizing $\delta_x - d_G(x,v)$ and y_v the second largest. In other words, $\delta_{x_v} - d_G(x_v,v) \geq \delta_{y_v} - d_G(y_v,v) \geq \max_{x\in N\setminus\{x_v,y_v\}}\{\delta_x - d_G(x,v)\}$. Let Ψ_v be the event that $(\delta_{x_v} - d_G(x_v,v)) - (\delta_{y_v} - d_G(y_v,v)) < 4\frac{\Delta}{\beta}$. Recall that the event that the ball of radius $2\frac{\Delta}{\beta}$ around v is cut contained in Ψ_v. Following the analysis of Theorem 4, $\Pr[\Psi_v] \leq 1 - e^{-O(\mathrm{ddim}\cdot 4\cdot\frac{\Delta}{\beta}/\Delta)} = 1 - 2^{-\mathrm{ddim}/t}$, where the equality follows by an appropriate choice of α.

Let $\hat{x}$ be the closest center to v. It holds that $\delta_{\hat{x}} - d_G(\hat{x},v) \geq -\Delta$, while for every center x at distance larger that 3Δ it holds that $\delta_x - d_G(x,v) \leq -2\Delta$. Therefore Ψ_v depends only on centers at distance at most 3Δ. In particular, by triangle inequality, if v and u are farther away than 6Δ, Ψ_v and Ψ_u are independent.

We take $m = \alpha_m \cdot 2^{\frac{\mathrm{ddim}}{t}} \cdot \mathrm{ddim}\cdot\log t$ independent partitions of X using Corollary 9, for $\alpha_m = \Theta(1)$ to be determined later. Denote by Ψ_v^i the event representing Ψ_v in the i'th partition. Let $\Phi_v = \bigwedge_{i=1}^m \Psi_v^i$ be the event that v "failed" in all the partitions. It holds that

$$\Pr[\Phi_v] \leq \left(1 - 2^{-\mathrm{ddim}/t}\right)^m \leq e^{-2^{-\mathrm{ddim}/t}\cdot m} = e^{-\alpha_m\cdot\mathrm{ddim}\cdot\log t}\ .$$

Note that if Ψ_v did not occurred, then the ball of radius $2\frac{\Delta}{\beta}$ around v was contained in a single cluster in at least one partition.

Let Y be an $\frac{\Delta}{\beta}$-net of X. Set $\mathcal{A} = \{\Phi_v\}_{v\in Y}$, to be a set of events determined by $\{\delta_x^i\}_{x\in N,1\leq i\leq m}$ (δ_x^i denotes δ_x in the i'th partition). Each event Φ_v might depend only on events Φ_u corresponding to vertices u at distance at most 6Δ from v. By Lemma 1, Φ_v is independent of all, but $\Gamma(\Phi_v) \leq \left(\frac{12\Delta}{\Delta/\beta}\right)^{O(\mathrm{ddim})} = 2^{O(\mathrm{ddim}\cdot\log t)}$ events. For every $\Phi_v \in \mathcal{A}$, set $x(\Phi_v) = p = 2^{-O(\mathrm{ddim}\cdot\log t)}$, such that $\max_{v\in Y}|\Gamma(v)| \leq \frac{1}{2p}$. Then, for every $\Phi_v \in \mathcal{A}$ it holds that,

$$x(\Phi_v)\cdot\Pi_{B\in\Gamma(\Phi_v)}(1 - x(B)) = p\cdot(1-p)^{|\Gamma(\Phi_v)|} \geq p\cdot(1-p)^{\frac{1}{2p}} \geq \frac{p}{e} \geq \Pr(\Phi_v)\ ,$$

where the last inequality holds for large enough α_m. By Lemma 11 we can efficiently find an assignment to $\{\delta_x^i\}_{x\in N,1\leq i\leq m}$ such that none of the events $\{\Phi_v\}_{v\in Y}$ occurred. Under this assumption, we argue that our sparse cover has the padding property. Consider some vertex $v \in V$. There is a net point $u \in Y$ at distance at most $\frac{\Delta}{\beta}$ from v. As the event Φ_u did not occur, there is some cluster C in the cover in which u is padded. In particular $B_G(v,\gamma\Delta) \subseteq B_G(u,2\gamma\Delta) \subseteq C$ as required.

Suppose that $|V| = n$, then the running time is $|Y|\cdot\frac{p}{1-p}\cdot\mathrm{poly}(|Y| + |Y|) = \mathrm{poly}(n)$. ◀

5 Minor Free Graphs

Our clustering algorithm is based on the clustering algorithm of [5], with a small modification. The clustering of [5] has two steps. In the first step the graph is partitioned into r-*Core* clusters (see Definition 12 bellow). While r-core clusters do not have bounded diameter, they do have a simple geometric structure. Moreover, this clustering also has the padding property for small balls. In the second step, each r-core cluster is partitioned into bounded diameter sub-clusters using Theorem 4.

▶ **Definition 12** (r-Core). *Given a weighted graph $G = (V, E, w)$, we say that G has an* r-core *with radius Δ, if there is a set of at most r shortest paths $\mathcal{I}_1, \dots, \mathcal{I}_{r'}$ such that for every $v \in V$, $d_G(v, \cup_i \mathcal{I}_i) \le \Delta$.*

Given a cluster $C \subseteq G$, we say that C is an r-core cluster with radius Δ, if $G[C]$ has an r-core *with radius Δ. Given a partition $\mathcal{P}$ of G, we say that it is an r-core partition with radius Δ if each cluster $C \in \mathcal{P}$, is an r-core cluster with radius Δ.*

The following theorem was proved implicitly in [5].

▶ **Lemma 13** (Core Clustering [5]). *Given a weighted graph $G = (V, E, w)$ that excludes K_r as a minor and a parameter $\Delta > 0$, there is a distribution $\mathcal{D}$ over r-core partitions with radius Δ, such that for every vertex $v \in V$ and $\gamma \in (0, \Omega(\frac{1}{r}))$ it holds that*

$$\Pr\left[B_G(v, \gamma\Delta) \subseteq P(v)\right] \ge e^{-O(r\cdot\gamma)} \ .$$

Even though we will not provide full details of the proof of Lemma 13, we will describe the algorithm itself and provide some intuition for the core clustering in Section 5.2. Our clustering algorithm will be executed in two steps: first we partition the graph into core clustering (Lemma 13) and then we partition each r-core cluster using Theorem 4.

Some historical notes: [5] presented two different algorithms for strong and weak padded decompositions. Each of these algorithms consisted of two steps. For weak decompositions, essentially they first partitioning the graph using r-core clustering. Secondly, instead of partition further each cluster, they pick a net from the r-cores in all the clusters, and iteratively grow balls around net points, ending with weak diameter guarantee. For strong decompositions, they partition the graph into 1-core clusters (instead of r-core), ending with a probability of only $e^{-O(r^2\cdot\gamma)}$ for a vertex x to be γ-padded.

5.1 Strong Padded Partitions for K_r Minor Free Graphs

▶ **Lemma 14.** *Let $G = (V, E, w)$ be a weighted graph that has an r-core with radius Δ. Then G admits a strong $(O(\log r), \Omega(1), \Delta)$-padded decomposition.*

Proof. Let $\mathcal{I}_1, \mathcal{I}_2, \dots, \mathcal{I}_{r'}$ be the r-core of G. For each i, let N_i be a $\frac{\Delta}{8}$-net of $\mathcal{I}_i$. Set $N = \cup_i N_i$. Every vertex $v \in V$ has some vertex in N at distance at most $\frac{\Delta}{4}$. Indeed, by definition of r-core, there is $x \in \mathcal{I}_i$ such that $d_G(v, x) \le \frac{\Delta}{8}$. Furthermore, there is a net point $y \in N_i$ at distance at most $\frac{\Delta}{8}$ from x. By triangle inequality $d_G(v, y) \le \frac{\Delta}{4}$. As $\mathcal{I}_i$ is a shortest path and N_i is a $\frac{\Delta}{8}$-net, there are at most $O(1)$ net points at distance $\frac{3}{4}\Delta$ from v in N_i. We conclude that in N there are at most $O(r)$ net points at distance $\frac{3}{4}\Delta$ from v. The lemma now follows by Theorem 4. ◀

▶ **Theorem 15.** *Let $G = (V, E, w)$ be a weighted graph that excludes K_r as a minor. Then G admits a strong $\left(O(r), \Omega(\frac{1}{r})\right)$-padded decomposition scheme.*

Proof. Let $\Delta > 0$ be some parameter. We construct the decomposition in two steps. First we sample an r-core partition $\mathcal{P}$ with radius parameter Δ using Lemma 13. Next, for every cluster $C \in \mathcal{P}$, we create a partition $\mathcal{P}_C$ using Lemma 14. The final partition is simply $\cup_{C\in\mathcal{P}}\mathcal{P}_C$, the union of all the clusters in all the created partitions. It is straightforward that the created partition has strong diameter Δ. To analyze the padding, consider a vertex $v \in V$ and parameter $0 < \gamma \le \Omega(\frac{1}{r})$. Denote by C_v the cluster containing v in $\mathcal{P}$, and by $P(v)$ the cluster of v in the final partition. Then,

$$\begin{aligned}\Pr\left[B_G(v,\gamma\Delta) \subseteq P(v)\right] &= \Pr\left[B_G(v,\gamma\Delta) \subseteq P(v) \mid B_G(v,\gamma\Delta) \subseteq C_v\right] \cdot \Pr\left[B_G(v,\gamma\Delta) \subseteq C_v\right] \\ &\ge e^{-O(\gamma\cdot r)} \cdot e^{-O(\gamma\cdot\log r)} = e^{-O(\gamma\cdot r)} \ ,\end{aligned}$$

where we used the fact that conditioning on $B_G(v, \gamma\Delta) \subseteq C_v$, it holds that $B_G(v, \gamma\Delta) = B_{G[C_v]}(v, \gamma\Delta)$. ◀

5.2 The Core Clustering Algorithm

In this section we describe the construction of the partition from Lemma 13. Afterwards, we will provide some intuition regarding the proof. For full details, we refer to [5]. Given two disjoint subsets $A, B \subseteq V$, we write $A \sim B$ if there exists an edge from a vertex in A to some vertex in B.

We denote the partition created by the algorithm by $\mathcal{S}$, and the clusters by $\{S_1, S_2, \dots\}$. The clusters are constructed iteratively. Initially $G_1 = G$. At step i, $G_i = G \setminus \cup_{j=1}^{i-1} S_j$. For a connected component $C \in G_i$, let $\mathcal{K}_{|C} = \{S_j \mid j < i \wedge C \sim S_j\}$ be the set of previously created clusters with a neighbor in C_i. To create S_i, pick arbitrary connected component C_i in G_i, and a vertex $x_i \in C_i$. For every neighboring cluster $S_j \in \mathcal{K}_{|C_i}$, pick arbitrary vertex $u_j \in C_i$ such that u_j has a neighbor in S_j. For each such u_j, let $\mathcal{I}_j$ be the shortest path in G_i from x_i to u_j. Let T_i be the tree created by the union of $\{\mathcal{I}_j\}_{S_j \in \mathcal{K}_{|C_i}}$ [6]. Sample a radius parameter R_i using truncated exponential distribution $\mathsf{Texp}_{[0,1]}(2r)$. The cluster S_i defined as $B_{G_i}(T_i, R_i\Delta)$, the set of all vertices at distance at most $R_i\Delta$ from T_i w.r.t. d_{G_i}. This finishes the construction of S_i. The algorithm halts when all the vertices are clustered. See pseudo-code in Algorithm 1. See also Figure 1 for illustration of the algorithm.

Algorithm 1 `Core-Partition`(G,Δ,r).

Let $G_1 \leftarrow G$, $i \leftarrow 1$.
Let $\mathcal{S} \leftarrow \emptyset$.
while G_i is non-empty **do**
  Let C_i be a connected component of G_i.
  Pick arbitrary $x_i \in C_i$. For each $S_j \in \mathcal{K}_{|C_i}$, let $u_j \in C_i$ be some vertex with a neighbor in S_j.
  Let T_i be a tree rooted at x_i, consisting of shortest paths towards $\{u_j \mid S_j \in \mathcal{K}_{|C_i}\}$.
  Let R_i be a random variable drawn independently from the distribution $\mathsf{Texp}_{[0,1]}(2r)$.
  Let $S_i \leftarrow B_{G_i}(T_i, R_i\Delta)$.
  Add S_i to $\mathcal{S}$.
  $G_{i+1} \leftarrow G_i \setminus S_i$.
  $i \leftarrow i + 1$.
end while
return $\mathcal{S}$.

Provided that the graph G excludes K_r as a minor, for every C_i it holds that $\left|\mathcal{K}_{|C_i}\right| \le r-2$. Indeed, by induction for every $S_j, S_{j'} \in \mathcal{K}_{|C_i}$, there is an edge between S_j to $S_{j'}$ [7]. Assume for contradiction that $\left|\mathcal{K}_{|C_i}\right| \ge r-1$. By contracting all the internal edges in C_i and in the clusters in $\mathcal{K}_{|C_i}$ we will obtain K_r as a minor, a contradiction. It follows that for every i, T_i is an r-core of S_i. In particular, Algorithm 1 indeed produces an r-core partitions with radius Δ.

Abraham et al. [5] called the core T_i of each cluster a *skeleton*. Their algorithm induce an iterative process that creates "skeletons" and removes their R_i neighborhoods (a buffer) from the graph. R_i was sampled according to truncated exponential distribution. They called such an algorithm a *threatening* skeleton-process. In general, they consider such a process where each R_i is drawn according to $\mathsf{Texp}_{[l,u]}(\frac{b}{u-l})$, for $0 = l < u \le 1$.

[6] Note that there is always a way to pick $\{\mathcal{I}_j\}_{S_j \in \mathcal{K}_{|C_i}}$ such that T_i will indeed be a tree.

[7] To see this note that there is a path between u_j to $u_{j'}$ in C_i. Therefore, when creating $S_{j'}$ (assuming $j < j'$), it was the case that $S_j \in \mathcal{K}_{|C_{j'}}$. In particular $T_{j'}$ contains a vertex with neighbor in S_j.

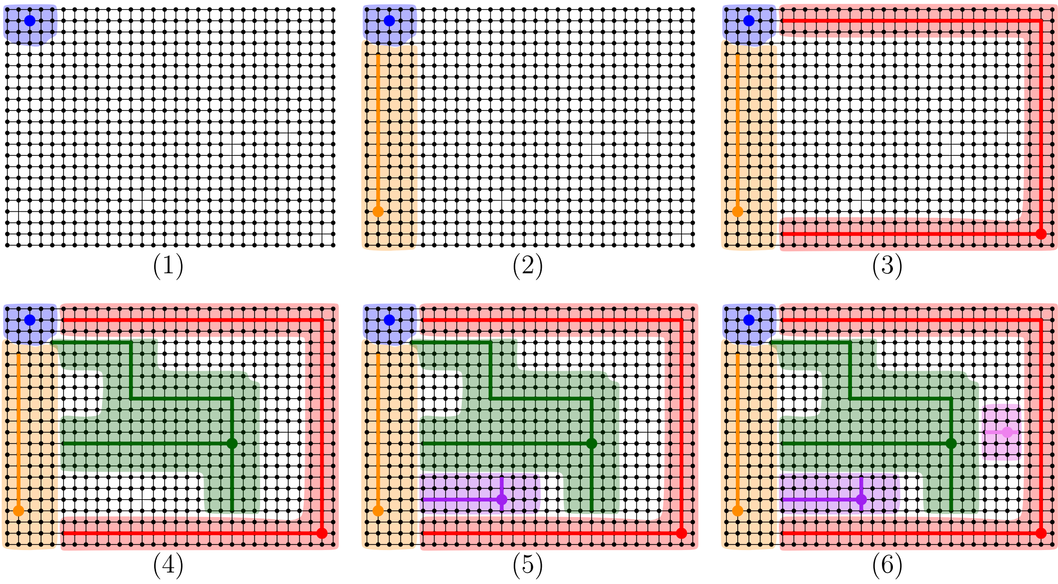

Figure 1 *The figure illustrates the 6 first steps in Algorithm 1. Here G is the (weighted) grid graph. Note that G excludes K_5 as a minor. In step (4), G_4 is the graph induced by all the vertices not colored in blue, orange or red. G_4 has a single connected component C_4. The green vertex defined as x_4. $\mathcal{K}_{|C_i}$ consist of 3 clusters S_1, S_2, S_3 colored respectively by blue, orange and red. T_4 is a tree rooted in x_4 colored in bold green, that consist of 3 shortest paths. Each of S_1, S_2, S_3 has a leaf of T_4 as a neighbor. R_4 is chosen according to $\mathsf{Texp}_{[0,1]}(10)$. The new cluster S_4, colored in green, consist of all vertices in C_4 at distance at most $R_4\Delta$ from T_4 w.r.t. d_{G_4}.*

Let $\gamma > 0$ be a padding parameter, fix some vertex $z \in V$ and set $B_z = B_G(z, \gamma\Delta)$. We say that a skeleton T_i threatens z if $d_{G_i}(z, T_i) \le (u+\gamma)\Delta$, in other words, if there is a positive probability that some vertex of B_z joins C_i. Let $\mathcal{J}_z = \{T_i \mid d_{G_i}(z, T_i) \le (u+\gamma)\Delta\}$ be the set of threatening skeletons. To bound the probability that B_z is cut, [5] first bound the expected number of threatening skeletons. A key lemma in [5] is that if we guaranteed that for every i, $|\mathcal{K}_{|C_i}| \le s$, and sample each radius R_i from $\mathsf{Texp}_{[l,u]}(\frac{b}{u-l})$ for $b = 2s$, it holds that

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

In a second key lemma, [5] argued that the probability that B_z is cut by a threatening skeleton-process, provided that $\tau = \mathbb{E}[|\mathcal{J}_z|]$, is at most

$$\left(1 - e^{-2b\gamma/(u-l)}\right)\left(1 + \frac{\tau}{e^b - 1}\right) .$$

In our case, as G is K_r free, thus we can pick $s = r - 2$. In Algorithm 1 we used the parameters $l = 0$, $u = 1$ and $b = 2r$. Therefore $\mathbb{E}[|\mathcal{J}_z|] \le 3e^{(2r+1)\cdot(1+\gamma)}$. Assuming that $\gamma = O(\frac{1}{r})$, we conclude that the probability that B_z is cut is at most

$$\left(1 - e^{-4r\gamma}\right)\left(1 + \frac{3e^{(2r+1)\cdot(1+\gamma)}}{e^{2r} - 1}\right) = O(r\gamma) .$$

In particular, the probability that B_z is padded is at least $1 - O(r\gamma) = e^{-O(r\gamma)}$.

6 Applications

In this section we present some applications of stochastic decompositions. Some applications are using a weaker type of decomposition called *separating* decompositions. The difference being that padding decompositions bound the probability for a ball to be cut, while separating decompositions bound the probability of an edge to be cut.

▶ **Definition 16** (Separating Decomposition). *A distribution $\mathcal{D}$ over partitions of a graph G is strongly (resp. weakly) (β, Δ)-separating decomposition if every $\mathcal{P} \in \mathrm{supp}(\mathcal{D})$ is strongly (resp. weakly) Δ-bounded and for every pair $u, v \in V$, $\Pr[P(v) \neq P(u)] \leq \beta \cdot \frac{d_G(u,v)}{\Delta}$.*

Note that in contrast to padding decomposition, there is no upper bound δ on the distance between u to v. Nevertheless, we argue that padded decompositions imply separating ones.

▶ **Lemma 17.** *Let $G = (V, E, w)$ be a weighted graph with a strongly (β, δ, Δ)-padded decomposition $\mathcal{D}$ such that $\delta \geq \frac{1}{\beta}$. Then $\mathcal{D}$ is also a strongly (β, Δ)-separating decomposition.*

Proof. Let $v, u \in V$ be a pair of vertices. If $d_G(u, v) \geq \frac{\Delta}{\beta}$, then obviously $\Pr[P(v) \neq P(u)] \leq 1 \leq \beta \cdot \frac{d_G(u,v)}{\Delta}$. Thus we can assume $d_G(u, v) \leq \frac{\Delta}{\beta} \leq \delta\Delta$. Set $\gamma = \frac{d_G(u,v)}{\Delta}$. It holds that

$$\Pr[P(v) = P(u)] \geq \Pr\left[B_G(v, \gamma\Delta) \subseteq P(v)\right] \geq e^{-\beta\gamma} \geq 1 - \beta\gamma \ .$$

In particular, $\Pr[P(v) \neq P(u)] \leq \beta\gamma = \beta \cdot \frac{d_G(u,v)}{\Delta}$ as required. ◀

Applying Lemma 17 on Corollary 9 and Theorem 15 we conclude,

▶ **Corollary 18.** *Let G be a weighted graph and $\Delta > 0$ some parameter.*

- *If G excludes K_r as a minor, it admits an efficient strongly $(O(r), \Delta)$-separating decomposition.*
- *If G has doubling dimension ddim, it admits an efficient strongly $(O(\mathrm{ddim}), \Delta)$-separating decomposition.*

6.1 Approximation for Unique Games on Minor Free Graphs

In the *Unique Games* problem we are give a graph $G = (V, E)$, an integer $k \geq 1$ and a set of permutations $\Pi = \{\pi_{uv}\}_{uv \in E}$ on $[k]$ satisfying $\pi_{uv} = \pi_{vu}^{-1}$. Given an assignment $x : V \to [k]$, the edge $uv \in E$ is satisfied if $\pi_{uv}(x(u)) = x(v)$. The problem is to find an assignment that maximizes the number of satisfied edges. The Unique Games Conjecture of Khot [32] postulates that it is NP-hard to distinguish whether a given instance of unique games is almost satisfiable or almost unsatisfiable. The unique games conjecture was thoroughly studied. The conjecture has numerous implications.

Alev and Lau [7] studied a special case of the unique games problem, where the graph G is K_r free. Given an instance (G, Π) where the optimal assignment violates ϵ-fraction of the edge constrains, Alev and Lau used an LP-based approach to efficiently find an assignment that violates at most $O(\sqrt{\epsilon} \cdot r)$-fraction. Specifically, in the rounding step of their LP, they used strong diameter separating decompositions with parameter $O(r^2)$. Using instead our decompositions from Corollary 18 with parameter $O(r)$ we obtain a quadratic improvement in the dependence on r.

▶ **Theorem 19.** *Consider an instance (G, Π) of the unique games problem, where the graph G is K_r free. Suppose that the optimal assignment violates at most an ϵ-fraction of the edge constrains. There is an efficient algorithm that find an assignment that violates at most an $O(\sqrt{\epsilon \cdot r})$-fraction.*

6.2 Spanner for Graphs with Moderate Doubling Dimension

Given a weighted graph $G = (V, E, w)$, a weighted graph $H = (V, E_H, w_H)$ is a *t-spanner* of G, if for every pair of vertices $v, u \in V$, $d_G(v, u) \leq d_H(v, u) \leq t \cdot d_X(v, u)$. If in addition H is a subgraph of G (that is $E_H \subseteq E$ and w_H agrees with w on E_H) then H is a *graph spanner*. The factor t is called the *stretch* of the spanner. The number of edges $|E_H|$ is the *sparsity* of the spanner. The weight of H is $w_H(H) = \sum_{e\in E_H} w_H(e)$ the sum of its edge weights. The *lightness* of H is $\frac{w_H(H)}{w(\mathrm{MST}(G))}$ the ratio between the weight of the spanner to the wight of the MST of G. The tradeoff between stretch and sparsity/lightness of spanners had been the focus of an intensive research effort, and low stretch graph spanners were used in a plethora of applications.

There is an extensive study of spanners for doubling metrics. Recently, for an n-vertex graph with doubling dimension ddim, Borradaile, Le and Wulff-Nilsen [13] contrasted a graph spanner with $1+\epsilon$ stretch, $\epsilon^{-O(\mathrm{ddim})}$ lightness and $n\cdot\epsilon^{-O(\mathrm{ddim})}$ sparsity (improving [43, 28, 27]). This result is also asymptotically tight. Note that the dependency on ddim is exponential, which is unavoidable for small, $1+\epsilon$ stretch. In cases where ddim is moderately large (say $\sqrt{\log n}$), it might be preferable to accept larger stretch in order to obtain reasonable lightness.

In a recent work, Filtser and Neiman [26], for every stretch parameter $t \geq 1$, constructed a spanner with stretch $O(t)$, lightness $O(2^{\frac{\mathrm{ddim}}{t}} \cdot t \cdot \log^2 n)$ and $O(n \cdot 2^{\frac{\mathrm{ddim}}{t}} \cdot \log n \cdot \log t)$ edges. However, this spanner was not a subgraph. Most applications require a graphic spanner. It is possible to transform [26] into a graphic spanner, but the number of edges becomes unbounded. The spanner construction of [26] is based on a variant of separating decompositions, where they used a weak-diameter version. If we replaced this with our strongly padded decompositions Corollary 9, and plug this into Theorem 3 from [26], we obtain a spanner with the same stretch to lightness ratio, but also with an additional sparsity guarantee.

▶ **Corollary 20.** *Let $G = (V, E, w)$ be an n vertex graph, with doubling dimension* ddim *and aspect ratio $\Lambda = \frac{\max_{e\in E} w(e)}{\min_{e\in E} w(e)}$. Then for every parameter $t > 1$ there is an graph-spanner of G with stretch $O(t)$, lightness $O(2^{\frac{\mathrm{ddim}}{t}} \cdot t \cdot \log^2 n)$ and $O(n \cdot 2^{\frac{\mathrm{ddim}}{t}} \cdot \log n \cdot \log \Lambda)$ edges.*

7 Conclusion and Open Problems

In this paper we closed the gap left in [5] between the padding parameters of strong and weak padded decompositions for minor free graphs. Our second contribution is tight strong padded decomposition scheme for graphs with doubling dimension ddim, which we also use to create sparse cover schemes. Some open questions remain:

1. Prove/disprove that K_r free graphs admit strong/weak decompositions with padding parameter $O(\log r)$, as conjectured by [5].
2. The question above is already open for the more restricted family of treewidth r graphs.
3. The δ parameter: [5] constructed weak $(O(r), \Omega(1))$-padded decomposition scheme, while we constructed strong $\left(O(r), \Omega(\frac{1}{r})\right)$-padded decomposition scheme. It will be nice to construct strong $(O(r), \Omega(1))$-padded decomposition scheme. Such a decomposition will imply a reacher spectrum of sparse covers (with $o(r)$ stretch).
4. Sparse covers for K_r free graphs: [6] constructed $(O(r^2), 2^r(r + 1)!)$-sparse cover scheme, while [15] constructed $(4, f(r) \cdot \log n)$-sparse cover scheme. An interesting open question is to create additional sparse cover schemes. Specifically, our padded decompositions suggest that an $(O(r), g(r))$-sparse cover scheme for some function g independent of n, should be possible. Currently it is unclear how to construct such a cover. Optimally, we would like to construct $(O(1), g(r))$-sparse cover scheme.

References

1 I. Abraham and O. Neiman. Using Petal-Decompositions to Build a Low Stretch Spanning Tree. *SIAM Journal on Computing*, 48(2):227–248, 2019. doi:10.1137/17M1115575.

2 Ittai Abraham, Yair Bartal, and Ofer Neiman. Advances in metric embedding theory. *Advances in Mathematics*, 228(6):3026–3126, 2011. doi:10.1016/j.aim.2011.08.003.

3 Ittai Abraham, Shiri Chechik, Michael Elkin, Arnold Filtser, and Ofer Neiman. Ramsey Spanning Trees and their Applications. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2018, New Orleans, LA, USA, January 7-10, 2018*, pages 1650–1664, 2018. doi:10.1137/1.9781611975031.108.

4 Ittai Abraham, Cyril Gavoille, Andrew V. Goldberg, and Dahlia Malkhi. Routing in Networks with Low Doubling Dimension. In *26th IEEE International Conference on Distributed Computing Systems (ICDCS 2006), 4-7 July 2006, Lisboa, Portugal*, page 75, 2006. doi:10.1109/ICDCS.2006.72.

5 Ittai Abraham, Cyril Gavoille, Anupam Gupta, Ofer Neiman, and Kunal Talwar. Cops, robbers, and threatening skeletons: padded decomposition for minor-free graphs. In *Symposium on Theory of Computing, STOC 2014, New York, NY, USA, May 31 - June 03, 2014*, pages 79–88, 2014. doi:10.1145/2591796.2591849.

6 Ittai Abraham, Cyril Gavoille, Dahlia Malkhi, and Udi Wieder. Strong-Diameter Decompositions of Minor Free Graphs. *Theory Comput. Syst.*, 47(4):837–855, 2010. doi:10.1007/s00224-010-9283-6.

7 Vedat Levi Alev and Lap Chi Lau. Approximating Unique Games Using Low Diameter Graph Decomposition. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2017, August 16-18, 2017, Berkeley, CA, USA*, pages 18:1–18:15, 2017. doi:10.4230/LIPIcs.APPROX-RANDOM.2017.18.

8 Baruch Awerbuch. Complexity of Network Synchronization. *J. ACM*, 32(4):804–823, 1985. doi:10.1145/4221.4227.

9 Baruch Awerbuch and David Peleg. Sparse Partitions (Extended Abstract). In *31st Annual Symposium on Foundations of Computer Science, St. Louis, Missouri, USA, October 22-24, 1990, Volume II*, pages 503–513, 1990. doi:10.1109/FSCS.1990.89571.

10 Yair Bartal. Probabilistic Approximations of Metric Spaces and Its Algorithmic Applications. In *37th Annual Symposium on Foundations of Computer Science, FOCS '96, Burlington, Vermont, USA, 14-16 October, 1996*, pages 184–193, 1996. doi:10.1109/SFCS.1996.548477.

11 Yair Bartal, Lee-Ad Gottlieb, Tsvi Kopelowitz, Moshe Lewenstein, and Liam Roditty. Fast, precise and dynamic distance queries. In *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2011, San Francisco, California, USA, January 23-25, 2011*, pages 840–853, 2011. doi:10.1137/1.9781611973082.66.

12 Guy E. Blelloch, Anupam Gupta, Ioannis Koutis, Gary L. Miller, Richard Peng, and Kanat Tangwongsan. Nearly-Linear Work Parallel SDD Solvers, Low-Diameter Decomposition, and Low-Stretch Subgraphs. *Theory Comput. Syst.*, 55(3):521–554, 2014. doi:10.1007/s00224-013-9444-5.

13 Glencora Borradaile, Hung Le, and Christian Wulff-Nilsen. Greedy spanners are optimal in doubling metrics. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2019, San Diego, California, USA, January 6-9, 2019*, pages 2371–2379, 2019. doi:10.1137/1.9781611975482.145.

14 Costas Busch, Chinmoy Dutta, Jaikumar Radhakrishnan, Rajmohan Rajaraman, and Srinivasagopalan Srivathsan. Split and Join: Strong Partitions and Universal Steiner Trees for Graphs. In *53rd Annual IEEE Symposium on Foundations of Computer Science, FOCS 2012, New Brunswick, NJ, USA, October 20-23, 2012*, pages 81–90, 2012. doi:10.1109/FOCS.2012.45.

15 Costas Busch, Ryan LaFortune, and Srikanta Tirthapura. Sparse Covers for Planar Graphs and Graphs that Exclude a Fixed Minor. *Algorithmica*, 69(3):658–684, 2014. doi:10.1007/s00453-013-9757-4.

16 Gruia Călinescu, Howard J. Karloff, and Yuval Rabani. Approximation Algorithms for the 0-Extension Problem. *SIAM J. Comput.*, 34(2):358–372, 2004. doi:10.1137/S0097539701395978.

17 Michael Elkin, Yuval Emek, Daniel A. Spielman, and Shang-Hua Teng. Lower-Stretch Spanning Trees. *SIAM J. Comput.*, 38(2):608–628, 2008. doi:10.1137/050641661.

18 Michael Elkin and Ofer Neiman. Efficient Algorithms for Constructing Very Sparse Spanners and Emulators. *ACM Trans. Algorithms*, 15(1):4:1–4:29, November 2018. doi:10.1145/3274651.

19 Michael Elkin, Ofer Neiman, and Christian Wulff-Nilsen. Space-efficient path-reporting approximate distance oracles. *Theor. Comput. Sci.*, 651:1–10, 2016. doi:10.1016/j.tcs.2016.07.038.

20 Michael Elkin and Seth Pettie. A Linear-Size Logarithmic Stretch Path-Reporting Distance Oracle for General Graphs. *ACM Trans. Algorithms*, 12(4):50:1–50:31, 2016. doi:10.1145/2888397.

21 Matthias Englert, Anupam Gupta, Robert Krauthgamer, Harald Räcke, Inbal Talgam-Cohen, and Kunal Talwar. Vertex Sparsifiers: New Results from Old Techniques. *SIAM J. Comput.*, 43(4):1239–1262, 2014. doi:10.1137/130908440.

22 Jittat Fakcharoenphol, Satish Rao, and Kunal Talwar. A tight bound on approximating arbitrary metrics by tree metrics. *J. Comput. Syst. Sci.*, 69(3):485–497, 2004. doi:10.1016/j.jcss.2004.04.011.

23 Jittat Fakcharoenphol and Kunal Talwar. An Improved Decomposition Theorem for Graphs Excluding a Fixed Minor. In *Approximation, Randomization, and Combinatorial Optimization: Algorithms and Techniques, 6th International Workshop on Approximation Algorithms for Combinatorial Optimization Problems, APPROX 2003 and 7th International Workshop on Randomization and Approximation Techniques in Computer Science, RANDOM 2003, Princeton, NJ, USA, August 24-26, 2003, Proceedings*, pages 36–46, 2003. doi:10.1007/978-3-540-45198-3_4.

24 Uriel Feige, MohammadTaghi Hajiaghayi, and James R. Lee. Improved Approximation Algorithms for Minimum Weight Vertex Separators. *SIAM J. Comput.*, 38(2):629–657, 2008. doi:10.1137/05064299X.

25 Arnold Filtser, Robert Krauthgamer, and Ohad Trabelsi. Relaxed Voronoi: A Simple Framework for Terminal-Clustering Problems. In *2nd Symposium on Simplicity in Algorithms, SOSA@SODA 2019, January 8-9, 2019 - San Diego, CA, USA*, pages 10:1–10:14, 2019. doi:10.4230/OASIcs.SOSA.2019.10.

26 Arnold Filtser and Ofer Neiman. Light Spanners for High Dimensional Norms via Stochastic Decompositions. In *26th Annual European Symposium on Algorithms, ESA 2018, August 20-22, 2018, Helsinki, Finland*, pages 29:1–29:15, 2018. doi:10.4230/LIPIcs.ESA.2018.29.

27 Arnold Filtser and Shay Solomon. The Greedy Spanner is Existentially Optimal. In *Proceedings of the 2016 ACM Symposium on Principles of Distributed Computing, PODC 2016, Chicago, IL, USA, July 25-28, 2016*, pages 9–17, 2016. doi:10.1145/2933057.2933114.

28 Lee-Ad Gottlieb. A Light Metric Spanner. In *IEEE 56th Annual Symposium on Foundations of Computer Science, FOCS 2015, Berkeley, CA, USA, 17-20 October, 2015*, pages 759–772, 2015. doi:10.1109/FOCS.2015.52.

29 Anupam Gupta, Robert Krauthgamer, and James R. Lee. Bounded Geometries, Fractals, and Low-Distortion Embeddings. In *44th Symposium on Foundations of Computer Science (FOCS 2003), 11-14 October 2003, Cambridge, MA, USA, Proceedings*, pages 534–543, 2003. doi:10.1109/SFCS.2003.1238226.

30 David S. Johnson. The NP-Completeness Column: An Ongoing Guide. *J. Algorithms*, 8(2):285–303, 1987. doi:10.1016/0196-6774(87)90043-5.

31 Lior Kamma and Robert Krauthgamer. Metric Decompositions of Path-Separable Graphs. *Algorithmica*, 79(3):645–653, 2017. doi:10.1007/s00453-016-0213-0.

32 Subhash Khot. On the power of unique 2-prover 1-round games. In *Proceedings on 34th Annual ACM Symposium on Theory of Computing, May 19-21, 2002, Montréal, Québec, Canada*, pages 767–775, 2002. doi:10.1145/509907.510017.

33 Philip N. Klein, Serge A. Plotkin, and Satish Rao. Excluded minors, network decomposition, and multicommodity flow. In *Proceedings of the Twenty-Fifth Annual ACM Symposium on Theory of Computing, May 16-18, 1993, San Diego, CA, USA*, pages 682–690, 1993. doi:10.1145/167088.167261.

34 Robert Krauthgamer, James R. Lee, Manor Mendel, and Assaf Naor. Measured Descent: A New Embedding Method for Finite Metrics. In *45th Symposium on Foundations of Computer Science (FOCS 2004), 17-19 October 2004, Rome, Italy, Proceedings*, pages 434–443, 2004. doi:10.1109/FOCS.2004.41.

35 James R. Lee and Anastasios Sidiropoulos. Genus and the Geometry of the Cut Graph. In *Proceedings of the Twenty-First Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2010, Austin, Texas, USA, January 17-19, 2010*, pages 193–201, 2010. doi:10.1137/1.9781611973075.18.

36 Nathan Linial and Michael E. Saks. Low diameter graph decompositions. *Combinatorica*, 13(4):441–454, 1993. doi:10.1007/BF01303516.

37 Jiri Matoušek. *Lectures on discrete geometry*. Springer-Verlag, New York, 2002. doi:10.1007/978-1-4613-0039-7.

38 Gary L. Miller, Richard Peng, Adrian Vladu, and Shen Chen Xu. Improved Parallel Algorithms for Spanners and Hopsets. In *Proceedings of the 27th ACM on Symposium on Parallelism in Algorithms and Architectures, SPAA 2015, Portland, OR, USA, June 13-15, 2015*, pages 192–201, 2015. doi:10.1145/2755573.2755574.

39 Robin A. Moser and Gábor Tardos. A constructive proof of the general lovász local lemma. *J. ACM*, 57(2):11:1–11:15, 2010. doi:10.1145/1667053.1667060.

40 Yuri Rabinovich. On average distortion of embedding metrics into the line and into L1. In *Proceedings of the 35th Annual ACM Symposium on Theory of Computing, June 9-11, 2003, San Diego, CA, USA*, pages 456–462, 2003. doi:10.1145/780542.780609.

41 Satish Rao. Small distortion and volume preserving embeddings for planar and Euclidean metrics. In *Proceedings of the fifteenth annual symposium on Computational geometry*, SCG '99, pages 300–306, New York, NY, USA, 1999. ACM. doi:10.1145/304893.304983.

42 Neil Robertson and Paul D. Seymour. Graph Minors. XVI. Excluding a non-planar graph. *J. Comb. Theory, Ser. B*, 89(1):43–76, 2003. doi:10.1016/S0095-8956(03)00042-X.

43 Michiel H. M. Smid. The Weak Gap Property in Metric Spaces of Bounded Doubling Dimension. In *Efficient Algorithms, Essays Dedicated to Kurt Mehlhorn on the Occasion of His 60th Birthday*, pages 275–289, 2009. doi:10.1007/978-3-642-03456-5_19.

A Path Reporting Distance Oracles

Given a weighted graph $G = (V, E, w)$, a *distance oracle* is a data structure that supports distance queries between pairs $u, v \in V$. The distance oracle has stretch t, if for every query $\{u, v\}$, the estimated distance $\text{est}(u, v)$ is within $d_G(u, v)$ and $t \cdot d_G(u, v)$. The studied objects are stretch, size the query time. An additional requirement that been recently studied [20] is *path reporting*: in addition to distance estimation, the distance oracle should also return a path of the promised length. In this case, we say that distance oracle has query time q, if answering a query when the reported path has m edges, takes $q + O(m)$ time.

Path reporting distance oracles were studied for general graphs [20, 19]. For the special case of graphs excluding K_r as a minor, Elkin, Neiman and Wulff-Nilsen [19] constructed a path reporting distance oracles with stretch $O(r^2)$, space $O(n \cdot \log \Lambda \cdot \log n)$ and query time $O(\log \log \Lambda)$, where $\Lambda = \frac{\max_{u,v} d_G(u,v)}{\min_{u,v} d_G(u,v)}$ is the aspect ratio. For this construction they used the strongly padded decomposition of [5] (in fact strong-diameter sparse covers). Implicitly, given a graph G that admits a strong (β, s)-sparse cover scheme, [19] constructs a path reporting distance oracle with stretch β, size $O(n \cdot s \cdot \log_\beta \Lambda)$ and query time $O(\log\log\Lambda)$. Following similar arguments to [19] [8], our padded decompositions from Theorem 15 implies that every K_r free graph admits a strong $(O(r), O(\log n))$-sparse cover scheme. We conclude:

▶ **Corollary 21.** *Given an n-vertex weighted graph $G = (V, E, w)$ which excludes K_r as a minor, with aspect ratio Λ , there is a path reporting distance oracle with stretch $O(r)$, space $O(n \cdot \log_r \Lambda \cdot \log n)$ and query time $O(\log\log\Lambda)$.*

It is interesting to mention that Busch et al. [15] constructed a $(4, O(f(r)\log n))$ sparse cover scheme for K_r free graphs, where $f(r)$ is an extremely large function of r. Using the framework of [19], it will imply a path reporting distance oracle with stretch 4, space $O(n \cdot \log\Lambda \cdot f(r))$ and query time $O(\log\log\Lambda)$. The value of $f(r)$ is larger that a square of the constant from the Robertson and Seymour structure theorem [42]. In particular, an estimation by Johnson [30] implies that $f(r)$ is larger than $2 \Uparrow (2 \Uparrow (2 \Uparrow (r/2)) + 3)$ [9]. This value is so big, that the [15]-based oracle is completely impractical already for quite small values of r.

For the case of graphs with doubling dimension ddim, we constructed the first strong-diameter sparse covers. Plugging our Theorem 10 into the framework of [19], we obtain the first path reporting distance oracle for doubling graphs. The only relevant previous distance oracle for doubling metrics is by Bartal et al. [11]. However, they focused on the $1+\epsilon$-stretch regime, where inherently the oracle size has exponential dependency on ddim.

▶ **Corollary 22.** *Given an n-vertex weighted graph $G = (V, E, w)$ with doubling dimension* ddim *and aspect ratio Λ, for every parameter $t \geq \Omega(1)$, there is a path reporting distance oracle with stretch $O(t)$, space $O(n \cdot 2^{\mathrm{ddim}/t} \cdot \mathrm{ddim} \cdot \log\Lambda)$ [10] and query time $O(\log\log\Lambda)$.*

In particular, there is a path reporting distance oracle with stretch $O(\mathrm{ddim})$, space $O(n \cdot \mathrm{ddim} \cdot \log\Lambda)$ and query time $O(\log\log\Lambda)$.

B Proof of Theorem 4 using Cones

We will prove a Theorem 4 with slightly weaker parameters. Specifically we will construct a strongly $\left(O(\ln\tau), \frac{1}{32}, 4\Delta\right)$-padded decomposition.

Order the vertices in $N = \{x_1, x_2, \dots\}$ arbitrarily. For every center $x_i \in N$, sample $\delta_i \in [0, 1]$ according to $\mathsf{Texp}(\lambda)$ truncated exponential distribution with parameter $\lambda = 2 + 2\ln\tau$. Set $R_i = \delta_i \cdot \Delta \in [0, \Delta]$. The clustering algorithm is executed in an iterative manner. We denote by S the set of unclustered vertices, which are also called *active* vertex. Initially $S = V$. As long as there is an active center $S \cap N \neq \emptyset$, pick active center $x_i \in N$ with minimal index and create the cluster

$$C_i = \left\{ v \in S \mid d_{G[S]}(v, x_i) - d_{G[S]}(v, N \cap S) \leq R_i \right\} .$$

This procedure halts when all the centers are clustered. See Algorithm 2 for pseudo code.

▷ Claim 23. For a vertex $v \in G$ let $x_v \in N$ be the closest center, and let $\mathcal{I}_v$ be the shortest path from v to x_v. Then if some vertex of $\mathcal{I}_v$ is clustered, so do v.

---

8 Taking $O(\log n)$ independent copies and using union bound,

9 $2 \Uparrow t$ denotes an exponential tower of t 2's. That is $2 \Uparrow 0 = 1$ and $2 \Uparrow t = 2^{2 \Uparrow (t-1)}$.

10 This is assuming $\Lambda > \log t$, otherwise simply using an arbitrary shortest path tree will provide a distance oracle with stretch $O(\log t)$.

Algorithm 2 `Partition-To-Cones`$(G=(V,E,w),N,\Delta,\tau)$.

Let $S \leftarrow V$, $\mathcal{S} \leftarrow \emptyset$.
Order the vertices in $N = x_1, x_2, \dots$ arbitrarily.
for $i = 1$ to $|N|$ **do**
  if $x_i \in S$ **then**
    Sample R_i independently from the distribution $\mathsf{Texp}(2 + 2\ln\tau)$.
    $C_i \leftarrow \emptyset$
    for all $v \in S$ **do**
      if $d_{G[S]}(v, x_i) - d_{G[S]}(v, N \cap S) \le R_i$ **then**
        Add v to C_i.
      end if
    end for
    $S \leftarrow S \setminus C_i$
    Add C_i to $\mathcal{S}$.
  end if
end for
return $\mathcal{S}$.

Proof. Suppose that $u \in \mathcal{I}_v$ joined the cluster of x_j while the set of active vertices were S (in particular $\mathcal{I}_v \subseteq S$). Then

$$\begin{aligned} d_{G[S]}(v, x_j) &\le d_{G[S]}(v, u) + d_{G[S]}(u, x_j) \\ &\le d_{G[S]}(v, u) + d_{G[S]}(u, x_v) + R_j = d_{G[S]}(v, x_v) + R_j \ . \end{aligned}$$ ◁

▶ **Corollary 24.** *All vertices are clustered.*

Proof. The vertex v will be clustered at the first time some vertex from $\mathcal{I}_v$ is clustered. As x_v itself necessarily clustered, the corollary follows. ◀

▷ Claim 25. Every cluster has strong diameter 4Δ.

Proof. Suppose that at the time we constructed C_i the set of active vertices was S. Let $v \in C_i$, and $x_v \in N$ the closest center to v. As v joined C_i and was active, all the vertices in $\mathcal{I}_v$ the shortest path from v to x_v were active as well. Therefore,

$$d_{G[S]}(v, x_i) \le d_{G[S]}(v, x_v) + R_i \le 2\Delta \ .$$

Let $\mathcal{I}$ be the shortest path from v to x_i in $G[S]$. We argue that all the vertices on $\mathcal{I}$ also joined C_i. Indeed, consider $u \in \mathcal{I}$. Then

$$\begin{aligned} d_{G[S]}(u, x_i) \ &= \ d_{G[S]}(v, x_i) - d_{G[S]}(v, u) \\ &\le \ d_{G[S]}(v, N \cap S) + R_i - d_{G[S]}(v, u) \ \le \ d_{G[S]}(u, N \cap S) + R_i \ . \end{aligned}$$

It follows that $d_{G[C_i]}(v, x_i) \le 2\Delta$. In particular C_i has strong diameter bounded by 4Δ. ◁

Consider some vertex $v \in V$ and parameter $\gamma \le \frac{1}{8}$. We will argue that the ball $B = B_G(v, \gamma\Delta)$ is fully contained in $P(v)$ with probability at least $2^{-O(\gamma \log \tau)}$, in other words that v is $\frac{\gamma}{4}$-padded. Let N_v be the set of centers x_i for which there is a non zero probability that C_i intersects B. Following the calculation in Claim 25, each vertex joins the cluster of a center at distance at most 2Δ. By triangle inequality, all the centers in N_v are at distance at most $(2+\gamma)\Delta \le 3\Delta$ from v. In particular $|N_v| \le \tau$.

For x_i, denote by $\mathcal{F}_i$ the event that some vertex of B joins the cluster C_i for the first time. I.e. $B\cap C_i\neq\emptyset$ and for all $j<i$, $B\cap C_j=\emptyset$. Denote by $\mathcal{C}_i$ the event that $\mathcal{F}_i$ occurred and B is cut by C_i. Note that for every $x_i\notin N_v$, $\mathcal{F}_i=\mathcal{C}_i=\emptyset$. To prove the theorem, it is enough to show that $\Pr\left[\cup_i\mathcal{C}_i\right]\le 1-e^{-O(\gamma\cdot\lambda)}$. Set $\alpha=e^{-4\gamma\cdot\lambda}$.

▷ Claim 26. For every i, $\Pr\left[\mathcal{C}_i\right]\le(1-\alpha)\left(\Pr\left[\mathcal{F}_i\right]+\frac{1}{e^{\lambda}-1}\right)$.

Proof. Let $S\subset V$ be the set of active vertices at the beginning of round i. If $B\cup\{x_i\}\not\subseteq S$ then $\Pr[\mathcal{C}_i]=0$ and we are done. Let ρ_S be the minimal value of δ_i such that if $\delta_i\ge\rho_S$, some vertex of B joins C_i. Formally $\rho_S=\frac{1}{\Delta}\cdot\min_{u\in B}\{d_{G[S]}(u,x_i)-d_{G[S]}(u,N\cap S)\}$. If $\rho_S>1$, then $\Pr[\mathcal{C}_i]=0$ and we are done, thus we assume $\rho_S\le 1$. Conditioning on S, it holds that

$$\Pr\left[\mathcal{F}_i\mid S\right]=\Pr\left[\delta_i\ge\rho_S\right]=\int_{\rho_S}^{1}\frac{\lambda\cdot e^{-\lambda y}}{1-e^{-\lambda}}dy=\frac{e^{-\rho_S\cdot\lambda}-e^{-\lambda}}{1-e^{-\lambda}}$$

Let $v'\in B$ some vertex that joins C_i if $\delta_i=\rho_S$. Then for every $u\in B$ it holds that

$$\begin{aligned}d_{G[S]}(u,x_i)&\le d_{G[S]}(v',x_i)+2\gamma\Delta\le d_{G[S]}(v',N\cap S)+\rho_S\cdot\Delta+2\gamma\Delta\\&\le d_{G[S]}(u,N\cap S)+(\rho_S+4\gamma)\cdot\Delta\ .\end{aligned}$$

Therefore, if $\delta_i\ge\rho_S+4\gamma$, the entire ball B will be contained in C_i. We conclude,

$$\begin{aligned}\Pr\left[\mathcal{C}_i\mid S\right]&\le\Pr\left[\rho_S\le\delta_i<\rho_S+4\gamma\right]\\&=\int_{\rho_S}^{\max\{1,\rho_S+4\gamma\}}\frac{\lambda\cdot e^{-\lambda y}}{1-e^{-\lambda}}dy\\&\le\frac{e^{-\rho_S\cdot\lambda}-e^{-(\rho_S+4\gamma)\cdot\lambda}}{1-e^{-\lambda}}\\&=\left(1-e^{-4\gamma\cdot\lambda}\right)\cdot\frac{e^{-\rho_S\cdot\lambda}}{1-e^{-\lambda}}\\&=(1-\alpha)\cdot\left(\Pr\left[\mathcal{F}_i\mid S\right]+\frac{1}{e^{\lambda}-1}\right)\ .\end{aligned}$$

By the low of total probability, we can remove the conditioning on S. Denote by f the density function of the distribution over all possible choices of S. It holds that,

$$\begin{aligned}\Pr\left[\mathcal{C}_i\right]&=\int_S\Pr\left[\mathcal{C}_i\mid S\right]\cdot f(S)\ dS\\&\le(1-\alpha)\cdot\int_S\left(\Pr\left[\mathcal{F}_i\mid S\right]+\frac{1}{e^{\lambda}-1}\right)\cdot f(S)\ dS\\&=(1-\alpha)\cdot\left(\Pr\left[\mathcal{F}_i\right]+\frac{1}{e^{\lambda}-1}\right)\ .\end{aligned}$$ ◁

We bound the probability that the ball B is cut,

$$\begin{aligned}\Pr\left[\cup_i\mathcal{C}_i\right]=\sum_{x_i\in N_v}\Pr\left[\mathcal{C}_i\right]&\le(1-\alpha)\cdot\sum_{x_i\in N_v}\left(\Pr\left[\mathcal{F}_i\right]+\frac{1}{e^{\lambda}-1}\right)\\&\le\left(1-e^{-4\gamma\cdot\lambda}\right)\cdot\left(1+\frac{\tau}{e^{\lambda}-1}\right)\\&\le\left(1-e^{-4\gamma\cdot\lambda}\right)\cdot\left(1+e^{-4\gamma\cdot\lambda}\right)=1-e^{-8\gamma\cdot\lambda}\ ,\end{aligned}$$

where the last inequality follows as $e^{-4\gamma\lambda}=\frac{e^{-4\gamma\lambda}\left(e^{\lambda}-1\right)}{e^{\lambda}-1}\ge\frac{e^{-4\gamma\lambda}\cdot e^{\lambda-1}}{e^{\lambda}-1}\ge\frac{e^{\frac{\lambda}{2}-1}}{e^{\lambda}-1}=\frac{\tau}{e^{\lambda}-1}$.