On Sparse Covers of Minor Free Graphs, Low Dimensional Metric Embeddings, and Other Applications

Arnold Filtser
Department of Computer Science, Bar-Ilan University, Ramat-Gan, Israel

Abstract

Given a metric space (X, d_X), a (β, s, Δ)-sparse cover is a collection of clusters $\mathcal{C} \subseteq P(X)$ with diameter at most Δ, such that for every point $x \in X$, the ball $B_X(x, \frac{\Delta}{\beta})$ is fully contained in some cluster $C \in \mathcal{C}$, and x belongs to at most s clusters in $\mathcal{C}$. Our main contribution is to show that the shortest path metric of every K_r-minor free graphs admits $(O(r), O(r^2), \Delta)$-sparse cover, and for every $\epsilon > 0$, $(4 + \epsilon, O(\frac{1}{\epsilon})^r, \Delta)$-sparse cover (for arbitrary $\Delta > 0$). We then use this sparse cover to show that every K_r-minor free graph embeds into $\ell_\infty^{\tilde{O}(\frac{1}{\epsilon})^{r+1} \cdot \log n}$ with distortion $3 + \epsilon$ (resp. into $\ell_\infty^{\tilde{O}(r^2) \cdot \log n}$ with distortion $O(r)$). Further, among other applications, this sparse cover immediately implies an algorithm for the oblivious buy-at-bulk problem in fixed minor free graphs with the tight approximation factor $O(\log n)$ (previously nothing beyond general graphs was known).

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

Keywords and phrases Sparse cover, minor free graphs, metric embeddings, ℓ_∞, oblivious buy-at-bulk

Digital Object Identifier 10.4230/LIPIcs.SoCG.2025.49

Related Version The reader is strongly encouraged to read the full version of the paper.
Full Version: `https://arxiv.org/abs/2401.14060` [43]

Funding This research was supported by the Israel Science Foundation (grant No. 1042/22).

1 Introduction

Given a metric space (X, d_X) a (β, s, Δ)-sparse cover is a collection of clusters $\mathcal{C} \subseteq P(X)$ all with diameter at most Δ, such that every point belongs to at most s clusters, and every ball $B_X(x, \frac{\Delta}{\beta})$ is fully contained in some cluster $C \in \mathcal{C}$. Sparse covers are very useful for algorithmic design, and in particular for divide and concur. Since their introduction by Awerbuch and Peleg [14], sparse covers found numerous applications. A partial list of which includes: compact routing schemes [81, 80, 86, 8, 7, 3, 23], distant-dependent distributed directories [15, 79, 80, 23], network synchronizers [11, 12, 73, 80, 23], distributed deadlock prevention [13], construction of spanners and ultrametric covers [55, 50, 70, 56, 48, 44], metric embeddings [82, 67], universal TSP and Stiner tree constructions [61, 22, 41, 24], and Oblivious buy-at-bulk [84].

We will study sparse covers in weighted graphs, where we distinguish between two different types of diameter. The *weak* diameter of a cluster $A \subseteq V$ in a weighted graph $G = (V, E, w)$ is the maximum pairwise distance $\max_{u,v \in A} d_G(u, v)$ w.r.t. the original shortest path distance, while the *strong* diameter is the maximum pairwise distance $\max_{u,v \in A} d_{G[A]}(u, v)$ in the induced subgraph. We continue with a formal definition:

▶ **Definition 1** (Sparse Cover). *Given a weighted graph $G = (V, E, w)$, a collection of clusters $\mathcal{C} = \{C_1, ..., C_t\}$ is called a weak/strong (β, s, Δ) sparse cover if the following conditions hold.*
1. *Bounded diameter: The weak/strong diameter of every cluster $C_i \in \mathcal{C}$ is bounded by Δ.*

41st International Symposium on Computational Geometry (SoCG 2025).
Editors: Oswin Aichholzer and Haitao Wang; Article No. 49; pp. 49:1–49:16
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

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. *Sparsity: For each $v \in V$, there are at most s clusters in $\mathcal{C}$ containing v.*

If the clusters $\mathcal{C}$ can be partitioned into s partitions $\mathcal{P}_1, \dots, \mathcal{P}_s$ s.t. $\mathcal{C} = \cup_{i=1}^{s} \mathcal{P}_i$, then $\{\mathcal{P}_1, \dots, \mathcal{P}_s\}$ is called a weak/strong (β, s, Δ) sparse partition cover. We say that a graph G admits a weak/strong (β, s) sparse (partition) cover scheme, if for every parameter $\Delta > 0$ it admits a weak/strong (β, s, Δ) sparse (partition) cover that can be constructed in expected polynomial time. Sparse partition cover scheme is abbreviated SPCS .

The notion of sparse cover scheme is the more common in the literature. However, SPCS provides additional structure that is crucial for different applications. [1] Obtaining strong diameter guarantee is also considerably more challenging, however it is frequently required in "subgraph based" applications. [2]

Sparse covers were introduced by Awerbuch and Peleg [14] who showed that for every $k \in \mathbb{N}$, every n vertex graph admits a strong $(4k-2, 2k \cdot n^{\frac{1}{k}})$-SPCS . Klein, Plotkin, and Rao [65] constructed a celebrated weak padded decomposition [3] for K_r-minor free graphs with padding parameter $O(r^3)$ (later improved to $O(r^2)$ [36]). It is folklore that padded decomposition with padding parameter ρ implies a $(\rho, O(\log n))$-SPCS (by taking the union of $O(\log n)$ independent partitions). Thus [65, 36] implied a weak $(O(r^2), O(\log n))$-SPCS for K_r-minor free graphs. Krauthgamer et al. [67] observed that one can use [65, 36] padded decomposition to construct a weak $(O(r^2), 2^r)$-SPCS for K_r-minor free graphs (see [41] for an explicit proof). This was the first sparse cover where all the parameters are independent from the cardinality of the vertex set n. Busch et al. [23] used shortest path separators to obtain a strong $(8, f(r) \cdot \log n)$-SPCS . [4] Abraham et al. [3] constructed a strong sparse cover. Specifically, they made some adaptations to [65] to obtain a strong $(O(r^2), 2^{O(r)} \cdot r!)$-sparse cover scheme (not a SPCS). The final major piece in our story is a weak padded decomposition for K_r-minor free graphs with padding parameter $O(r)$ by Abraham et al. [6]. Which was later improved to strong padded decomposition with the same parameter by Filtser [38]. In particular, this implies a strong $(O(r), O(\log n))$-sparse cover scheme. Interestingly, it was unknown how to turn this padded decomposition into a sparse cover with sparsity independent of n. Indeed, Filtser explicitly asked whether it is possible to construct $(O(r), g(r))$-sparse cover scheme [38] . The main result of this paper is an affirmative answer to this question. A decade after the publication of [2] we finally managed to turn this padded decomposition into a sparse cover.

▶ **Theorem 2** (Cover for Minor Free Graphs). *Every K_r-minor free graph admits the following:*

- *Strong $\big(O(r), O(r^2)\big)$-SPCS .*
- *For $\epsilon \in (0, \frac{1}{2})$, strong $\big(4+\epsilon, O(\frac{1}{\epsilon})^r\big)$-SPCS .*

The first bullet in Theorem 2 improves over the previous SPCS state of the art [36, 67, 41] in a threefold manner: (1) the padding is improved quadratically from $O(r^2)$ to $O(r)$, (2)

---

[1] In this paper we require the partition property of SPCS for metric embeddings into ℓ_∞, and for the oblivious buy-at-bulk. This property is also crucial in the construction of ultrametric covers [48, 42, 44].

[2] In this paper we use the strong diameter guarantee for routing, buy-at-bulk, and path reporting distance oracle.

[3] Roughly, padded decomposition with padding parameter ρ is a random partition into clusters of diameter at most Δ such that every ball of radius $\frac{\Delta}{\rho}$ is fully contained in a single cluster with probability at least $\frac{1}{2}$. See [38].

[4] $f(r)$ is the enormous constant hiding in the Robertson Seymour [83] structure theorem. Johnson [62] estimated $f(r) \geq 2 \Uparrow (2 \Uparrow (2 \Uparrow \frac{r}{2})) + 3)$ where $2 \Uparrow t$ is the exponential tower function ($2 \Uparrow 0 = 1$ and $2 \Uparrow t = 2^{2 \Uparrow (t-1)}$) .

Table 1 Summary of new and previous work on sparse covers.

Family	Padding	Sparsity	SPCS ?	Diameter	Ref
General	$4k-2$	$2k \cdot n^{\frac{1}{k}}$	yes	strong	[14]
Planar	32	18	yes (implicit)	strong	[23]
K_r-minor free	$O(r^3)$	$O(\log n)$	yes	weak	[65]
	$O(r^2)$	$O(\log n)$	yes	weak	[36]
	$O(r)$	$O(\log n)$	yes	weak	[6]
	$O(r)$	$O(\log n)$	yes	strong	[38]
	8	$f(r) \cdot \log n$ (4)	yes (implicit)	strong	[23]
	$O(r^2)$	2^r	yes	weak	[36, 67]
	$O(r^2)$	$2^{O(r)} \cdot r!$	no	strong	[3]
	$O(r)$	$O(r^2)$	yes	strong	Theorem 2
	$4+\epsilon$	$O(\frac{1}{\epsilon})^r$	yes	strong	

the sparsity is improved exponentially from 2^r to $O(r^2)$, (3) the diameter guarantee is now strong. Alternately, the second bullet improves the padding parameter from $O(r^2)$ to $O(1)$ while keeping a similar sparsity. See Table 1 for a comparison of ours and previous work.

1.1 Low dimensional metric embeddings into ℓ_∞

Metric embedding is a map between two metric spaces that approximately preserves pairwise distances. Given a (finite) metric space (X, d_X), a map $\phi : V \to \mathbb{R}^k$, and a norm $\|\cdot\|$, the *contraction* and *expansion* of the map ϕ are the smallest $\xi, \rho \geq 1$, respectively, such that for every pair $x, y \in X$,

$$\frac{1}{\xi} \cdot d(x,y) \leq \|\phi(x) - \phi(y)\| \leq \rho \cdot d(x,y) \quad .$$

The *distortion* of the map is then $\xi \cdot \rho$. The expansion ρ is also called the Lipschitz constant of the embedding ϕ. Metric embeddings into norm spaces were thoroughly studied [21, 72, 74, 4], and have a plethora of applications.

There is a special interest for metric embeddings into ℓ_∞. From an algorithmic viewpoint, there is a significant advantage in the additional structure the norm space is providing. One example being nearest neighbor search (NNS) [58, 19], where ℓ_∞ enjoys succinct data structures. NNS for many other spaces works by first embedding the space into ℓ_∞, and then using the ℓ_∞ NNS data structure to answer queries in the original metric space (see e.g. [37, 59]). From a geometric view point, ℓ_∞ is a special norm as it is a universal host metric space. That is, every finite metric spaces embeds isometrically (that is with distortion 1) into ℓ_∞ (the so called Fréchet embedding). However, there are n-point metric spaces of which every isometric embedding into ℓ_∞ requires $\Omega(n)$ dimensions [72]. It is desirable to have low dimensional metric embeddings, as these are much more useful for algorithmic design. Matoušek [74] showed that for every integer $t \geq 2$, every metric space embeds with distortion $2t-1$ into ℓ_∞ of dimension $O(n^{1/t} \cdot t \cdot \log n)$ (which is almost tight assuming the Erdős girth conjecture [34]). For distortion $O(\log n)$, Abraham et al. [4] later improved the dimension to $O(\log n)$.

For more restrictive metric spaces better results are known. Linial et al. [72] showed that every n-point tree metric embeds isometrically into $\ell_\infty^{O(\log n)}$, while Neiman [75] showed that every metric space with doubling dimension d embeds into $\ell_\infty^{\epsilon^{-O(d)} \cdot \log n}$ with distortion $1+\epsilon$. Krauthgamer et al. [67] showed that every n-point K_r-minor free graph G embeds

Table 2 Summary of new and previous work on metric embeddings into ℓ_∞.

Family	Distortion	Dimension	Ref
General Metric	1	$n-1$	Fréchet
	$2k-1$	$O(k \cdot n^{1/k} \cdot \log n)$	[74]
	$O(\log n)$	$O(\log n)$	[4]
Tree	1	$\Theta(\log n)$	[72]
K_r-Minor Free	$O(r^2)$	$\tilde{O}(3^r) \cdot \log n$	[67]
	$O(r)$	$\tilde{O}(r^2) \cdot \log n$	Corollary 3
	$3+\epsilon$	$\tilde{O}(\frac{1}{\epsilon})^{r+1} \cdot \log n$	Theorem 4

into $\ell_\infty^{\tilde{O}(3^r \cdot \log n)}$ with distortion $O(r^2)$. Their embedding follows from a SPCS based on [65, 36]. However, their construction uses additional properties of that cover, and we cannot simply plug in our SPCS to get an improved embedding. We show that under very general conditions,[5] one can use SPCS in a black box manner to obtain a metric embedding into ℓ_∞ (see full version [43]). As a corollary, we improve both the distortion and dimension compared to [67]. We than slightly tailor the embedding for minor free graphs to push the distortion down all the way to $3+\epsilon$. See Table 2 for a comparison of new and old results.

▶ **Corollary 3.** *Every n vertex K_r-minor free graph embeds into $\ell_\infty^{O(r^2 \cdot \log r \cdot \log n)}$ with distortion $O(r)$.*

▶ **Theorem 4.** *For every $\epsilon \in (0, \frac{1}{2})$, every n vertex K_r-minor free graph G can be embedded into $\ell_\infty^{O(\frac{1}{\epsilon})^{r+1} \cdot \log \frac{1}{\epsilon} \cdot \log \frac{n}{\epsilon}}$ with distortion $3+\epsilon$.*

1.2 Oblivious Buy-at-Bulk

Given a weighted graph $G = (V, E, w)$ and a *canonical fusion function* $f : \mathbb{N} \to \mathbb{R}_{\geq 0}$ (see full version [43] for a definition), in the oblivious buy-at-bulk problem, the goal is to pick a route P_i for every possible demand pair $\delta_i = (s_i, t_i) \in \binom{V}{2}$. Then, given a specific set of demands $A = \{\delta_1, \ldots, \delta_k\}$, the cost of our oblivious solution $\mathcal{P} = \{P_1, \ldots, P_k\}$ is $\mathrm{cost}(\mathcal{P}) = \sum_{e \in E} f(\varphi_e) \cdot w(e)$, where φ_e is the number of paths in $\mathcal{P}$ using e. The solution is said to have approximation ratio ρ if for every subset of demands, the induced cost of the oblivious solution is at most ρ times the optimal solution. Due to the concavity of the canonical fusion function f, it is advantageous for the chosen paths to intersect as much as possible. The best known approximation for general graphs is $O(\log^2 n)$ [53], while for planar graphs $O(\log n)$ approximation is known [84], which is tight [57]. Srinivasagopalan et al. [84] left it as an explicit open problem to "obtain efficient solutions to other related network topologies, such as minor-free graphs." More than a decade later, compared with general graphs, nothing better for K_r-minor free graphs is known. Using our sparse covers on top of a black box reduction from [84], we obtain the following tight result ([57]):

▶ **Corollary 5.** *For every n-vertex weighted K_r-minor free graph $G = (V, E, w)$ admits an efficiently commutable solution to the oblivious buy-at-bulk problem with approximation ratio $O(r^6 \cdot \log n)$. Furthermore, the solution is also oblivious to the concave faction f.*

[5] In fact the reduction hold in general with no condition, while the dimension slightly increases. See full version [43].

1.3 Further Applications

Sparse partition and universal TSP / Steiner tree

Given a weighted graph $G = (V, E, w)$, a (α, τ, Δ)-sparse partition is a partition $\mathcal{C}$ of V into clusters with weak diameter at most Δ, such that every ball of radius $\frac{\Delta}{\alpha}$ intersects at most τ clusters from $\mathcal{C}$. G admits a (α, τ)-sparse partition scheme if it admits (α, τ, Δ)-sparse partition for every $\Delta > 0$. Using our Theorem 2, we construct $\big(O(r), O(r^2)\big)$-sparse partition scheme for K_r minor free graphs, improving over the previous state of the art of $\big(O(r^2), 2^r\big)$-sparse partition scheme [61, 41]. For further details, see full version [43].

In the universal TSP problem, we are given a metric space (X, d_X) and the goal is to choose a single permutation π (a universal TSP) of X, such that given a subset $S \subseteq X$, we visit the points in S w.r.t. the order in π. This is the induced TSP tour by π. The permutation π has stretch ρ, if the length of the induced tour for every subset S is at most ρ times larger than the optimal tour for S. There is a general reduction from sparse partition scheme to universal TSP. Using this reduction and our sparse partitions, given a shortest path metric of an n point K_r minor free graph, we construct universal TSP with stretch $O(r^4) \cdot \log n$, exponentially improving the dependence on r compared with previous results $(O(1)^r \cdot \log n)$. The same phenomena occurs also for the universal Steiner tree problem. See full version [43] for further details.

Name Independent Routing

Here we are given an unweighted graph $G = (V, E)$ where the names (and ports) of all nodes are fixed. The goal is to design a compact routing scheme that will allow sending packages in the network, where routing decisions are made using small local routing tables, and the resulting routing paths are approximate shortest paths. This regime is considered more challenging and practical from the regime where nodes names and ports could be chosen by the routing scheme designer. Given a hereditary graph family that has α-orientation (see full version [43] for a definition), where each graph possess strong (τ, β)-sparse cover scheme, Abraham et al. [3] constructed name independent compact routing scheme with stretch $O(\beta)$, $O(\log n + \log \tau)$-bit headers and tables of $O(\frac{\log^3 n}{\log\log n} + \alpha \cdot \log n) \cdot \tau \cdot \log D$ bits. We use our strong sparse covers (Theorem 2) to construct name independent compact routing scheme significantly improving over previous work. See full version [43] for a comparison.

Path reporting distance oracle

A path reporting distance oracle (PRDO) for a weighted graph $G = (V, E, w)$ is a succinct data structure that given a query $\{x, y\}$, efficiently returns an approximate x-y shortest path P. We say that a PRDO has stretch k and query time t, if for every query (x, y), the oracle returns a path P of weight at most $k \cdot d_G(x, y)$ in $O(|P|) + t$ time. PRDO were first studied for general graphs. Later, Elkin et al. [33] constructed a PRDO for K_r-minor free graphs based on strong sparse covers. We plug in our strong sparse cover from Theorem 2 to obtain improvements in both space and stretch, see full version [43] for details.

1.4 Related Work

We provided background on sparse covers in the introduction. We refer to the cited papers for additional background on sparse partitions, UTSP, UST, routing and distance oracles. Here we provide additional background on metric embeddings in order to put our results (Corollary 3, Theorem 4) in a wider context. We begin with metric embeddings into ℓ_p spaces.

Every n point metric space embeds into $\ell_2^{O(\log n)}$ with distortion $O(\log n)$ [21], which is also tight [72]. Planar graphs, and more generally fixed minor free graphs, embed into $\ell_2^{O(\log n)}$ with distortion $O(\sqrt{\log n})$ [82, 5], which is also tight [76]. The big open question here is regarding the embedding of such graphs into ℓ_1. The upper bounds are the same as for ℓ_2, while the only lower bound is 2 [71]. A long standing conjecture by Gupta et al. [54] states that every graph family excluding a fixed minor, and in particular planar graphs, can be embedded into ℓ_1 with constant distortion. Some partial progress for planar graph was made in the cases where we care only about vertices laying on a small number of faces [68, 40], or only about vertex pairs laying on the same face [78, 69].

Refined notion of distortion in metric embeddings were studied, such as scaling distortion [4, 18], and terminal/prioritized distortion [30, 31]. In particular, there been study of prioritized low dimensional embeddings into ℓ_∞ [45, 32]. Online metric embeddings into normed spaces were also studied [60, 77, 20].

A different venue of research is metric embeddings into trees, or more generally low treewidth graphs. Every n-point metric space stochastically embeds into trees with expected distortion $O(\log n)$ [16, 35]. This result is tight even when the metric is the shortest path metric of a planar graph [9]. Every planar graph with diameter Δ can be (deterministically) embedded into a graph with treewidth $\tilde{O}(\epsilon^{-3})$ with additive distortion $\epsilon\cdot\Delta$ [51, 49, 25]. Every K_r-minor free graphs stochastically embeds into graphs with treewidth $f(r)\cdot O(\frac{\log\log n}{\epsilon})^2$ with expected additive distortion $\epsilon\cdot\Delta$ [28, 49]. Clan embeddings and Ramsey-type embeddings of K_r-minor free graphs were also studied [48]. Finally, recently it was shown that every K_r-minor free graphs stochastically embeds into graphs with treewidth $f(r)\cdot\tilde{O}(\frac{1}{\epsilon})\cdot\text{poly}(\log(n\cdot\Phi))$ with multiplicative expected distortion $1+\epsilon$ [29] (here Φ is the aspect ratio).

2 Preliminaries

$\tilde{O}$ notation hides poly-logarithmic factors, that is $\tilde{O}(g)=O(g)\cdot\text{polylog}(g)$. All logarithms are at base 2 (unless specified otherwise), ln stand for the natural logarithm. Given a set A, $\binom{A}{2}=\{\{x,y\}\mid x,y\in A, x\neq y\}$ denotes all the subsets of size 2. For a number $k\in\mathbb{N}$, $[k]=\{1,2,\ldots,k\}$.

We consider connected undirected graphs $G=(V,E,w)$ 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 closed 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 $A\subseteq V$, $G[A]$ denotes the induced graph on A, and $G\setminus A:=G[V\setminus A]$. The *diameter* of a graph G is $\text{diam}(G)=\max_{v,u\in V}d_G(v,u)$, i.e. the maximal distance between a pair of vertices. Given a subset $A\subseteq V$, the *weak*-diameter of A is $\text{diam}_G(A)=\max_{v,u\in A}d_G(v,u)$, i.e. the maximal distance between a pair of vertices in A, w.r.t. to d_G. The *strong*-diameter of A is $\text{diam}(G[A])$, the diameter of the graph induced by A.

The *aspect ratio* of a metric space (X,d_X) is usually defined as the ratio between the maximum and minimum distances. However, here we mainly work with the shortest path distance in graphs, where we allow 0-weights. Formally, here the shortest path metric is actually a pseudometric. Accordingly, we will define aspect ratio of a graph $G=(V,E,w)$ as the ratio between the maximum distance to the minimum non zero distance: $\Phi(G)=\frac{\max_{u,v\in V}d_G(u,v)}{\min_{u,v\in V\text{ s.t. }d_G(u,v)>0}d_G(u,v)}$.

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 $K_{3,3}$ minor-free), outer-planar graphs (K_4 and $K_{3,2}$ minor-free), series-parallel graphs (K_4 minor-free) and trees (K_3 minor-free).

The ℓ_∞-norm of a vector $x = (x_1, \dots, x_k) \in \mathbb{R}^k$ is $\|x\|_\infty := \max_{i\in[k]} |x_i|$. An embedding from a metric space (X, d_X) into ℓ_∞ is a function $f : X \to \mathbb{R}^k$. The embedding f has distortion $c \cdot t$ if for every $x, y \in X$, $\frac{1}{c} \cdot d_X(x, y) \le \|f(x) - f(y)\|_\infty \le t \cdot d_X(x, y)$. t is the *expansion* (also known as a *Lipschitz* constant) of f, while c is that *contraction* of f. An embedding with distortion 1 (where $c = t = 1$) is called *isometric*. Embedding $f : X \to \ell_\infty^k$ can be viewed as a collection of embeddings $\{f_i\}_{i=1}^k$ into the line $\mathbb{R}$, where $f_i(x)$ equals to the i'th coordinate of $f(x)$. We will also denote $(f(x))_i = f_i(x)$. Using this notation, f has expansion t if for every $x, y \in X$ and $i \in [k]$, $|f_i(x) - f_i(y)| \le t \cdot d_X(x, y)$. Similarly, f has contraction c if for every $x, y \in X$ there is some $i \in [k]$ such that $|f_i(x) - f_i(y)| \ge \frac{1}{c} \cdot d_X(x, y)$. In this case, we will say that the pair x, y is *satisfied* by the coordinate i.

3 Technical Overview

Cop Decomposition

Abraham et al. [6] constructed a padded decomposition for K_r-minor free graphs based on the cops-and-robbers game [10]. Fix the scale parameter $\Delta > 0$. The process works as follows: pick arbitrary x_1, and let the ball $B_G(x_1, r_1)$ be the first cluster η_1, where $r_1 \in [0, \Delta]$ is sampled using truncated exponential distribution. To construct the second cluster, pick an arbitrary connected component C_2 of $G \setminus \eta_1$, and arbitrary $x_2 \in C_2$. Let T_2 be a shortest path from x_2 to some vertex y with a neighbor in η_1. Then the second cluster $\eta_2 = B_{G[C_2]}(T_2, r_2)$ is a ball around T_2 in the graph induced by the connected component, where the radius $r_2 \in [0, \Delta]$ is sampled using truncated exponential distribution. In general, suppose that we already constructed clusters $\eta_1, \dots, \eta_{k-1}$. Let C_k be an arbitrary connected component of $G \setminus \cup_{i<k}\eta_i$, and $x_k \in C_k$ arbitrary vertex. Let $\mathcal{K}_{C_k} \subseteq \{\eta_1, \dots, \eta_{k-1}\}$ be all the previously created clusters η_i, such that there is an edge from η_i to C_k. Let T_k be a shortest path tree in C_k, with x_k as a root, and such that for every $\eta_i \in \mathcal{K}_{C_k}$, there is an edge from a vertex in T_k to η_i. In particular, T_k will have at most $|\mathcal{K}_{C_k}|$ leaves. The k'th cluster $\eta_k = B_{G[C_k]}(T_k, r_k)$ is a ball around T_k in the induced graph $G[C_k]$, where $r_k \in [0, \Delta]$ is sampled using truncated exponential distribution. See Figure 1 for illustration.

One can run the cop decomposition on general graphs without getting any interesting structure. What makes it particularly interesting for K_r-minor free graphs is the fact that the size of the set $\mathcal{K}_{C_k}$ of neighboring previously created clusters is always bounded by $r - 2$. Indeed, one can argue that if $|\mathcal{K}_{C_k}| \ge r - 1$, than one can contract all the internal edges inside each cluster in $\mathcal{K}_{C_k}$, and the connected component C_k, and obtain K_r as a minor. Thus T_k is a shortest path tree (w.r.t. $G[C_k]$) with at most $r - 2$ leaves, and the cluster C_k is a ball of radius at most Δ around this tree. We will call each such cluster a *supernode*, and the shortest path tree T_k it's *skeleton*. In addition, we construct a tree $\mathcal{T}$ over the supernodes. Here each supernode η_k will be the child in $\mathcal{T}$ of the last created supernode $\eta_i \in \mathcal{K}_{C_k}$. Note that the only outgoing edges from η_k are either to its descendants or ancestors in $\mathcal{T}$. Furthermore, η_k have at most $r - 2$ "neighbor" ancestors. We will call the connected component C_k where we constructed the supernode η_k (with skeleton T_k) the *domain* of η_k, denoted $\text{dom}(\eta_k)$. Note that the vertices in $\text{dom}(\eta_k)$ will either belong to η_k, or to the descendants of η_k w.r.t. $\mathcal{T}$. See Figure 1 for illustration.

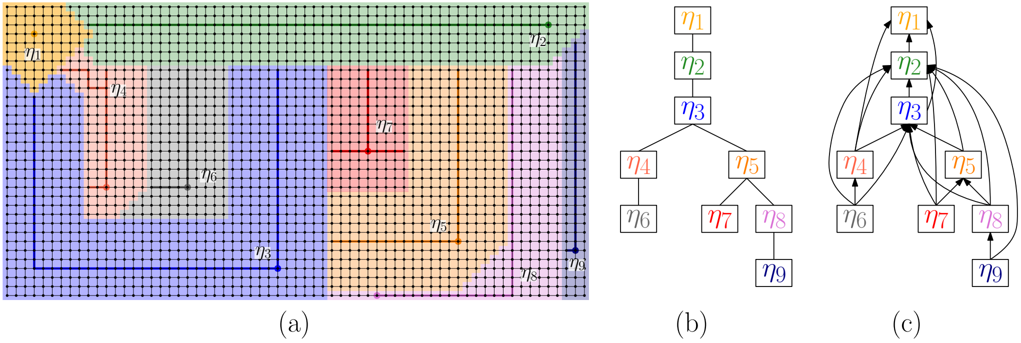

Figure 1 (a) Illustration of an (Δ, γ, w)-buffered cop decomposition of the unweighted grid graph together with (b) - the associated tree $\mathcal{T}$. There are 9 different supernodes $\eta_1, \ldots, \eta_9$, all colored with different colors. Each supernode η contains a shortest path tree T_η (the bold lines) with at most 3 leaves, where all the vertices $x \in \eta$ in the super node are at distance at most $\Delta = 6$ from T_η. The domain of each supernode consist of all the supernodes in its subtree. For example $\mathrm{dom}(\eta_5) = G\left[\eta_5 \cup \eta_7 \cup \eta_8 \cup \eta_9\right]$, and $\mathrm{dom}(\eta_3) = G\left[V \setminus (\eta_1 \cup \eta_2)\right]$. As η_3 and η_9 are not adjacent, the distance from η_3 to any vertex in η_9 (w.r.t. $\mathrm{dom}(\eta_3)$) is at least γ. The associated digraph $\overrightarrow{G}_{\mathcal{C}}$ is illustrated in (c).

Due to the truncated exponential distribution, Abraham et al. [6] showed that a small ball $B_G(x, \gamma\Delta)$ is likely to be fully contained in a single supernode. [6] Unfortunately, the supernodes $\{\eta_i\}_{i\geq 1}$ do not have a bounded diameter. Nevertheless, following [38], using the skeleton T_k, one can partition each supernode η_k into clusters of diameter $O(\Delta)$, while cutting each small ball only with a small probability. Combining these two processes together, one obtains a strong $(O(r), \Omega(\frac{1}{r}), O(\Delta))$-padded decomposition. However, it is unclear if it is possible to use the cop decomposition to create a sparse cover. Indeed, the "dependence tree" $\mathcal{T}$ does not have a bounded depth, and the entire process looks very chaotic. Indeed, every small change in the sampling of the radii leads to a completely different outcome. In contrast, the [65] (as well as [3]) clustering process had only a depth of r, and thus [67] enumerated all the possible choices in the process leading to a sparse cover of exponential sparsity.

Buffered Cop Decomposition

In a recent work, Chang et al. [26] obtained a new "separation"-property in the cop-decomposition. Instead of growing a ball with a random radius around the skeleton T_k, Chang et al. constructed the supernode T_K deterministically. The new separation property is the following, consider a supernode η, and a vertex $v \in \mathrm{dom}(\eta)$ such that v belongs to a supernode η', which is descendant of η, but there is no edge from η to η'. Then $d_{G[\mathrm{dom}(\eta)]}(\eta, v) > \frac{\Delta}{r} = \gamma$. In other words, for every descendant η' of η, either they are neighbors, or every path in $\mathrm{dom}(\eta)$ from η to η' is of length at least γ. [7] Chang et al. called the new partition *Buffered Cop Decomposition*, because there is now a buffer between

[6] Specifically, using a sophisticated argument, based on a potential function, Abraham et al. [6] showed that with probability at least $e^{-O(r)\cdot\gamma}$, the ball $B_G(x, \gamma\Delta)$ is fully contained in a single supernode.

[7] Roughly speaking, [26] begin with the supernode η_k being equal to the skeleton T_k. Then, as the algorithm progresses, each time the buffer property is violated we add the violating vertices to a previously created supernode. One can argue that the depth of this process is bounded by r (once for each neighboring ancestor supernode), and thus the cluster vertices are all within Δ distance from the skeleton.

non-neighboring clusters. They used the new buffered cop decomposition to construct a *shortcut-partition* (which is a generalization of scattering partition [41]). Roughly, a shortcut-partition is a partition of the vertices into clusters of diameter at most $\epsilon \cdot D$, such that for every pair of vertices u, v at distance at most D, there is an approximate shortest path going through at most $O_r(\frac{1}{\epsilon})$ clusters. Chang et al. used their shortcut-partitions to construct tree covers [17, 25], distance oracles [85, 66, 1], to solve the Steiner point removal problem [39, 41, 64, 27, 46], and to construct additive embedding of apex minor free graphs into low treewidth graphs [51, 28, 47, 49, 25].

Sparse Covers

The starting point of this paper is the buffered cop decomposition of [26]. We begin by observing some additional properties. Let $\overrightarrow{G}_{\mathcal{C}}$ be a digraph where the supernodes are the vertices, and there is a directed edge from a supernode η to its ancestor η' (w.r.t. $\mathcal{T}$) iff there is an edge between vertices in η and η'. $\overrightarrow{G}_{\mathcal{C}}$ is a DAG (directed acyclic graph) with maximum out-degree $r-2$, and it has additional crucial property: if v has outgoing edges towards u, z than there has to be an edge between u and z (in one way or another). We call a graph with this property a *transitive DAG*. In transitive DAG's, neighboring vertices tend to share many of their neighbors. Denote by $B_{\overrightarrow{G}_{\mathcal{C}}}(\eta, q)$ the set of supernodes towards which there is a directed path from η of length at most q. We use the transitive DAG property to show that the size of $B_{\overrightarrow{G}_{\mathcal{C}}}(\eta, q)$ is bounded by $\binom{r+q}{r}$ [8]. Note that crucially, for $q \geq r$, $\binom{r+q}{r} \approx O(q)^r$ the growth rate is sub-exponential in q. Next, we generalize the buffer property to argue that for every ancestor supernode η' of η such that $\eta' \notin B_{\overrightarrow{G}_{\mathcal{C}}}(\eta, 2q+1)$ it holds that the distance from every vertex $v \in \eta$ to η' in $\mathrm{dom}(\eta')$ is at least $(q+1) \cdot \frac{\Delta}{r}$. Combining these two properties together, it follows that for every vertex v there are at most $\binom{r+q}{r}$ ancestor supernodes at distance $q \cdot \frac{\Delta}{r}$.

Similar to the process of padded decomposition [6, 38], our sparse cover is constructed in two steps. First we cover the vertices using *enlarged* supernodes, and then we separately cover each enlarged supernode. Fix $q \geq 0$, and for every supernode η, let $\hat{\eta} = B_{G[\mathrm{dom}(\eta)]}(\eta, q \cdot \frac{\Delta}{r})$ be all the vertices at distance at most $q \cdot \frac{\Delta}{r}$ from η in $\mathrm{dom}(\eta)$. In particular, a vertex $v \in \eta$ can join only to the enlarged supernodes which are ancestors of η at distance at most $q \cdot \frac{\Delta}{r}$. Consider the ball $B = B_G(v, \frac{q}{2} \cdot \frac{\Delta}{r})$, and let η be the first supernode containing some vertex from B. By the minimality of η, and the triangle inequality, it will follow that the ball $B \subseteq \hat{\eta}$ is contained in the enlarged supernode. From the other hand, due to the properties discussed above, each vertex v will belong to at most $\binom{r+q}{r}$ enlarged supernodes. Thus we get both the sparsity and the padding properties we wanted. The only missing property at this point is the bounded diameter. Next we cover each enlarged supernode $\hat{\eta}$ using the skeleton T_η. T_η consist of at most r shortest path. We go over these shortest paths, and choose a Δ-net N. Specifically, a set such that every two net points are at distance at least Δ, and every point has a net point at distance at most Δ. Fix $R = 2\Delta + q \cdot \frac{\Delta}{r}$. Due to the properties of shortest paths, every point has at most $O(r+q)$ net points at distance $2R$. Now taking all the balls of radius R around net points provides us the desired sparse cover for the enlarged supernode. Taking the union of all the covers for all the enlarged supernodes we obtain the sparse cover for the graph. Fixing $q = 1$ we obtain padding $O(r)$ and sparsity $O(r^2)$, while by taking $q = \Theta(\frac{r}{\epsilon})$ we obtain padding $4+\epsilon$ and sparsity $O(\frac{1}{\epsilon})^r$.

[8] It was brought to our attention that this fact actually follows from Theorem 4.2. in [52].

Metric embedding into ℓ_∞

Our metric embedding into ℓ_∞ is based on our SPCS . The first to construct a sparse cover based metric embedding was Rao [82], who used the [65] padded decomposition to embed K_r minor free graphs into ℓ_2 with distortion $O(r^3 \cdot \sqrt{\log n})$. Given a partition $\mathcal{P}$, for a vertex v belonging to cluster $v \in C_v \in \mathcal{P}$, let $\partial_{\mathcal{P}}(v) = d_X(v, V \setminus C_v)$ be the distance between v to the boundary of the cluster C_v. Note that if v is padded $B_G(v, \frac{\Delta}{\beta}) \subseteq C_v$, then $\partial_{\mathcal{P}}(v) > \frac{\Delta}{\beta}$. For each cluster $C \in \mathcal{P}$, sample $\alpha_C \in \{\pm 1\}$ u.a.r. . For every vertex v, send v to $\alpha_{C_v} \cdot \partial_{\mathcal{P}}(v)$. By the triangle inequality, it follows that the expansion is at most 2. But for which vertex pairs can we guarantee small contraction?

Consider a pair u, v at distance $d_G(v, u) \in (\Delta, 2\Delta]$, and suppose that $\mathcal{P}$ has diameter at most Δ, and v is padded in $\mathcal{P}$. Then u and v must belong to different clusters C_u, C_v. If it so happened that $\alpha_{C_u} \neq \alpha_{C_v}$, then $|\alpha_{C_v} \cdot \partial_{\mathcal{P}}(v) - \alpha_{C_u} \cdot \partial_{\mathcal{P}}(u)| = \partial_{\mathcal{P}}(v) + \partial_{\mathcal{P}}(u) \geq \frac{\Delta}{\beta} \geq \frac{d_G(u,v)}{2\beta}$, and we obtain a bound on the contraction! Rao took $O(\log n)$ independent samples of the coefficients $\{\alpha_C\}_{C \in \mathcal{P}}$, and gets the contraction guarantee in a constant fraction of the samples. Next, Rao also took $O(\log n)$ independent partitions to get that v is padded in a constant fraction of them. Finally, Rao sampled partitions for all possible distance scales, concatenated the resulting embeddings of them all, and obtained the desired $O(r^3 \cdot \sqrt{\log n})$ distortion. Assuming all the distances are in $[1, \text{poly}(n)]$, there are $O(\log n)$ different distances scales, and the resulting dimension is $O(\log^3 n)$: one log for the number of scales, one log to sample many partitions for each scale, and one log to sample the coefficients $\{\alpha_C\}_{C \in \mathcal{P}}$. Nevertheless, in ℓ_2, using dimension reduction [63], one can reduce the dimension to $O(\log n)$ without significantly increasing the distortion.

The focus of our paper is embeddings into ℓ_∞. One can repeat Rao's embedding exactly as is, and get embedding into $\ell_\infty^{O(\log^3 n)}$ with distortion $O(r^3)$ (or $O(r)$ using the improved padding parameter from [6]). Unfortunately, there is no general dimension reduction in ℓ_∞ (ala [63]). Nevertheless, the dimension still can be dramatically improved. First, observe that when embedding into ℓ_∞, we don't need to succeed on a constant fraction of the partitions (or the α coefficients), it is enough to be successful only once! Thus, instead taking $O(\log n)$ independent samples from a padded decomposition, one can use a SPCS . Indeed, Krauthgamer et al. [67] constructed an $(O(r^2), 3^r)$-SPCS for K_r minor free graphs. Using this SPCS immediately leads to an embedding into $\ell_\infty^{\tilde{O}(3^r) \cdot \log^2 n}$ with distortion $O(r^2)$. To remove additional $\log n$ factor, [67] used an additional property of the [65] based SPCS that does not holds in general: it is possible to create 3^r partitions such that for every pair u, v, there will be a single partition where both u and v will be padded **simultaneously**. [67] heavily relied on this property, while our SPCS (and actually all the others as well) lacking it. Hence we cannot apply [67] as is.

Our solution follow similar lines to [67], but avoids using the additional special structure of [65]. Consider a graph G with a (β, τ)-SPCS . First, for every scale $\Delta_i = \rho^i$, for $\rho = O(\frac{\beta}{\epsilon})$, create τ partitions $\mathcal{P}_i^1, \mathcal{P}_i^2, \ldots, \mathcal{P}_i^\tau$, all with diameter Δ_i, and such that every vertex is β-padded in one of them (that is $\forall v,\ B_G(v, \frac{\Delta}{\beta})$ is contained in some cluster). Next, create, laminar partitions that closely resemble the original partitions. Specifically, we create $\{\tilde{\mathcal{P}}_i^1\}_{i \in \mathbb{Z}}, \{\tilde{\mathcal{P}}_i^2\}_{i \in \mathbb{Z}}, \ldots, \{\tilde{\mathcal{P}}_i^\tau\}_{i \in \mathbb{Z}}$, where for every i, j, $\mathcal{P}_i^j$ refines $\mathcal{P}_{i+1}^j$, $\mathcal{P}_i^j$ has radius at most $(1+\epsilon) \cdot \Delta_i$, and for every $v \in V$, and $i \in \mathbb{Z}$, $B_G(v, \frac{\Delta}{(1+\epsilon)\cdot\beta})$ is fully contained in a cluster of one of $\tilde{\mathcal{P}}_i^1, \ldots, \tilde{\mathcal{P}}_i^\tau$. In other words, for every $v \in V$, and $i \in \mathbb{Z}$, $\max_{j \in [\tau]} \partial_{\mathcal{P}_i^j}(v) > \frac{\Delta_i}{(1+\epsilon)\cdot\beta}$. Similar laminar partitions were also created in [67].

Next, our goal is to embed w.r.t. each laminar partition $\{\tilde{\mathcal{P}}_i^j\}_{i\in\mathbb{Z}}$ independently, such that for every pair u, v which is separated in partition $\tilde{\mathcal{P}}_i^j$ it will hold that $\|f(u) - f(v)\|_\infty \geq \partial_{\tilde{\mathcal{P}}_i^j}(v) + \partial_{\tilde{\mathcal{P}}_i^j}(u)$. [9] Note that given such embedding for all the laminar partitions, we can concatenate them all to obtain the desired distortion. Indeed, constant expansion follows by triangle inequality, while for every u, v, $\|f(u) - f(v)\|_\infty \geq \max_{i,j}\left(\partial_{\tilde{\mathcal{P}}_i^j}(v) + \partial_{\tilde{\mathcal{P}}_i^j}(u)\right) = \Omega(\frac{d_G(u,v)}{\beta})$. To create the embedding w.r.t. to the laminar partition $\{\tilde{\mathcal{P}}_i^j\}_{i\in\mathbb{Z}}$, let $i_0 \in \mathbb{Z}$ be a the maximum index such that $\tilde{\mathcal{P}}_{i_0}^j$ is partitioned into singletons, and let $k \in \mathbb{N}$ be the maximum such that $\tilde{\mathcal{P}}_{i_0+k}^j$ is not the trivial partition into a single cluster $\{V\}$. Clearly it is enough to embed only w.r.t. the laminar partition $\{\tilde{\mathcal{P}}_i^j\}_{i=i_0}^{i_0+k}$. Let $\tilde{\mathcal{P}}_{i_0+k}^j = \{A_1, \dots, A_m\}$ be the clusters in the top partition in our hierarchy. We create a prefix free code $\alpha : \tilde{\mathcal{P}}_{i_0+k}^j \to \{\pm 1\}^*$. Specifically, each cluster A_q is assigned a string of ± 1 of length at most $2 \cdot \left\lceil \log \frac{|V|}{|A_q|} \right\rceil$. The strings of different clusters might be of different length. However, for every $A_q, A_{q'}$ there is an index s such that $\alpha_s(A_q), \alpha_s(A_{q'})$ exist and differ. Then the embedding of each vertex $v \in A_q$ defined by concatenating the coordinates $\left(\alpha_1(A_q) \cdot \partial_{\tilde{\mathcal{P}}_{i_0+k}^j}(v), \alpha_2(A_q) \cdot \partial_{\tilde{\mathcal{P}}_{i_0+k}^j}(v), \dots\right)$ with an embedding created inductively for A_q w.r.t. the induced partition by $\{\tilde{\mathcal{P}}_i^j\}_{i=i_0}^{i_0+k-1}$. To bound the contraction, consider a pair u, v and suppose that they are first separated in level $k' \in [i_0, i_0 + k]$. Then the embeddings of u and v are "aligned" in scales $[k' + 1, i_0 + k]$, while in scale k' they belong to respective clusters $A_v, A_u \in \tilde{\mathcal{P}}_{k'}^j$. The respective codes $\alpha(A_v), \alpha(A_u)$ will differ in some coordinate s and thus $\|f(u) - f(v)\|_\infty \geq \left|\alpha_s(A_u) \cdot \partial_{\tilde{\mathcal{P}}_{k'}^j}(u) - \alpha_s(A_v) \cdot \partial_{\tilde{\mathcal{P}}_{k'}^j}(v)\right| = \partial_{\tilde{\mathcal{P}}_{k'}^j}(u) + \partial_{\tilde{\mathcal{P}}_{k'}^j}(v)$. To conclude a bound on the contraction it remains to observe that for every partition $\tilde{\mathcal{P}}_l^j$ that refines $\tilde{\mathcal{P}}_{k'}^j$ it holds that $\partial_{\tilde{\mathcal{P}}_{k'}^j}(v) + \partial_{\tilde{\mathcal{P}}_{k'}^j}(u) \geq \partial_{\tilde{\mathcal{P}}_l^j}(v) + \partial_{\tilde{\mathcal{P}}_l^j}(u)$.

The overall number of coordinates used for the embedding of the laminar partition $\{\tilde{\mathcal{P}}_i^j\}_{i=i_0}^{i_0+k}$ is $2 \cdot \lceil \log n \rceil + 2(k + 1)$, where k have to be bounded by a logarithm of the aspect ratio $O(\log \Phi)$. If we started from a (β, τ)-SPCS , there are τ laminar partitions and thus the overall dimension is $O(\tau \cdot \log(n \cdot \Phi))$. We remove the dependence on the aspect ratio using fairly standard techniques. Specifically, by observing that for far enough scales, we can use the same coordinate. For this to hold, when treating scale ρ^i, we need to "contract" all vertex pairs at distance at most $\frac{\rho^i}{n^2}$ (as otherwise a pair can accumulate error in unbounded number of scales). Hence we can remove the dependence on the aspect ratio only if the contracted graph still admits a SPCS .

References

1 I. Abraham and C. Gavoille. Object location using path separators. In *Proceedings of the Twenty-fifth Annual ACM Symposium on Principles of Distributed Computing*, PODC '06, pages 188–197, 2006. Full version: `https://www.cse.huji.ac.il/~ittaia/papers/AG-TR.pdf`. `doi:10.1145/1146381.1146411`.

2 I. Abraham, C. Gavoille, A. Gupta, O. Neiman, and K. Talwar. Cops, robbers, and threatening skeletons: Padded decomposition for minor-free graphs. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing*, STOC '14, pages 79–88, 2014. `doi:10.1145/2591796.2591849`.

[9] For comparison, [67] used the simultaneous padding property of [65] and only guaranteed $\|f(u) - f(v)\|_\infty \geq \min\left\{\partial_{\tilde{\mathcal{P}}_i^j}(v), \partial_{\tilde{\mathcal{P}}_i^j}(u)\right\}$.

3 I. Abraham, C. Gavoille, D. Malkhi, and U. Wieder. Strong-diameter decompositions of minor free graphs. *Theory of Computing Systems*, 47(4):837–855, 2010. doi:10.1007/s00224-010-9283-6.

4 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.

5 Ittai Abraham, Arnold Filtser, Anupam Gupta, and Ofer Neiman. Metric embedding via shortest path decompositions. *SIAM J. Comput.*, 51(2):290–314, 2022. a priliminary version apperared in the proceedings of STOC 18. doi:10.1137/19m1296021.

6 Ittai Abraham, Cyril Gavoille, Anupam Gupta, Ofer Neiman, and Kunal Talwar. Cops, robbers, and threatening skeletons: Padded decomposition for minor-free graphs. *SIAM J. Comput.*, 48(3):1120–1145, 2019. preliminary version published in STOC 2014. doi:10.1137/17M1112406.

7 Ittai Abraham, Cyril Gavoille, and Dahlia Malkhi. Compact routing for graphs excluding a fixed minor. In Pierre Fraigniaud, editor, *Distributed Computing, 19th International Conference, DISC 2005, Cracow, Poland, September 26-29, 2005, Proceedings*, volume 3724 of *Lecture Notes in Computer Science*, pages 442–456. Springer, 2005. doi:10.1007/11561927_32.

8 Ittai Abraham, Cyril Gavoille, Dahlia Malkhi, Noam Nisan, and Mikkel Thorup. Compact name-independent routing with minimum stretch. *ACM Trans. Algorithms*, 4(3):37:1–37:12, 2008. doi:10.1145/1367064.1367077.

9 Noga Alon, Richard M. Karp, David Peleg, and Douglas B. West. A graph-theoretic game and its application to the k-server problem. *SIAM J. Comput.*, 24(1):78–100, 1995. preliminary version published in On-Line Algorithms 1991. doi:10.1137/S0097539792224474.

10 Thomas Andreae. On a pursuit game played on graphs for which a minor is excluded. *J. Comb. Theory, Ser. B*, 41(1):37–47, 1986. doi:10.1016/0095-8956(86)90026-2.

11 Hagit Attiya and Jennifer L. Welch. *Distributed computing - fundamentals, simulations, and advanced topics (2. ed.).* Wiley series on parallel and distributed computing. Wiley, 2004.

12 B. Awerbuch and D. Peleg. Network synchronization with polylogarithmic overhead. In *Proc. 31st IEEE Symp. on Foundations of Computer Science*, pages 514–522, 1990.

13 Baruch Awerbuch, Shay Kutten, and David Peleg. On buffer-economical store-and-forward deadlock prevention. *IEEE Trans. Commun.*, 42(11):2934–2937, 1994. doi:10.1109/26.328973.

14 Baruch Awerbuch and David Peleg. Sparse partitions. In *Proceedings of the 31st IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 503–513, 1990. doi:10.1109/FSCS.1990.89571.

15 Baruch Awerbuch and David Peleg. Concurrent online tracking of mobile users. In Lyman Chapin, editor, *Proceedings of the Conference on Communications Architecture & Protocols, SIGCOMM 1991, Zürich, Switzerland, September 3-6, 1991*, pages 221–233. ACM, 1991. doi:10.1145/115992.116013.

16 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.

17 Yair Bartal, Nova Fandina, and Ofer Neiman. Covering metric spaces by few trees. In *46th International Colloquium on Automata, Languages, and Programming, ICALP 2019, July 9-12, 2019, Patras, Greece*, pages 20:1–20:16, 2019. doi:10.4230/LIPIcs.ICALP.2019.20.

18 Yair Bartal, Arnold Filtser, and Ofer Neiman. On notions of distortion and an almost minimum spanning tree with constant average distortion. *J. Comput. Syst. Sci.*, 105:116–129, 2019. preliminary version published in SODA 2016. doi:10.1016/j.jcss.2019.04.006.

19 Yair Bartal and Lee-Ad Gottlieb. Approximate nearest neighbor search for ℓ_p-spaces ($2 < p < \infty$). *Theor. Comput. Sci.*, 757:27–35, 2019. doi:10.1016/J.TCS.2018.07.011.

20 Sujoy Bhore, Arnold Filtser, and Csaba D. Toth. Online duet between metric embeddings and minimum-weight perfect matchings. In *Proceedings of the 2024 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 4564–4579, 2024. doi:10.1137/1.9781611977912.162.

21 J. Bourgain. On lipschitz embedding of finite metric spaces in hilbert space. *Israel Journal of Mathematics*, 52(1-2):46–52, 1985. `doi:10.1007/BF02776078`.

22 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. IEEE Computer Society, 2012. `doi:10.1109/FOCS.2012.45`.

23 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`.

24 Ostas Busch, Da Qi Chen, Arnold Filtser, Daniel Hathcock, D Ellis Hershkowitz, and Rajmohan Rajaraman. One tree to rule them all: Poly-logarithmic universal steiner tree. In *2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 60–76, 2023. `doi:10.1109/FOCS57990.2023.00012`.

25 Hsien-Chih Chang, Jonathan Conroy, Hung Le, Lazar Milenkovic, Shay Solomon, and Cuong Than. Covering planar metrics (and beyond): O(1) trees suffice. In *2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 2231–2261, 2023. `doi:10.1109/FOCS57990.2023.00139`.

26 Hsien-Chih Chang, Jonathan Conroy, Hung Le, Lazar Milenković, Shay Solomon, and Cuong Than. Shortcut partitions in minor-free graphs: Steiner point removal, distance oracles, tree covers, and more. In *Proceedings of the 2024 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 5300–5331, 2024. `doi:10.1137/1.9781611977912.191`.

27 Yun Kuen Cheung. Steiner point removal - distant terminals don't (really) bother. In Artur Czumaj, editor, *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2018, New Orleans, LA, USA, January 7-10, 2018*, pages 1353–1360. SIAM, 2018. `doi:10.1137/1.9781611975031.89`.

28 Vincent Cohen-Addad, Arnold Filtser, Philip N. Klein, and Hung Le. On light spanners, low-treewidth embeddings and efficient traversing in minor-free graphs. *CoRR*, abs/2009.05039, 2020. To appear in FOCS 2020,`https://arxiv.org/abs/2009.05039`. `arXiv:2009.05039`.

29 Vincent Cohen-Addad, Hung Le, Marcin Pilipczuk, and Michał Pilipczuk. Planar and minor-free metrics embed into metrics of polylogarithmic treewidth with expected multiplicative distortion arbitrarily close to 1*. In *2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 2262–2277, 2023. `doi:10.1109/FOCS57990.2023.00140`.

30 M. Elkin, A. Filtser, and O. Neiman. Terminal embeddings. *Theoretical Computer Science*, 697:1–36, 2017. `doi:10.1016/j.tcs.2017.06.021`.

31 Michael Elkin, Arnold Filtser, and Ofer Neiman. Prioritized metric structures and embedding. *SIAM J. Comput.*, 47(3):829–858, 2018. `doi:10.1137/17M1118749`.

32 Michael Elkin and Ofer Neiman. Lossless prioritized embeddings. *SIAM J. Discret. Math.*, 36(3):1529–1550, 2022. `doi:10.1137/21M1436221`.

33 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`.

34 P. Erdős. Extremal problems in graph theory. In *in "Theory of Graphs and Its Applications," Proc. Sympos. Smolenice*, pages 29–36, 1964.

35 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, November 2004. preliminary version published in STOC 2003. `doi:10.1016/j.jcss.2004.04.011`.

36 Jittat Fakcharoenphol and Kunal Talwar. An improved decomposition theorem for graphs excluding a fixed minor. In *RANDOM-APPROX*, pages 36–46, 2003. `doi:10.1007/978-3-540-45198-3_4`.

37 Martin Farach-Colton and Piotr Indyk. Approximate nearest neighbor algorithms for hausdorff metrics via embeddings. In *40th Annual Symposium on Foundations of Computer Science, FOCS '99, 17-18 October, 1999, New York, NY, USA*, pages 171–180. IEEE Computer Society, 1999. `doi:10.1109/SFFCS.1999.814589`.

38 Arnold Filtser. On strong diameter padded decompositions. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2019, September 20-22, 2019, Massachusetts Institute of Technology, Cambridge, MA, USA*, pages 6:1–6:21, 2019. doi:10.4230/LIPIcs.APPROX-RANDOM.2019.6.

39 Arnold Filtser. Steiner point removal with distortion $O(logk)$ using the relaxed-voronoi algorithm. *SIAM J. Comput.*, 48(2):249–278, 2019. doi:10.1137/18M1184400.

40 Arnold Filtser. A face cover perspective to ℓ_1 embeddings of planar graphs. In Shuchi Chawla, editor, *Proceedings of the 2020 ACM-SIAM Symposium on Discrete Algorithms, SODA 2020, Salt Lake City, UT, USA, January 5-8, 2020*, pages 1945–1954. SIAM, 2020. doi:10.1137/1.9781611975994.120.

41 Arnold Filtser. Scattering and sparse partitions, and their applications. In Artur Czumaj, Anuj Dawar, and Emanuela Merelli, editors, *47th International Colloquium on Automata, Languages, and Programming, ICALP 2020, July 8-11, 2020, Saarbrücken, Germany (Virtual Conference)*, volume 168 of *LIPIcs*, pages 47:1–47:20. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2020. doi:10.4230/LIPIcs.ICALP.2020.47.

42 Arnold Filtser. Labeled nearest neighbor search and metric spanners via locality sensitive orderings. In Erin W. Chambers and Joachim Gudmundsson, editors, *39th International Symposium on Computational Geometry, SoCG 2023, June 12-15, 2023, Dallas, Texas, USA*, volume 258 of *LIPIcs*, pages 33:1–33:18. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023. doi:10.4230/LIPIcs.SoCG.2023.33.

43 Arnold Filtser. On sparse covers of minor free graphs, low dimensional metric embeddings, and other applications. *CoRR*, abs/2401.14060, 2024. doi:10.48550/arXiv.2401.14060.

44 Arnold Filtser, Yuval Gitlitz, and Ofer Neiman. Light, reliable spanners. *CoRR*, abs/2307.16612, 2023. doi:10.48550/arXiv.2307.16612.

45 Arnold Filtser, Lee-Ad Gottlieb, and Robert Krauthgamer. Labelings vs. embeddings: On distributed and prioritized representations of distances. *Discret. Comput. Geom.*, 2023. doi:10.1007/s00454-023-00565-2.

46 Arnold Filtser, Robert Krauthgamer, and Ohad Trabelsi. Relaxed voronoi: A simple framework for terminal-clustering problems. In Jeremy T. Fineman and Michael Mitzenmacher, editors, *2nd Symposium on Simplicity in Algorithms, SOSA 2019, January 8-9, 2019, San Diego, CA, USA*, volume 69 of *OASIcs*, pages 10:1–10:14. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2019. doi:10.4230/OASICS.SOSA.2019.10.

47 Arnold Filtser and Hung Le. Clan embeddings into trees, and low treewidth graphs. In Samir Khuller and Virginia Vassilevska Williams, editors, *STOC '21: 53rd Annual ACM SIGACT Symposium on Theory of Computing, Virtual Event, Italy, June 21-25, 2021*, pages 342–355. ACM, 2021. doi:10.1145/3406325.3451043.

48 Arnold Filtser and Hung Le. Locality-sensitive orderings and applications to reliable spanners. In Stefano Leonardi and Anupam Gupta, editors, *STOC '22: 54th Annual ACM SIGACT Symposium on Theory of Computing, Rome, Italy, June 20 - 24, 2022*, pages 1066–1079. ACM, 2022. doi:10.1145/3519935.3520042.

49 Arnold Filtser and Hung Le. Low treewidth embeddings of planar and minor-free metrics. In *63rd IEEE Annual Symposium on Foundations of Computer Science, FOCS 2022, Denver, CO, USA, October 31 - November 3, 2022*, pages 1081–1092. IEEE, 2022. doi:10.1109/FOCS54457.2022.00105.

50 Arnold Filtser and Ofer Neiman. Light spanners for high dimensional norms via stochastic decompositions. *Algorithmica*, 2022. doi:10.1007/s00453-022-00994-0.

51 E. Fox-Epstein, P. N. Klein, and A. Schild. Embedding planar graphs into low-treewidth graphs with applications to efficient approximation schemes for metric problems. In *Proceedings of the 30th Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '19, pages 1069–1088, 2019. doi:10.1137/1.9781611975482.66.

52 Martin Grohe, Stephan Kreutzer, Roman Rabinovich, Sebastian Siebertz, and Konstantinos S. Stavropoulos. Coloring and covering nowhere dense graphs. *SIAM J. Discret. Math.*, 32(4):2467–2481, 2018. doi:10.1137/18M1168753.

53 Anupam Gupta, Mohammad Taghi Hajiaghayi, and Harald Räcke. Oblivious network design. In *Proceedings of the Seventeenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2006, Miami, Florida, USA, January 22-26, 2006*, pages 970–979. ACM Press, 2006. URL: http://dl.acm.org/citation.cfm?id=1109557.1109665.

54 Anupam Gupta, Ilan Newman, Yuri Rabinovich, and Alistair Sinclair. Cuts, trees and l_1-embeddings of graphs. *Comb.*, 24(2):233–269, 2004. doi:10.1007/S00493-004-0015-X.

55 Sariel Har-Peled, Piotr Indyk, and Anastasios Sidiropoulos. Euclidean spanners in high dimensions. In *Proceedings of the Twenty-Fourth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2013, New Orleans, Louisiana, USA, January 6-8, 2013*, pages 804–809, 2013. doi:10.1137/1.9781611973105.57.

56 Sariel Har-Peled, Manor Mendel, and Dániel Oláh. Reliable spanners for metric spaces. *ACM Trans. Algorithms*, 19(1):7:1–7:27, 2023. doi:10.1145/3563356.

57 Makoto Imase and Bernard M. Waxman. Dynamic steiner tree problem. *SIAM J. Discret. Math.*, 4(3):369–384, 1991. doi:10.1137/0404033.

58 Piotr Indyk. On approximate nearest neighbors in non-euclidean spaces. In *39th Annual Symposium on Foundations of Computer Science, FOCS '98, November 8-11, 1998, Palo Alto, California, USA*, pages 148–155. IEEE Computer Society, 1998. doi:10.1109/SFCS.1998.743438.

59 Piotr Indyk. Approximate nearest neighbor algorithms for Fréchet distance via product metrics. In *Proceedings of the 8th Symposium on Computational Geometry*, pages 102–106, Barcelona, Spain, June 2002. ACM Press. doi:10.1145/513400.513414.

60 Piotr Indyk, Avner Magen, Anastasios Sidiropoulos, and Anastasios Zouzias. Online embeddings. In Maria J. Serna, Ronen Shaltiel, Klaus Jansen, and José D. P. Rolim, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, 13th International Workshop, APPROX 2010, and 14th International Workshop, RANDOM 2010, Barcelona, Spain, September 1-3, 2010. Proceedings*, volume 6302 of *Lecture Notes in Computer Science*, pages 246–259. Springer, 2010. doi:10.1007/978-3-642-15369-3_19.

61 Lujun Jia, Guolong Lin, Guevara Noubir, Rajmohan Rajaraman, and Ravi Sundaram. Universal approximations for tsp, steiner tree, and set cover. In Harold N. Gabow and Ronald Fagin, editors, *Proceedings of the 37th Annual ACM Symposium on Theory of Computing, Baltimore, MD, USA, May 22-24, 2005*, pages 386–395. ACM, 2005. doi:10.1145/1060590.1060649.

62 D. S. Johnson. The NP-completeness column: An ongoing guide (column 19). *Journal of Algorithms*, 8(3):438–448, 1987.

63 William Johnson and Joram Lindenstrauss. Extensions of Lipschitz mappings into a Hilbert space. *Contemporary Mathematics*, 26:189–206, 1984. see here.

64 Lior Kamma, Robert Krauthgamer, and Huy L. Nguyen. Cutting corners cheaply, or how to remove steiner points. *SIAM J. Comput.*, 44(4):975–995, 2015. doi:10.1137/140951382.

65 Philip 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*, STOC '93, pages 682–690, New York, NY, USA, 1993. ACM. doi:10.1145/167088.167261.

66 Philip N. Klein. Preprocessing an undirected planar network to enable fast approximate distance queries. In *Proceedings of the Thirteenth Annual ACM-SIAM Symposium on Discrete Algorithms, January 6-8, 2002, San Francisco, CA, USA*, pages 820–827, 2002. see here. URL: http://dl.acm.org/citation.cfm?id=545381.545488.

67 Robert Krauthgamer, James R. Lee, Manor Mendel, and Assaf Naor. Measured descent: a new embedding method for finite metrics. *Geometric and Functional Analysis*, 15(4):839–858, 2005. preliminary version published in FOCS 2004. doi:10.1007/s00039-005-0527-6.

68 Robert Krauthgamer, James R. Lee, and Havana Rika. Flow-cut gaps and face covers in planar graphs. In Timothy M. Chan, editor, *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2019, San Diego, California, USA, January 6-9, 2019*, pages 525–534. SIAM, 2019. doi:10.1137/1.9781611975482.33.

69 Nikhil Kumar. An approximate generalization of the okamura-seymour theorem. In *63rd IEEE Annual Symposium on Foundations of Computer Science, FOCS 2022, Denver, CO, USA, October 31 - November 3, 2022*, pages 1093–1101. IEEE, 2022. doi:10.1109/FOCS54457.2022.00106.

70 Hung Le and Shay Solomon. A unified framework for light spanners. In Barna Saha and Rocco A. Servedio, editors, *Proceedings of the 55th Annual ACM Symposium on Theory of Computing, STOC 2023, Orlando, FL, USA, June 20-23, 2023*, pages 295–308. ACM, 2023. doi:10.1145/3564246.3585185.

71 James R. Lee and Prasad Raghavendra. Coarse differentiation and multi-flows in planar graphs. *Discret. Comput. Geom.*, 43(2):346–362, 2010. doi:10.1007/S00454-009-9172-4.

72 Nathan Linial, Eran London, and Yuri Rabinovich. The geometry of graphs and some of its algorithmic applications. *Comb.*, 15(2):215–245, 1995. preliminary version published in FOCS 1994. doi:10.1007/BF01200757.

73 Nancy A. Lynch. *Distributed Algorithms.* Morgan Kaufmann, 1996.

74 Jiří Matoušek. On the distortion required for embedding finite metric spaces into normed spaces. *Israel Journal of Mathematics*, 93(1):333–344, 1996. doi:10.1007/BF02761110.

75 Ofer Neiman. Low dimensional embeddings of doubling metrics. *Theory Comput. Syst.*, 58(1):133–152, 2016. doi:10.1007/S00224-014-9567-3.

76 Ilan Newman and Yuri Rabinovich. A lower bound on the distortion of embedding planar metrics into euclidean space. In *SCG '02: Proceedings of the eighteenth annual symposium on Computational geometry*, pages 94–96. ACM Press, 2002. doi:10.1145/513400.513412.

77 Ilan Newman and Yuri Rabinovich. Online embedding of metrics. In Susanne Albers, editor, *17th Scandinavian Symposium and Workshops on Algorithm Theory, SWAT 2020, June 22-24, 2020, Tórshavn, Faroe Islands*, volume 162 of *LIPIcs*, pages 32:1–32:13. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2020. doi:10.4230/LIPICS.SWAT.2020.32.

78 Haruko Okamura and P.D. Seymour. Multicommodity flows in planar graphs. *Journal of Combinatorial Theory, Series B*, 31(1):75–81, 1981. doi:10.1016/S0095-8956(81)80012-3.

79 David Peleg. Distance-dependent distributed directories. *Inf. Comput.*, 103(2):270–298, 1993. doi:10.1006/INCO.1993.1020.

80 David Peleg. *Distributed Computing: A Locality-Sensitive Approach.* SIAM, Philadelphia, PA, 2000. doi:10.1137/1.9780898719772.

81 David Peleg and Eli Upfal. A trade-off between space and efficiency for routing tables. *J. ACM*, 36(3):510–530, 1989. doi:10.1145/65950.65953.

82 Satish Rao. Small distortion and volume preserving embeddings for planar and Euclidean metrics. In *Proceedings of the Fifteenth Annual Symposium on Computational Geometry, Miami Beach, Florida, USA, June 13-16, 1999*, pages 300–306, 1999. doi:10.1145/304893.304983.

83 N. Robertson and P. D. Seymour. Graph minors. XVI. Excluding a non-planar graph. *Journal of Combinatoral Theory Series B*, 89(1):43–76, 2003. doi:10.1016/S0095-8956(03)00042-X.

84 Srinivasagopalan Srivathsan, Costas Busch, and S. Sitharama Iyengar. Oblivious buy-at-bulk in planar graphs. In Naoki Katoh and Amit Kumar, editors, *WALCOM: Algorithms and Computation - 5th International Workshop, WALCOM 2011, New Delhi, India, February 18-20, 2011. Proceedings*, volume 6552 of *Lecture Notes in Computer Science*, pages 33–44. Springer, 2011. doi:10.1007/978-3-642-19094-0_6.

85 Mikkel Thorup. Compact oracles for reachability and approximate distances in planar digraphs. *J. ACM*, 51(6):993–1024, November 2004. doi:10.1145/1039488.1039493.

86 Mikkel Thorup and Uri Zwick. Compact routing schemes. In *Proceedings of the Thirteenth Annual ACM Symposium on Parallel Algorithms and Architectures, SPAA 2001, Heraklion, Crete Island, Greece, July 4-6, 2001*, pages 1–10, 2001. doi:10.1145/378580.378581.