Low Treewidth Embeddings of Planar and Minor-Free Metrics

[1], Arnold Filtser*[1], and Hung Le†[2]

[1]Bar-Ilan University, arnold.filtser@biu.ac.il
[2]University of Massachusetts at Amherst, hungle@cs.umass.edu

Abstract

Cohen-Addad, Filtser, Klein and Le [FOCS'20] constructed a stochastic embedding of minor-free graphs of diameter D into graphs of treewidth $O_\epsilon(\log n)$ with expected additive distortion $+\epsilon D$. Cohen-Addad *et al.* then used the embedding to design the first quasi-polynomial time approximation scheme (QPTAS) for the capacitated vehicle routing problem. Filtser and Le [STOC'21] used the embedding (in a different way) to design a QPTAS for the metric Baker's problems in minor-free graphs. In this work, we devise a new embedding technique to improve the treewidth bound of Cohen-Addad *et al.* exponentially to $O_\epsilon(\log\log n)^2$. As a corollary, we obtain the first efficient PTAS for the capacitated vehicle routing problem in minor-free graphs. We also significantly improve the running time of the QPTAS for the metric Baker's problems in minor-free graphs from $n^{O_\epsilon(\log(n))}$ to $n^{O_\epsilon(\log\log(n))^3}$.

Applying our embedding technique to planar graphs, we obtain a deterministic embedding of planar graphs of diameter D into graphs of treewidth $O((\log\log n)^2)/\epsilon)$ and additive distortion $+\epsilon D$ that can be constructed in nearly linear time. Important corollaries of our result include a bicriteria PTAS for metric Baker's problems and a PTAS for the vehicle routing problem with bounded capacity in planar graphs, both run in *almost-linear* time. The running time of our algorithms is significantly better than previous algorithms that require quadratic time.

A key idea in our embedding is the construction of an (exact) emulator for tree metrics with treewidth $O(\log\log n)$ and hop-diameter $O(\log\log n)$. This result may be of independent interest.

*This research was supported by the ISRAEL SCIENCE FOUNDATION (grant No. 1042/22).

†The research was supported by the start-up grant of UMass Amherst and and by the National Science Foundation under Grant No. CCF-2121952.

Contents

1 Introduction

Metric embedding is an influential algorithmic tool that has been applied to many different settings, for example, approximation/sublinear/online/distributed algorithms [LLR95, AMS99, BCL+18, KKM+12], machine learning [GKK17], computational biology [HBK+03], and computer vision[AS03]. The fundamental idea of metric embedding in solving an algorithmic problem is to embed an input metric space to a host metric space that is "simpler" than the input metric space, solve the problem in the (simple) host metric space and then map the solution back to a solution of the input metric space. In this algorithmic pipeline, the structure of the host metric space plays a decisive role.

In their seminal result, Fakcharoenphol, Rao and Talwar [FRT04] (improving over Bartal [Bar96, Bar98], see also [Bar04]) constructed a stochastic embedding of an arbitrary n-point metric space to a tree with expected multiplicative distortion $O(\log n)$; the distortion was shown to be optimal [Bar96]. One may hope to get better distortion by constraining the structure of the input metric space, or enriching the host space. Shattering such hopes, Carroll and Goel [CG04] (implicitly) showed that there in an infinite family of planar graphs, such that every deterministic embedding into treewidth $n^{\frac{1}{3}}$ graphs will have multiplicative distortion $\Omega(n^{\frac{1}{3}})$.[1] Furthermore, Chakrabarti, Jaffe, Lee, and Vincent [CJLV08] showed that any stochastic embedding of planar graphs into graphs with *constant* treewidth requires expected distortion $\Omega(\log n)$. Carroll and Goel [CG04] showed a more general trade-off: any stochastic embedding of planar graphs with expected distortion $c \geq 1$ requires treewidth $\Omega(\log(n)/c)$. In particular, it any stochastic embedding with expected distortion $\sqrt{\log n}$ has treewidth $O(\sqrt{\log n})$.

Bypassing this roadblock, Fox-Epstein, Klein, and Schild [FKS19] studied *additive embeddings*: a Δ-additive embedding $f : V(G) \to V(H)$ of a graph G to a graph H is an embedding such that:

$$\text{for every } u, v \in V(G), \qquad d_G(u,v) \leq d_H(f(u), f(v)) \leq d_G(u,v) + \Delta \ .$$

The parameter Δ is the *additive distortion* of the Δ-additive embedding f. Fox-Epstein *et al.* [FKS19] showed that planar graphs of diameter D admit a (deterministic) (ϵD)-additive embedding into graphs of treewidth $O(\epsilon^{-c})$ for some universal constant c. The constant treewidth bound (for a constant ϵ) sharply contrasts additive embeddings with multiplicative embeddings (where the treewidth is polynomial). Their motivation for developing the additive embedding was to design PTASes for the metric Baker's problems in planar graphs.

In a seminal paper [Bak94], Baker designed PTASes for several problems in planar graphs such as independent set, dominating set, and vertex cover, where vertices have non-negative measures (or weights). Note that these problems are APX-hard in general graphs. Baker's results subsequently inspired the development of powerful algorithmic frameworks for planar graphs, such as deletion decomposition [Bak94, Epp00, DHK05], contraction decomposition [Kle05, DHM07, DHK11], and bidimensionality [DH05, FLRS11]. Metric Baker's problems generalize Baker's problems in that vertices in the solution must be at least/at most a distance ρ from each other for some input parameter ρ. The most well-studied examples of metric Baker's problems include ρ-independent set, ρ-dominating set, (k,r)-center. Metric Baker problems have been studied in the context of parameterized complexity [DFHT05, MP15, BL16, KLP20] where (a) the input graphs are restricted to subclasses of minor-free graphs, such as planar and bounded treewidth graphs, and (b) parameters such as ρ and/or the size of the optimal solution are small. When ρ is a constant, Baker's layering technique can be applied to obtain a linear time PTAS for unweighted

[1] This lower bound is achieved by applying Lemma 24 on the $n-1$-subdivision of the $n \times n$ grid.

planar graphs [Bak94] and efficient PTASes for unweighted minor-free graphs [DHK05]. However, the most challenging case is when the graph is weighted, and ρ is part of the input; even when restricted to bounded treewidth graph, (single-criteria) PTASes are not known for metric Baker problems.[2] Marx and Pilipczuk [MP15] showed that, under Exponential Time Hypothesis (ETH), ρ-independent/dominating set problems cannot be solved in time $f(k)n^{o(\sqrt{k})}$ when the solution size is at most k. As observed by Fox-Epstein *et al.* [FKS19], the result of Marx and Pilipczuk [MP15] implies that, under ETH, there is no (single-criteria) efficient PTAS for ρ-independent/dominating set problems in planar graphs. However, for the case of uniform measure (i.e. $\forall v$, $\mu(v) = 1$), a (non-efficient) PTAS can be obtained via local search [FL21].

Fox-Epstein *et al.* [FKS19] bypassed the ETH lower bound by designing a *bi-criteria* efficient PTAS for ρ-independent/dominating set problems in planar graphs using their additive distortion embedding of planar graphs into bounded treewidth graphs. Since the treewidth of their embedding is $O(\epsilon^{-c})$ for some constant $c \geq 19$ (see Section 6.5 in [FKS19]), and the running time to construct the embedding is $n^{O(1)}$ for an unspecified constant in the exponent due to the embedding step, the running time of their PTAS is $2^{O(\epsilon^{-(c+1)})}n^{O(1)}$ [FKS19]. In their paper, they noted:

> "Admittedly, in our current proof, the treewidth is bounded by a polynomial of very high degree in $1/\epsilon$. There is some irony in the fact that our approach to achieving an efficient PTAS yields an algorithm that is inefficient in the constant dependence on ϵ."

Given the state of affairs, the following problem naturally arises:

Question 1. *Can we design a PTAS for metric Baker's problems with (almost) linear running time? Can we obtain a PTAS with a more practical dependency on ϵ?*

Cohen-Addad, Filtser, Klein, and Le [CFKL20] studied additive embeddings in a more general setting of K_r-minor-free graphs for any fixed r. They proved a strong lower bound on treewidth against *deterministic* additive embeddings. Specifically, they showed (Theorem 3 in [CFKL20]) that there is an n-vertex K_6-minor-free graph such that any deterministic additive embedding into a graph of treewidth $o(\sqrt{n})$ must incur a distortion at least $\frac{D}{20}$. On the other hand, they showed that randomness helps reduce the treewidth exponentially. In particular, they constructed a *stochastic* additive embedding of K_r-minor-free graphs into graphs with treewidth $O(\frac{\log n}{\epsilon^2})$ and *expected* additive distortion $+\epsilon D$. Specifically, there is distribution $\mathcal{D}$ over dominating embeddings (i.e. no distances shrink) into treewidth $O(\frac{\log n}{\epsilon^2})$ graphs such that $\forall u, v,\ \mathbb{E}_{f,H_f}[d_{H_f}(f(u), d(v))] \leq d_G(u, v) + \epsilon D$.

Their primary motivation was to design a quasi-polynomial time approximation scheme (QPTAS) for the bounded-capacity vehicle routing problem (VRP) in K_r-minor-free graphs. In this problem, we are given an edge-weighted graph $G = (V, E, w)$, a set of clients $K \subseteq V$, a depot $r \in V$, and the capacity $Q \in \mathbb{Z}_+$ of the vehicle. We are tasked with finding a collection of tours $\mathcal{S} = \{R_1, R_2, \ldots\}$ of *minimum cost* such that each tour, starting from r and ending at r, visits at most Q clients and every client is visited by at least one tour; the cost of $\mathcal{S}$ is the total weight of all tours in $\mathcal{S}$. (See Definition 2 for a more formal definition.) The VRP was introduced by Dantzig and Ramser [DR59] and has been extensively studied since then; see the survey by Fisher [Fis95].

[2] In bounded treewidth graphs, no single-criteria algorithm for metric Baker problems are known where the objective of the approximation is the measure of the set. However, there is a single-criteria PTAS where the objective of the approximation is the parameter ρ (and the measure is fixed). See Lemma 9.

The problem is APX-hard, as it is a generalization of the Travelling Salesperson Problem (TSP) when $|Q| = |V|$, which is APX-hard [PY93], and admits a constant factor approximation [HR85]. To get a $(1+\epsilon)$-approximation, it is necessary to restrict the structures of the input graph. Fundamental graph structures that have long been studied are low dimensional Euclidean (or doubling) spaces, planarity and minor-freeness.

When the capacity Q is a part of the input, a QPTAS for VRP in Euclidean space of constant dimension is known [DM15, ACL09]; it remains a major open problem to design a PTAS for VRP even for the Euclidean plane. For trees, a PTAS was only obtained by a recent work of Mathieu and Zhou [MZ21], improving upon the QPTAS of Jayaprakash and Salavatipour [JS22]. No PTAS is known beyond trees, such as planar graphs or bounded treewidth graphs. For the *unsplittable* demand version of the problem on trees, Becker [Bec18] showed that the problem is APX-hard.

Going beyond trees, it is natural to restrict the problem further by considering *constant* Q. In this regime, Becker *et al.* [BKS19] designed the first (randomized) PTAS for bounded-capacity VRP in planar graphs with running time $n^{O_\epsilon(1)}$, improving upon the earlier QPTAS by Becker, Klein and Saulpic [BKS17]. Recently, Cohen-Addad *et al.* [CFKL20] obtained the first efficient PTAS for the problem in planar graphs with running time $O_\epsilon(1)n^{O(1)}$. Their algorithm uses the embedding of [FKS19] as a blackbox, and hence, suffers the drawback of the embedding: the exponent of n in the running time is unspecified due to the embedding step.

In K_r-minor-free graphs, Cohen-Addad *et al.* [CFKL20] designed a QPTAS for bounded-capacity VRP in K_r-minor-free graphs using their (stochastic) embedding of K_r-minor-free graphs into graphs with treewidth $O(\frac{\log n}{\epsilon^2})$. More precisely, their algorithm has running time $(\log n)^{\mathsf{tw}} n^{O(1)}$ where tw is the treewidth of the embedding. Thus, any significant improvement over treewidth bound $O(\frac{\log n}{\epsilon^2})$, such as $O(\log n/\log\log n)$, would lead to a PTAS. Their work left the following questions as open problems.

Question 2. *Can we devise an additive embedding of K_r-minor-free graphs into graphs with treewidth $O(\log(n)/\log\log(n))$? Can we design a PTAS for the bounded-capacity VRP in K_r-minor-free graphs? Can we improve the running time of the PTAS for the bounded-capacity VRP in planar graphs to (almost) linear?*

We remark that the dependency on ϵ of all known approximation schemes for bounded-capacity VRP problem is *doubly exponential* in $1/\epsilon$ [BKS19, CFKL20]. Reducing the dependency to singly exponential in $1/\epsilon$ is a fascinating open problem.

One drawback of the stochastic embedding with additive distortion is that it cannot be applied directly to design (bicriteria) PTASes for metric Baker problems in minor-free graphs. To remedy this drawback, the authors[FL21] recently introduced *clan embedding* and *Ramsey type embedding* with additive distortions. A clan embedding of a graph G to a graph H is pair of maps (f,χ) where f is a *one-to-many* embedding $f : V(G) \to 2^{V(H)}$ that maps each vertex $x \in V(G)$ to a subset of vertices $f(x) \subseteq V(H)$ in H, called *copies* of x (where each set $f(x) \neq \emptyset$ is not empty, and every two sets of $x \neq y$ are $f(x) \cap f(y) = \emptyset$ are disjoint), and $\chi : V(H) \to V(H)$ maps each vertex x to a vertex $\chi(x) \in f(x)$, called the *chief* of x. Furthermore, f must be *dominating*:

$$\forall x,y \in V(G), \qquad d_G(x,y) \leq \min_{x'\in f(x), y'\in f(y)} d_H(x',y') \ .$$

A clan embedding (f,χ) has additive distortion $+\Delta$ if for every $x,y \in V(G)$:

$$\min_{x'\in f(x)} d_H(x',\chi(y)) \leq d_G(x,y) + \Delta \tag{1}$$

That is, there is some vertex in the clan of x which is close to the chief of y. Note that the distortion guarantee is in the worst case. That is, Equation (1) holds for every pair $x, y \in V(G)$, and every embedding $f \in \mathrm{supp}(\mathcal{D})$ in the support.

A Ramsey type embedding is a (stochastic) one-to-one embedding in which there is a subset of vertices $M \subseteq V$ such that every vertex is included in M with a probability at least $1-\delta$, for a given parameter $\delta \in (0,1)$, and for every vertex $u \in M$, the additive distortion of the distance from u to every other vertex in V is $+\Delta$. See Definition 9 for a formal definition.

The authors [FL21] showed previously that for any given K_r-minor-free graph and parameters $\epsilon, \delta \in (0,1)$, one can construct a distribution $\mathcal{D}$ over clan embeddings into $O_r(\frac{\log^2 n}{\epsilon\delta})$-treewidth graphs with additive distortion $+\epsilon D$ (D being the diameter) and such that the expected clan size $\mathbb{E}[|f(x)|]$ of every $x \in V(G)$ is bounded by $1+\delta$. For Ramsey type embeddings, the treewidth is also $O_r(\frac{\log^2 n}{\epsilon\delta})$ for an additive distortion $+\epsilon D$ [FL21].

Using the clan embedding and Ramsey type embedding, the authorsobtained a QPTAS for metric Baker's problem in K_r-minor-free graphs with running time $n^{O_r(\epsilon^{-2}\log(n)\log\log(n))}$ [FL21]. The precise running time of the algorithm is $(\log n)^{O(\mathsf{tw})} n^{O(1)}$ where tw is the treewidth of the embeddings. The key questions are:

Question 3. *Can we improve the treewidth bound $O(\log^2(n))$ of the clan embedding and Ramsey embedding? Can we design a PTAS for metric Baker's problems in K_r-minor-free graphs?*

1.1 Our Results

In this paper, we provide affirmative answers to Question 1 and Question 2, while making significant progress toward Question 3.

We construct a new stochastic additive embedding of K_r-minor-free graphs into graphs with treewidth $O_r(\frac{(\log\log n)^2}{\epsilon^2})$. Our treewidth bound improves *exponentially* over the treewidth bound of Cohen-Addad *et al.* [CFKL20]. See Table 1 for a summary of new and previous embeddings.

Theorem 1 (Embedding Minor-free Graphs to Low Treewidth Graphs)**.** *Given an n-vertex K_r-minor-free graph $G(V,E,w)$ of diameter D, we can construct in polynomial time a stochastic additive embedding $f : V(G) \to H$ into a distribution over graphs H of treewidth at most $O_r(\epsilon^{-2}(\log\log n)^2)$ and expected additive distortion ϵD.*

The main bottleneck for not having an almost-linear time algorithm for the embedding in minor-free graphs is that the best algorithm for computing Robertson-Seymour decomposition takes quadratic time [KKR12].

Using our stochastic additive embedding, we design the first PTAS for the bounded-capacity vehicle routing problem in K_r-minor-free graphs; our PTAS indeed is efficient. This resolves Question 2 in the affirmative (more on the planar case later in Theorem 6).

Theorem 2 (PTAS for Bounded-Capacity VRP in Minor-free Graphs)**.** *There is a randomized polynomial time approximation scheme for the bounded-capacity VRP in K_r-minor-free graphs that runs in $O_{\epsilon,r}(1) \cdot n^{O(1)}$ time.*

Our proof of Theorem 2 follows the embedding framework of Cohen-Addad *et al.* [CFKL20]. In a nutshell, they showed that if planar graphs with *one vortex* (see Section 5 for a formal definition) have an additive embedding with treewidth $k(n,\epsilon)$, then K_r-minor-free graphs have a stochastic

Family	Type	Treewidth	Runtime	Ref
Planar	Deterministic	$O_r(\epsilon^{-1} \cdot \log n)$	$O_\epsilon(n^{O(1)})$	[EKM14]
		$O_r(\epsilon^{-c}), c \geq 19$	$O_\epsilon(n^{O(1)})$	[FKS19]
		$O_r(\epsilon^{-1} \cdot (\log\log n)^2)$	$\epsilon^{-2} \cdot \tilde{O}(n)$	Theorem 3
K_r-minor free	Stochastic	$O_r(\epsilon^{-2} \cdot \log n)$	$O_{\epsilon,r}(n^{O(1)})$	[CFKL20]
		$O_r(\epsilon^{-2} \cdot (\log\log n)^2)$	$O_{\epsilon,r}(n^{O(1)})$	Theorem 1
	Ramsey type / Clan	$(\epsilon\delta)^{-1} \cdot O_r(\log^2 n)$	$O_{\epsilon,r}(n^{O(1)})$	[FL21]
		$(\epsilon\delta)^{-1} \cdot \tilde{O}_r(\log n)$	$O_{\epsilon,r}(n^{O(1)})$	Theorems 7 and 8

Table 1: *Summary of current and previous embeddings into low treewidth graphs with additive distortion $+\epsilon D$. For planar graphs, we get an embedding whose treewidth has a minor depdency on n; however our running time and dependency on ϵ is much improved. For stochastic embeddings of K_r-minor-free graphs, we obtain an exponential improvement. For Ramsey-type and clan embeddings, we obtain a quadratic improvement. All known results for planar graphs, including ours, can be extended to bounded genus graphs with only a constant loss in the treewidth using the cutting technique described in [CFKL20].*

additive embedding with treewidth roughly $O_r(k(n, O_r(\epsilon^2)) + \log(n))$. That is, the reduction incurs an additive factor of $O(\log n)$. Thus, there are two issues one has to resolve to reduce the treewidth to $o(\log(n))$: (i) construct an embedding of planar graphs with one vortex that has treewidth $k(n,\epsilon) = o(\log(n))$ and (ii) remove the loss $O(\log(n))$ in the reduction of Cohen-Addad *et al.* [CFKL20]. By a relatively simple idea (see Section 6), we could remove the additive term $O(\log(n))$ in the reduction. We are left with constructing an embedding of planar graphs with one vortex, which is the main barrier we overcome in this work.

One can intuitively think of planar graphs with one vortex as *noisy* planar graphs, where the vortex is a kind of low-complexity[3] noise that affects the planar embedding in a local area, i.e., a face. From this point of view, we need an embedding of planar graphs that is *robust* to the noise. The embedding for planar graphs of Fox-Epstein *et al.* [FKS19] relies heavily on planarity to perform topological operations such as cutting paths open, and decomposing the graphs into so-called *bars and cages*. As a result, adding a vortex to a planar graph makes their embedding inapplicable. Cohen-Addad *et al.* [CFKL20] instead use *balanced shortest path separators* in their embedding of planar graphs with one vortex. Using shortest path separators results in an embedding technique that is robust to the noise caused by the vortex. However, their embedding has treewidth $O(\log(n)/\epsilon)$ due to that recursively decomposing the graphs using balanced separators gives a decomposition tree of depth $O(\log n)$. This is a universal phenomenon in almost all techniques that rely on balanced separators: the depth $O(\log n)$ factors in many known algorithms in planar graphs [GKP95, AGK+98, EKM14, FKM15, CAdVK+16]. It seems that to get a treewidth $o(\log n)$, one needs to avoid using balanced separators in the construction.

Surprisingly perhaps, we can still use balanced shortest path separators to get an embedding with treewidth $O((\log\log n)^2/\epsilon)$ for planar graphs with one vortex (see Lemma 16 for a formal statement). Our key idea is to "shortcut" the decomposition tree of depth $O(\log n)$ by adding edges between nodes in such a way that, for every two nodes in the tree, there is a path in the shortcut tree with $O(\log\log n)$ edges, a.k.a. $O(\log\log n)$ *hops*. To keep the treewidth of the embedding small, we require that the resulting treewidth of the decomposition tree (after adding

[3] Low-complexity means that the vortex has bounded pathwidth.

the shortcuts) remains small; the decomposition tree is a tree and hence has treewidth 1. We show that we can add shortcuts in such a way that the resulting treewidth is $O(\log\log(n))$. The two factors of $O(\log\log(n))$ — one from the hop length and one from the treewidth blowup due to shortcutting — result in treewidth of $O((\log\log n)^2/\epsilon)$ of the embedding. We formulate these ideas in terms of constructing an emulator for trees with treewidth $O(\log\log n)$ and hop-diameter $O(\log\log n)$; see Section 1.2 for a more formal discussion. The main conceptual message of our technique is that it is possible to get around the depth barrier $O(\log n)$ of balanced separators by shortcutting the decomposition tree. We belive that this idea would find further use in designing planar graph algorithms.

Applying our technique to planar graphs, we obtain an additive embedding with treewidth $O((\log\log n)^2/\epsilon)$ that can be constructed in nearly linear time.

Theorem 3 (Embedding Planar Graphs to Bounded Treewidth Graphs)**.** *Given an n-vertex planar graph $G(V,E,w)$ of diameter D and a parameter $\epsilon \in (0,1)$, there is a deterministic embedding $f : V(G) \to H$ into a graph H of treewidth $O(\epsilon^{-1}(\log\log n)^2)$ and additive distortion $+\epsilon D$. Furthermore, f can be deterministically constructed in $O(n \cdot \frac{\log^3 n}{\epsilon^2})$ time.*

Our embedding, while having a minor dependency on n, offers three advantages over the embedding of [FKS19]. First, our embedding can be constructed in nearly linear running time, removing the running time bottleneck of algorithms that uses the embedding of [FKS19]. Second, our embedding algorithm is much simpler than that of [FKS19] as it only uses the recursive shortest path separator decomposition. Third, the dependency on ϵ of the treewidth is linear, which we show to be optimal by the following theorem.

Theorem 4 (Embedding Planar Graphs Lower Bound)**.** *For any $\epsilon \in (0, \frac{1}{2})$ and any $n = \Omega(1/\epsilon^2)$, there exists an unweighted n-vertex planar graph $G(V,E)$ of diameter $D \le (1/\epsilon + 2)$ such that for any deterministic dominating embedding $f : V(G) \to H$ into a graph H with additive distortion $+\epsilon D$, the treewidth of H is $\Omega(1/\epsilon)$.*

We note that Eisenstat, Klein, and Mathieu [EKM14] implicitly constructed an embedding with additive distortion $+\epsilon D$ and treewidth $O(\epsilon^{-1} \cdot \log n)$ (see also [CFKL20]). However, in the applications of designing PTASes, the running time is at least exponential in the treewidth, and hence, this embedding only implies ineffcient PTASes or QPTASes. The dependency on n of the treewidth of our embedding in Theorem 3 is exponentially smaller.

Our new embedding result (Theorem 3) leads to an almost linear time PTAS for metric Baker's problems in planar graphs, thereby giving an affirmative answer to Question 1. See Table 2 for a comparison with existing results.

Theorem 5 (PTAS for Metric Baker Problems in Planar Graphs)**.** *Given an n-vertex planar graph $G(V,E,w)$, two parameters $\epsilon \in (0,1)$ and $\rho > 0$, and a measure $\mu : V \to \mathbb{R}^+$, one can find in $2^{\tilde{O}(1/\epsilon^{2+\kappa})} \cdot n^{1+o(1)}$ time for* any fixed $\kappa > 0$*: (1) a $(1-\epsilon)\rho$-independent set I such that for every ρ-independent set $\tilde{I}$, $\mu(I) \ge (1-\epsilon)\mu(\tilde{I})$ and (2) a $(1+\epsilon)\rho$-dominating set S such that for every ρ-dominating set $\tilde{S}$, $\mu(S) \le (1+\epsilon)\mu(\tilde{S})$.*

We could choose $\kappa = 0.01$ in Theorem 5 to get a PTAS with running time $2^{\tilde{O}(1/\epsilon^{2.01})} \cdot n^{1+o(1)}$ for the ρ-independent set/dominating set problems. The dependency on $1/\epsilon$ of the running time of our PTAS is much smaller than that of [FKS19].

Bounded-capacity vehicle routing problem

	Running time	Ref
Euclidean plane	$2^{O(\epsilon^{-3})} + O(n \log n)$	[AKTT97]
Planar graph	$n^{\mathrm{poly}(\epsilon^{-1} \log n)}$	[BKS17]
	$n^{O_\epsilon(1)}$	[BKS19]
	$O_\epsilon(1) \cdot n^{O(1)}$	[CFKL20]
	$O_\epsilon(1) \cdot n^{1+o(1)}$	Theorem 6
K_r-minor free	$n^{O_{r,\epsilon}(\log \log n)}$	[CFKL20]
	$O_{\epsilon,r}(1) \cdot n^{O(1)}$	Theorem 2

Metric ρ dominating/independent set problems

	Running time	Ref
Planar graph	$n^{O_\epsilon(1)}$	[EKM14]
	$2^{O(\epsilon^{-c+1})} \cdot n^{O(1)}$ for $c \geq 19$	[FKS19]
	$2^{\tilde{O}(\epsilon^{-(2+\kappa)})} \cdot n^{1+o(1)}$, any fixed $\kappa > 0$	Theorem 5
K_r-minor free	$n^{\tilde{O}(\epsilon^{-2} \log(n))}$	[FL21]
	$n^{\tilde{O}(\epsilon^{-2} (\log \log(n))^3)}$	Theorem 9

Table 2: *On the left illustrated the running time of all the approximation schemes for the bounded-capacity vehicle routing problem. On the right illustrated the running time of the bicriteria approximation schemes for the metric ρ dominating/independent set problems in their full generality.*

By using our new embedding in Theorem 3 and some additional ideas, we significantly improve the running time of the PTAS by Cohen-Addad *et al.* [CFKL20] to almost linear time as asked in Question 2. See Table 2 for a summary.

Theorem 6 (PTAS for Bounded-Capacity VRP in Planar Graphs)**.** *There is a randomized polynomial time approximation scheme for the bounded-capacity VRP in planar graphs that runs in time $O_\epsilon(1) \cdot n^{1+o(1)}$.*

We note that all results stated for planar graphs so far can be extended to bounded genus graphs with the same bounds using the cutting technique of [CFKL20].

In clan embeddings and Ramsey type embeddings, we observe that the construction of the authors[FL21], in combination with the construction of Cohen-Addad *et al.* [CFKL20], can be seen as providing a reduction from planar graphs with one vortex to K_r-minor-free graphs. Roughly speaking, the reduction implies that if planar graphs with one vortex have an embedding with treewidth $t(\epsilon, n)$ and additive distortion $+\epsilon D$, then K_r-minor-free graphs has a clan embedding/Ramsey type embedding with treewidth $t(O_r(\frac{\delta\epsilon}{\log(n)}), n) + O_r(\log(n))$ and additive distortion $+\epsilon D$. The authors[FL21] used the embedding of planar graphs with one vortex by Cohen-Addad *et al.* [CFKL20], which has treewidth $t(\epsilon, n) = O(\frac{\log n}{\epsilon})$, to get a clan embedding/Ramsey type embedding of treewidth $O_r(\frac{\log n}{(\delta\epsilon)/\log(n)}) + O_r(\log n) = O_r(\frac{\log^2 n}{\delta\epsilon})$.

There are two fundamental barriers if one wants to improve the treewidth bound of $O_r(\frac{\log^2 n}{\delta\epsilon})$: improving the embedding of planar graphs with one vortex and removing the (both multiplicative and additive) loss $O(\log n)$ in the reduction step. Our embedding for planar graphs with one vortex (Lemma 16) overcomes the former barrier. Specifically, by plugging in our embedding for planar graphs with one vortex, we obtain treewidth $O_r(\frac{\log n (\log \log n)^2}{\epsilon\delta})$ in both clan embedding and Ramsey type embedding; the improvement is from quadratic in $\log(n)$ to nearly linear in $\log(n)$. While we are not able to fully answer Question 3 due to the second barrier, we are one step closer to its resolution.

Theorem 7 (Clan Embeddings of Minor-free Graphs into Low Treewidth Graphs)**.** *Given a K_r-minor-free n-vertex graph $G = (V, E, w)$ of diameter D and parameters $\epsilon \in (0, \frac{1}{4})$, $\delta \in (0, 1)$, there is* *a distribution $\mathcal{D}$ over clan embeddings (f,χ) with additive distortion $+\epsilon D$ into graphs of treewidth $O_r(\frac{\log n(\log\log n)^2}{\epsilon\delta})$ such that for every $v\in V$, $\mathbb{E}[|f(v)|]\le 1+\delta$.*

Theorem 8 (Ramsey Type Embeddings of Minor-free Graphs into Low Treewidth Graphs)**.** *Given an n-vertex K_r-minor-free graph $G=(V,E,w)$ with diameter D and parameters $\epsilon\in(0,\frac{1}{4})$, $\delta\in(0,1)$, there is a distribution over dominating embeddings $g:G\to H$ into treewidth $O_r(\frac{\log n(\log\log n)^2}{\epsilon\delta})$ graphs, such that there is a subset $M\subseteq V$ of vertices for which the following claims hold:*

1. *For every $u\in V$, $\Pr[u\in M]\ge 1-\delta$.*
2. *For every $u\in M$ and $v\in V$, $d_H(g(u),g(v))\le d_G(u,v)+\epsilon D$.*

Using our new clan and Ramsey embeddings in Theorem 7 and Theorem 8 on top of the algorithm from [FL21], we improve the running time of the QPTAS for the metric Baker's problems in K_r-minor-free graphs [FL21] from $n^{\tilde{O}(\epsilon^{-2}\log(n))}$ to $n^{\tilde{O}(\epsilon^{-2}(\log\log(n))^3)}$.

Theorem 9 (QPTAS for Metric Baker Problems in Minor-free Graphs)**.** *Given an n-vertex K_r-minor-free graph $G(V,E,w)$, two parameters $\epsilon\in(0,1)$ and $\rho>0$, and a vertex weight function $\mu:V\to\mathbb{R}^+$, one can find in $n^{\tilde{O}_r(\epsilon^{-2}(\log\log(n))^3)}$ time:*
(1) a $(1-\epsilon)\rho$-independent set I such that for every ρ-independent set $\tilde{I}$, $\mu(I)\ge(1-\epsilon)\mu(\tilde{I})$, and
(2) a $(1+\epsilon)\rho$-dominating set S such that for every ρ-dominating set $\tilde{S}$, $\mu(S)\le(1+\epsilon)\mu(\tilde{S})$.

1.2 Techniques

A key technical component in all the embeddings is an emulator for trees with *low treewidth* and *low hop path* between every pair of vertices. For a path P, denote by $\mathrm{hop}(P)$ the number of edges in the path P. For a pair of vertices $u,v\in V$, and parameter $h\in\mathbb{N}$, denote by

$$d_G^{(h)}(u,v)=\min\{w(P)\mid P\text{ is a }u-v\text{ path, and }\mathrm{hop}(P)\le h\}$$

the weight of a minimum $u-v$ path with at most h edges. If no such path exist, $d_G^{(h)}(u,v)=\infty$.

Given an edge-weighted tree $T=(V_T,E_T,w_T)$, we say that an edge-weighted graph $K=(V_K,E_K,w_T)$ is an *h-hop emulator* for T if $V_K=V_T$ and $d_K^{(h)}(u,v)=d_T(u,v)$ for every pair of vertices $u,v\in V$. We show that every tree admits a low-hop emulator with small treewidth.

Theorem 10. *Given an edge-weighted n-vertex tree T, there is an $O(\log\log n)$-hop emulator K_T for T that has treewidth $O(\log\log n)$. Furthermore, K_T can be constructed in $O(n\cdot\log\log n)$ time.*

Constructing a low hop emulator for trees is a basic problem that was studied by various authors [Cha87, AS87, BTS94, Tho97, NS07, CG08, Sol11, FL22] in different contexts. The goal of these constructions is to minimize the number of edges of the emulator for a given hop bound, with no consideration to the treewidth of the resulting graph. To the best of our knowledge, the only previously constructed low-hop emulator with bounded treewidth is a folklore recursive construction [4] of a 2-hop emulator with treewidth $O(\log n)$. Our emulator seeks to balance between the treewidth and the hop bound, and achieves an exponential reduction in the treewidth (at the expense of an increase in the hop bound). Due to its fundamental nature, we believe that Theorem 10 will find more applications in other contexts.

[4] Consider the following construction: take the centroid v of the tree (a vertex v such that each connected component in $T\smallsetminus\{v\}$ has at most $\frac{2}{3}n$ vertices), add edges from v to all the vertices, and recurse on each connected component of $T\smallsetminus\{v\}$. The result is a 2-hop emulator with treewidth $O(\log n)$.

Embedding of planar graphs. To avoid notational clutter, we describe our embedding technique for planar graphs instead of addressing planar graphs with one vortex directly (which is the main motivation of our technique). We will use the low hop emulator in Theorem 10 to obtain a simple and efficient construction as claimed in Theorem 3. Our construction not only addresses Question 1, but it is also robust enough to be easily extensible to planar graphs with one vortex, which, as discussed above, is the fundamental barrier towards an embedding of K_r-minor-free graphs with low treewidth.

Given a planar graph $G = (V, E, w)$ with diameter D, we construct a recursive decomposition Φ of G using shortest path separators [Tho04, Kle02]. Specifically, we pick a global root vertex $r \in V$. Φ has a tree structure of depth $O(\log n)$, and each node $\alpha \in \Phi$ is associated with a cluster of G, the "border" of which consists of $O(1)$ shortest paths rooted at r. This set of border shortest paths is denoted by $\mathcal{Q}_\alpha$. For each node α, all the vertices not in $\mathcal{Q}_\alpha$ are called *internal vertices*, and denoted I_α. Each leaf note $\alpha \in \Phi$ contains only a constant number of internal vertices.

Following the standard techniques [EKM14, CFKL20], for each path $Q \in \mathcal{Q}_\alpha$ associated with a node $\alpha \in \Phi$, one would place a set of equally-spaced vertices of size $O(\frac{\log(n)}{\epsilon})$, called *portals*, such that the distance between any two nearby portals is $O(\epsilon D/\log(n))$. Since $|\mathcal{Q}_\alpha| = O(1)$, each node of Φ has $O(\log(n)/\epsilon)$ portals. We form a tree embedding of *treewidth* $O(\log(n)/\epsilon)$ from portals of Φ by considering every edge $\{\alpha, \beta\}$ in Φ and adding all pairwise edges between portals of α and β. For the distortion argument, we consider any shortest path P of G from u to v. Let $\Phi[\alpha_u, \alpha_v]$ be the path in Φ between a node α_u containing u and α_v containing v. As Φ has depth $O(\log n)$, $\Phi[\alpha_u, \alpha_v]$ has $O(n)$ nodes. Thus, P intersects at most $O(\log n)$ shortest paths on the boundary of $O(\log n)$ nodes in $\Phi[\alpha_u, \alpha_v]$. For any shortest path $Q \in \mathcal{Q}_\alpha$ that intersects P for $\alpha \in \Phi[\alpha_u, \alpha_v]$, we route P through the portal closest to the intersection vertex of P and Q, incurring an additive distortion of $O(\epsilon D/\log(n))$. This means that the total distortion due to routing through portals is $O(\log n) \cdot O(\epsilon D/\log(n)) = O(\epsilon)D$.

Our key idea is to use Theorem 10 to construct an $O(\log\log n)$-hop emulator K_Φ for Φ along with a tree decomposition $\mathcal{T}$ for K_Φ of width $\mathsf{tw}(\mathcal{T}) = O(\log\log n)$. For each path $Q \in \mathcal{Q}_\alpha$ in each node $\alpha \in \Phi$, we now only place $O(\frac{\log\log(n)}{\epsilon})$ portals such that the distance between any two nearby portals is $O(\epsilon D/\log\log(n))$. For each edge $\{\alpha, \beta\}$ *in* K_Φ, as opposed to only consider edges in Φ in the standard technique, we add all pairwise edges between portals of α and β to obtain the host graph H. By replacing each node $\alpha \in \Phi$ in each bag of $\mathcal{T}$ with the portals of α, we obtain a tree decomposition $\mathcal{X}_H$ of H. Since $\mathcal{T}$ has width $O(\log\log n)$ and each node has $O(\log\log(n)/\epsilon)$ portals, the treewidth of $\mathcal{X}_H$ is $O(\log\log(n)) \cdot O(\log\log(n)/\epsilon) = O((\log\log n)^2/\epsilon)$.

For the distortion argument, we crucially use the property that K_Φ has hop diameter $O(\log\log(n))$. Consider any shortest path P of G, whose endpoints are internal vertices of two leaf nodes, say α_1 and α_2, of Φ. There is a (shortest) path consisting of $k = O(\log n \log n)$ nodes from α_1 to α_2 in K_Φ. This means that it suffices to consider k shortest paths $\{Q_1, \ldots, Q_k\}$ associated with these nodes intersecting P. By the construction, there are edges in H between portals of Q_i and Q_{i+1} for every $1 \le i \le k-1$. Thus, by routing an intersection vertex of $P \cap Q_i$ to the nearest portal of Q_i for every $i \in [k]$, we obtain a path P_H in H between the endpoints of P. As the portal distance is $O(\epsilon D/\log\log(n))$, each routing step incurs $O(\epsilon D/\log\log(n))$ additive distortion. Since the number of paths is $k = O(\log\log n)$, the total distortion is $k \cdot O(\epsilon D/\log\log(n)) = O(\epsilon)D$, as desired.

Conceptually, one can think of our shortcutting technique as an effective way to reduce the recursion depth $O(\log n)$ by balanced separators to $O(\log\log(n))$. Since using balanced separators is a fundamental technique in designing planar graphs algorithms, we believe that our idea or its variants could be used to improve other algorithms where the depth $O(\log n)$ of the recursion tree is the main barrier.

Extension to K_r-minor-free graphs. Since planar graphs with one vortex have balanced shortest path separators [CFKL20], the embedding algorithm for planar graphs described above is readily applicable to these graphs; the treewidth of the embedding is $O((\log\log n)^2/\epsilon)$. If we plug this embedding into the framework of Cohen-Addad *et al.* [CFKL20], we obtain a stochastic embedding of K_r-minor-free graphs into $O_r((\log\log n)^2\epsilon^{-2}) + O_r(\log n)$ treewidth graphs. The additive $O_r(\log n)$ term is due to a reduction step involving *clique-sums* (Lemma 19 in [CFKL20]).

Our idea to remove the $+O_r(\log n)$ additive term using Robertson-Seymour decomposition is as follows. (See Section 5 for the formal definitions of concepts in this paragraph.) By Robertson-Seymour decomposition, G can be decomposed into a tree $\mathbb{T}$ of nearly embeddable graphs, called *pieces*, that are glued together via *adhesions* of constant size. Our new embedding of planar graphs with one vortex implies that nearly embeddable graphs have embeddings into treewidth $O((\log\log n)^2/\epsilon)$. To glues these embeddings together, we first root the tree $\mathbb{T}$. Then for each piece G_i, we simply add the adhesion J_0 between G_i and its parent bag G_0 to every bag in the tree decomposition of the host graph H_i in the embedding of G_i; this increases the treewidth of G_0 by only $+O(h(r))$, and the overall treewidth of the final embedding is also increased only by $+O(h(r))$. For the stretch, let Q_{uv} be a shortest path between two vertices u and v. We show by induction that for every vertex $u \in G_a$ and $v \in G_b$ for any two pieces G_a, G_b such that G_i is their lowest common ancestor, there is a vertex $x \in Q_{uv} \cap G_i$ and $y \in Q_{uv} \cap G_i$, the distances between u and x, and between v and y are preserved *exactly*. Thus, the distortion for $d_G(u,v)$ is just $+\epsilon D$. Since we apply this construction to *every* piece of $\mathbb{T}$, $+\epsilon D$ is also the distortion of the final embedding.

Bounded-capacity VRP in planar graphs. For bounded-capacity VRP in planar graphs, the main ingredient is a rooted-stochastic embedding (see Definition 4), where given a root r (the depot in the VRP problem), there is a distribution over embeddings such that every pair of vertices u, v, have an additive distortion of $\epsilon \cdot (d_G(r,u) + d_G(r,v))$ in expectation. Efficient construction of such embedding is the main bottleneck in constructing near-linear time PTAS for the bounded-capacity VRP. The stochastic embedding was introduced by Becker *et al.* [BKS19]. The main idea of their embedding is to randomly slice graphs into vertex-disjoint bands, construct an additive embedding into a bounded treewidth graph for each band, and finally combine the resulting embeddings. There are two steps in the algorithm of Becker *et al.* [BKS19] that incur running time $\Omega(n^2)$. First, the algorithm uses the additive embedding of Fox-Epstein *et al.* [FKS19], which has a large (but polynomial) running time. We improve this step by using the embedding in Theorem 3. Second, for each band B, they construct a new graph G_B by taking the union of the shortest paths of all pairs of vertices in B. Since these shortest paths may contain vertices not in B, the size of G_B could be $\Omega(n)$; as a result, the total size of all subgraphs constructed from the bands is $\Omega(n^2)$. We improve this step by showing that it suffices to embed the graph induced by the band $G[B]$ plus one more vertex, and hence the total number of vertices of all subgraphs is $O(n)$. The key idea in our proof is to observe that if a vertex falls far from the boundary of the band it belongs to, it has small additive distortion w.r.t. all other vertices. We then carefully choose the parameters so that each vertex is successful in this aspect with probability $1-\epsilon$. Interestingly, we obtain here a strong Ramsey-type guarantee: each vertex v, with probability $1-\epsilon$, has a small additive distortion w.r.t.

all other vertices (instead of a small expected distortion w.r.t. each specific vertex).

1.3 Related Work

Different types of embeddings were studied for planar and minor free graphs. K_r-minor-free graphs embed into ℓ_p space with multiplicative distortion $O_r(\log^{\min\{\frac{1}{2},\frac{1}{p}\}} n)$ [Rao99, KLMN05, AGG+19, AFGN18, Fil20a], in particular, they embed into ℓ_∞ of dimension $O_r(\log^2 n)$ with a constant multiplicative distortion. They also admit spanners with multiplicative distortion $1+\epsilon$ and $\tilde{O}_r(\epsilon^{-3})$ lightness [BLW17]. On the other hand, there are other graph families that embed well into bounded treewidth graphs. Talwar [Tal04] showed that graphs with doubling dimension d and aspect ratio Φ [5] stochastically embed into graphs with treewidth $\epsilon^{-O(d\log d)}\cdot\log^d\Phi$ and expected distortion $1+\epsilon$. Similar embeddings are known for graphs with highway dimension h [FFKP18] (into treewidth $(\log\Phi)^{-O(\log^2\frac{h}{\epsilon})}$ graphs), and graphs with correlation dimension k [CG12] (into treewidth $\tilde{O}_{k,\epsilon}(\sqrt{n})$ graphs).

Ramsey type embeddings of general graphs into ultrametrics and trees were extensively studied [BFM86, BBM06, BLMN05, MN07, NT12, BGS16, ACE+20, FL21]. Clan embeddings were also studied for trees [FL21]. Clan embeddings are somewhat similar to the previously introduced multi-embeddings [BM04, Bar21] in that each vertex is mapped to a subset of vertices. However, multi-embeddings lack the notion of chief, which is crucial in all our applications. See also [HHZ21a]. Finally, clan and Ramsey type embeddings were also studied in the context of hop-constrained metric embeddings [HHZ21b, Fil21].

2 Preliminaries

$\tilde{O}$ notation hides poly-logarithmic factors, that is $\tilde{O}(g) = O(g)\cdot\mathrm{polylog}(g)$, while O_r notation hides factors in r, e.g. $O_r(m) = O(m)\cdot f(r)$ for some function f of r. All logarithms are at base 2 (unless specified otherwise).

A (quasi-)polynomial time approximation scheme, or PTAS (QPTAS), for a problem is a family of algorithms that, given any *fixed* $\epsilon > 0$, there is an algorithm A_ϵ in the family that runs in (quasi-)polynomial time and finds a $(1+\epsilon)$-approximation for the problem. A PTAS is *efficient* if the running time of A_ϵ is of the form $f(\epsilon)n^{O(1)}$ for some function $f(\epsilon)$ that only depends on ϵ.

We consider connected undirected graphs $G = (V,E)$ with edge weights $w_G : E \to \mathbb{R}_{\geq 0}$. A graph is called unweighted if all its edges have unit weight. Additionally, we denote G's vertex set and edge set by $V(G)$ and $E(G)$, respectively. Often, we will abuse notation and write G instead of $V(G)$. Let $T(V_T,E_T)$ be a tree. For any two vertices $u,v\in V_T$, we denoted by $T[u,v]$ the unique path between u and v in T. d_G denotes the shortest path metric in G, i.e., $d_G(u,v)$ is the shortest distance between u to v in G. The *metric complement* of $G(V,E,w)$ is an edge-weighted complete graph with vertex set V and the weight of each edge is the distance between its endpoints in $G(V,E,w)$. Note that every metric space can be represented as the shortest path metric of a weighted complete graph. We will use the notions of metric spaces, and weighted graphs interchangeably. When the graph is clear from the context, we might use w to refer to w_G, and d to refer to d_G. $G[S]$ denotes the induced subgraph by S.

[5] The aspect ratio of a metric space (X,d) is the ratio between the maximum and minimum pairwise distances $\frac{\max_{x,y} d(x,y)}{\min_{x,y} d(x,y)}$.

Definition 1 (Tree decomposition). *A tree decomposition of a graph $G(V,E)$ is a tree $\mathcal{T}(\mathcal{B},\mathcal{E})$ where $\mathcal{B}$ is a collection of subset of vertices of V, called* bags*, that satisfies the following conditions:*

1. $\cup_{B\in\mathcal{B}} B = V$.
2. *For every edge $(u,v) \in E$, there is a bag $B \in \mathcal{B}$ such that $\{u,v\} \subseteq B$.*
3. *For every vertex $u \in V$, the set of bags $\{B \in \mathcal{B} : u \in B\}$ containing u induces a connected subtree of $\mathcal{T}(\mathcal{B},\mathcal{E})$.*

The *width* of a tree decomposition $\mathcal{T}$ is $\max_{i\in V(\mathcal{T})} |X_i| - 1$ and the treewidth of G, denoted by tw, is the minimum width among all possible tree decompositions of G. A *path decomposition* of a graph $G(V,E)$ is a tree decomposition where the underlying tree is a path. The pathwidth of G, denoted by pw, is defined accordingly.

A metric embedding is a function $f : X \to Y$ between the points of two metric spaces (X, d_X) and (Y, d_Y). A metric embedding f is said to be *dominating* if for every pair of points $x, y \in X$, it holds that $d_X(x,y) \le d_Y(f(x), f(y))$. We say the a metric embedding f has additive distortion $+\Delta$ if f is dominating, and for every $x, y \in X$, $d_Y(f(x), f(y)) \le d_X(x,y) + \Delta$. Stochastic embedding is a distribution $\mathcal{D}$ over dominating embeddings $f : V(G) \to V(H_f)$. Stochastic embedding $\mathcal{D}$ has expected additive distortion $+\Delta$, if for every pair of points $x, y \in X$, it holds that $\mathbb{E}_{f\sim\mathcal{D}}[d_{H_f}(f(x), f(y))] \le d_G(x,y) + \Delta$.

Bounded-Capacity Vehicle Routing Problem. A *tour* R in a graph $G(V,E,w)$ is a closed walk, and the cost of R is $\mathsf{Cost}(R) = \sum_{e\in E(R)} x_e w(e)$ where x_e is the number of times that R visits edge e.

Definition 2 (Vehicle Routing Problem (VRP)). *An instance of the vehicle routing problem (VRP) is a tuple $(G(V,E,w), K, r, Q)$ where $K \subseteq V$ is a set of* clients*, $r \in V$ is the* depot*, and $Q \in \mathbb{Z}_+$ is the* capacity *of the vehicle. A feasible solution of an instance of the VRP, denoted by $\mathcal{S}$,is a tuple $(\mathcal{K}, \mathcal{R}, q)$ where:*

- $\mathcal{K} = \{K_1, K_2, \dots, K_q\}$ *is a family of q subset of terminals such that $\cup_{j=1}^{q} K_j = K$ and $|K_j| \le Q$ for every $j \in [q]$.*

- $\mathcal{R} = \{R_1, \dots, R_q\}$ *is the set of q tours such that each tour R_j (a) starts and ends at the depot r and (2) visits all terminals in K_j for every $j \in [q]$.*

The cost of $\mathcal{S}$ is: $\mathsf{Cost}(\mathcal{S}) = \sum_{j=1}^{q} \mathsf{Cost}(R_j)$. *The goal is to find a solution with minimum cost for the VRP in $G(V,E,w)$.*

3 Spanners for Trees of Low Treewidth and Hop Diameter

In this section, we focus on proving Theorem 10 that we restate here for convenience.

Theorem 10. *Given an edge-weighted n-vertex tree T, there is an $O(\log\log n)$-hop emulator K_T for T that has treewidth $O(\log\log n)$. Furthermore, K_T can be constructed in $O(n \cdot \log\log n)$ time.*

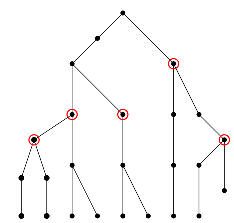
Our proof of Theorem 10 uses the following lemma, which states that in every tree, we can remove a small number of vertices, such that every remaining connected component has at most two boundary vertices. This lemma was used, sometimes implicitly, in the literature [AHLT05, FLSZ18, NPSW22]. We include a proof for completeness.

Lemma 1. *Given any parameter $\ell \in \mathbb{N}$ and an n-vertex tree T, then in $O(n)$ time we can find a subset X of at most $\frac{2}{\ell+1}n - 1 = O(\frac{n}{\ell})$ vertices such that every connected component C of $T \smallsetminus X$ is of size $|C| \le \ell$, and C has at most two outgoing edges towards X.*

Proof. We begin the proof of Lemma 1 with the following claim:

Claim 1. *Given an n-vertex tree T, in $O(n)$ time we can find a set A of at most $\frac{n}{\ell+1}$ vertices such that every connected component in $T \smallsetminus A$ contains at most ℓ vertices.*

Proof. If $n \le \ell$ then we are done. Otherwise, fix an arbitrary root r, and visit the tree in a bottom-up fashion from the leaves towards the root until we find the first vertex v with subtree T_v of size at least $\ell + 1$. We then add v to A, and continue the process on $T \smallsetminus T_v$. By induction, we will add at most $\frac{|T \smallsetminus T_v|}{\ell+1} \le \frac{n-(\ell+1)}{\ell+1} = \frac{n}{\ell+1} - 1$ additional vertices to A. It follows that $|A| \le \frac{n}{\ell+1}$ as required. A can be constructed in linear time by a simple tree traversal. In the illustration on right $\ell = 3$, and the vertices in A are encircled in red. □



We continue the proof of Lemma 1. We first apply Claim 1 to obtain a set A of size $\frac{n}{\ell+1}$ in $O(n)$ time such that every connected component of $T \smallsetminus A$ has at most ℓ vertices. The set X contains A and the least common ancestor of every pair of vertices in A. X is called the *least common ancestor closure* of A (see [FLSZ18]). Given A, X can be found in $O(n)$ time by traversing T in post order. Clearly, every connected component C of $T \smallsetminus X$ has size at most ℓ. Note that if C has outgoing edges towards $x, y \in X$, then necessarily x is an ancestor of y, or vice versa (as otherwise their least common ancestor will be in C, but this is a contradiction as $C \cap X = \emptyset$). As it is impossible that C will have three outgoing edges towards $x, y, z \in X$ where x is an ancestor of y, and y is an ancestor of z, we conclude that C has at most two outgoing edges towards X.

Finally, we argue by induction on n that $|X| \le 2|A| - 1$; the lemma will then follow as $2|A| - 1 \le \frac{2}{\ell+1}n - 1$. The base of the induction when $|A| = 1$ is clear. Consider a tree T rooted at a vertex r with children $v_1, \dots, v_s$ whose subtrees are denoted by $T_1, \dots, T_s$, respectively. First, assume that $r \notin A$. If all the vertices of A belong to a single subtree T_i, then $|X| \le 2|A| - 1$ by the induction hypothesis on T_i. Otherwise, suppose w.l.o.g. that A intersects $T_1, \dots, T_{s'}$ where $s' \ge 2$. Then X will contain the root r, along with some vertices in each subtree T_i for $i \in [1, s']$. Using the induction hypothesis we conclude that:

$$|X| \le 1 + \sum_{i=1}^{s'} (2|T_i \cap A| - 1) = 2|A| + 1 - s' \le 2|A| - 1 \ .$$

The case where $r \in X$ is similar. □

We are now ready to prove Theorem 10.

Proof of Theorem 10. Fix $c = \frac{2}{\log \frac{3}{2}}$. We will prove by induction on n, that every n-vertex tree T for $n \ge 2$ has a $(1 + c \cdot \log \log n)$-hop emulator K_T, where K_T has treewidth at most $1 + c \cdot \log \log n$. The base case is when $n = 1$; we simply return the original tree $K_T = T$ (a singleton emulator with treewidth 0). For the general case, using Lemma 1 with parameter $\ell = \sqrt{2n} - 1$,[6] we obtain a set X of at most $\frac{2}{\ell+1}n - 1 = \ell$ vertices, such that $T \smallsetminus X$ has a set of connected components $\mathcal{C}$, where every $C \in \mathcal{C}$ is of size $|C| \le \ell$, and has at most two outgoing edges towards X. Note that

[6] For the sake of simplicity, we will ignore integral issues.

for $n \geq 1$, $|X| \leq \ell = \sqrt{2n} - 1 < n^{\frac{2}{3}}$. For each such connected component C, let $T_C = T[C]$ be the induced subtree, and K_C be the emulator constructed using the induction hypothesis, with the tree decomposition $\mathcal{T}_C$. Next, we create a new tree T_X with X as a vertex set as follows. For every connected component C with two outgoing edges towards two vertices $u, v \in X$, we add an edge $\{u, v\}$ to T_X of weight $d_T(u, v)$. One can easily verify that T_X is indeed a tree (in fact it is a minor of T), and furthermore, for every two vertices $x, y \in X$, $d_{T_X}(x, y) = d_T(x, y)$. We recursively construct an emulator K_X along with a tree decomposition $\mathcal{T}_X$ for T_X.

We construct the emulator K_T for T as follows: K_T contains the emulator K_X, the union of all the emulators $\cup_{C \in \mathcal{C}} K_C$, and in addition, for every $C \in \mathcal{C}$ with outgoing edges towards $\{u, v\}$ (we allow $u = v$), and for every vertex $z \in C$, we add edges $\{z, u\}, \{z, v\}$ to K_T. (The weight of every edge added to K_T is the distance between their endpoints in T.)

First, we argue that K_T has low hop. Consider a pair of vertices u, v in T, and suppose that $u \in C^u$ and $v \in C^v$ where C^u and C^v are in $\mathcal{C}$. (The cases where either u or v is in X follows by the same argument.) If $C^v = C^u$, then by the induction hypothesis on K_{C^u}, $d_{K_T}^{(1+c \cdot \log\log n)}(u, v) \leq d_{K_{C^u}}^{(1+c \cdot \log\log |C_u|)}(u, v) = d_{T_{C^u}}(u, v) = d_T(u, v)$. Else, the unique $u - v$ path P in T must go though X. Let $x_u, x_v \in P \cap X$ be the closest vertices to u and v, respectively. Since K_T contains the edges $\{u, x_u\}$, $\{v, x_v\}$, by the induction hypothesis on K_X, we have

$$d_{K_T}^{(3+c \cdot \log\log |X|)}(u, v) \leq d_T(u, x_u) + d_{K_X}^{(1+c \cdot \log\log |X|)}(x_u, x_v) + d_T(x_v, v) = d_T(u, v) \ .$$

Observe that $3 + c \cdot \log\log |X| \leq 3 + c \cdot \log\log n^{\frac{2}{3}} = 3 + c \cdot \log \frac{2}{3} + c \cdot \log\log n = 1 + c \cdot \log\log n$. This means that $d_{K_T}^{(1+c \log\log(n))}(u, v) \leq d_{K_T}^{(3+c \cdot \log\log |X|)}(u, v) \leq d_T(u, v)$, as desired.

Next, we argue that K_T has treewidth at most $1 + c \cdot \log\log n$. We define a tree-decomposition $\mathcal{T}$ as follows: For every cluster $C \in \mathcal{C}$ with outgoing edges towards $\{u, v\}$ (we allow $u = v$), let $\mathcal{T}_C$ be the tree decomposition obtained by the induction hypothesis on $T[C]$. Let $\mathcal{T}'_C$ be the tree decomposition obtained from $\mathcal{T}_C$ by adding the vertices u, v to every bag. Consider the tree decomposition $\mathcal{T}_X$ of K_X. Since K_X contains the edge $\{u, v\}$, there is a bag $B_{u,v} \in \mathcal{T}_X$ containing both u and v. We add an edge in $\mathcal{T}$ between $B_{u,v}$ and an arbitrary bag in $\mathcal{T}_C$. It follows directly from the construction that $\mathcal{T}$ is a valid tree decomposition of K_T. Furthermore, the width of the decomposition is bounded by

$$\max\{1 + c \cdot \log\log |X|, 3 + c \cdot \log\log \ell\} \leq 1 + c \cdot \log\log n \ ,$$

as required.

Finally, we bound the running time. Note that by Lemma 1, we can find X in $O(n)$ time. Constructing the set of connected components $\mathcal{C}$ and the tree T_X takes $O(n)$ time. To find the weights for edges of T_X, we use an exact distance oracle for trees with construction time $O(n)$ and query time $O(1)$. It is folklore that such a distance oracle can be constructed by a simple reduction to the least common ancestor (LCA) data structure [BFC00] (see also [HT84, SV88]). Constructing K_T from K_X and $\{K_C\}_{C \in \mathcal{C}}$ takes $O(n)$ time as well. As the depth of the recursion is $O(\log\log n)$, and the total running time of each level is $O(n)$, the overall running time is $O(n \log\log n)$. □

4 Embedding Planar Graphs and Applications (Proof of Theorem 3)

In this section, we focus on proving Theorem 3 that we restate here for convenience.

Theorem 3 (Embedding Planar Graphs to Bounded Treewidth Graphs)**.** *Given an n-vertex planar graph $G(V,E,w)$ of diameter D and a parameter $\epsilon \in (0,1)$, there is a deterministic embedding $f : V(G) \to H$ into a graph H of treewidth $O(\epsilon^{-1}(\log\log n)^2)$ and additive distortion $+\epsilon D$. Furthermore, f can be deterministically constructed in $O(n \cdot \frac{\log^3 n}{\epsilon^2})$ time.*

Throughout, we fix a planar drawing of G. W.l.o.g, we assume that G is triangulated; otherwise, we can triangulate G in linear time and set the weight of the new edges to be $+\infty$. Let r be an (arbitrary) vertex of G and T_r be a shortest path tree rooted at r. We say that a shortest path Q in G is an *r-path* if Q is a path in T_r with r as an endpoint.

We start by defining *η-rooted shortest path decomposition.*

Definition 3 (η-rooted shortest path decomposition (η-RSPD))**.** *An* η-rooted shortest path decomposition *with root r, denoted by Φ of a graph is a binary tree with the following properties:*

(P1.) *Φ has height $O(\log n)$ and $O(n)$ nodes.*

(P2.) *Each node $\alpha \in \Phi$ is associated with a subgraph X_α of G, called a* piece*, such that:*

- *(a) X_α contains at most η shortest paths rooted at r, called* boundary paths*. Denote by $\mathcal{Q}_\alpha$ be the set of all boundary paths. We will abuse notation and denote by $\mathcal{Q}_\alpha$ the union of all the vertices in all the boundary paths. Every vertex in X_α which does not lie on a boundary path, is called an* internal vertex*.*
- *(b) If α is the root of Φ, then $X_\alpha = G$; if α is a leaf of Φ, then X_α has at most η internal vertices. Otherwise, α is an internal node with exactly two children β_1, β_2. It holds that $X_\alpha = X_{\beta_1} \cup X_{\beta_2}$ and $X_{\beta_1} \cap X_{\beta_2}$ are the boundary paths shared by X_{β_1} and $X_{\beta 2}$.*
- *(c) For any vertex $u \in V(X_\alpha)$ and $v \in V(G) \setminus V(X_\alpha)$, (any) path between u and v must intersect some vertex that lies on a boundary path of X_α.*

In planar graphs, an η-RSPD is also known as a *recursive decomposition* with shortest path separators, see, e.g., [AGK+98, Kle02, KKS11]. Abraham *et al.* [AFGN18] (see also [Fil20a]) defined a related notion of shortest path decomposition (SPD). There are several diffences between RSPD and SPD : SPD the height of the tree is a general parameter k; the number of boundary paths in each node is equal to its depth (distance to the root); the tree is not necessarily binary; and the leaves contain no internal node. SPD was used to construct multiplicative embeddings into ℓ_1 [AFGN18, Fil20a] and scattering partitions [Fil20b].

The following observation follows directly from the definition.

Observation 1. *Let Φ be an η-RSPD of G. Then $G = \cup_{\alpha \in \mathsf{leaves}(\Phi)} X_\alpha$ where $\mathsf{leaves}(\Phi)$ is the set of leaves of the tree Φ.*

For any two nodes $\alpha, \beta \in \Phi$, let $\Phi[\alpha,\beta]$ be the subpath between α and β in Φ. A crucial property of of an RSPD in our construction is the *separation property*: for any path P_{uv} between two vertices $u \in \mathcal{Q}_\alpha$ and $v \in \mathcal{Q}_\beta$ of two given nodes α and β, for any node λ on the path between α and β in the tree Φ, $P_{uv} \cap \mathcal{Q}_\lambda \neq \emptyset$.

An explicit representation of Φ could take $\Omega(n^2)$ bits and hence is not computable in nearly linear time. However, Φ has a *compact representation* that only takes $O(n\eta)$ words of space. Specifically, we store at each node $\alpha \in \Phi$ at most η vertices $B_\alpha = \{v_1, v_2, \ldots, v_{\eta'}\}$ such that $\{T_r[r, v_i]\}_{i=1}^{\eta'}$ is the set of r-paths $\mathcal{Q}_\alpha$. If α is a leaf node of Φ, then α is associated with an extra set of $O(1)$ vertices, denoted by I_α, that are internal vertices of X_α. Using shortest path separators, Thorup [Tho04] showed that one can compute a compact representation of an $O(1)$-RSPD of G in time $O(n \log n)$. (See Section 2.5 in [Tho04] for details; the RSPD is called *frame separator* decomposition of $G(V, E, w)$ in Thorup's paper.)

Lemma 2 (Thorup [Tho04]). *Given a planar graph $G(V, E, w)$, a (compact representation of a) $O(1)$-RSPD Φ of $G(V, E, w)$ can be computed in $O(n \log n)$ time.*

In Section 4.1 below, we present an embedding of $G(V, E, w)$ into a low-treewidth graph in nearly linear running time. In Section 4.2, we construct a rooted stochastic embedding for planar graphs, which is used for approximating the vehicle routing problem. In Section 4.3, we present applications of the two embeddings in designing almost linear time PTASes.

4.1 The Embedding Construction

We refer readers to Section 1.2 for an overview of the proof. We will construct a one-to-many embedding $f : V \to 2^{V(H)}$ with additive distortion $+O(\epsilon) \cdot D$; we can recover distortion ϵD by scaling ϵ. By Definition 5, we are required to guarantee that for any two copies of a single vertex $v_1, v_2 \in f(v)$, $d_H(v_1, v_2) = O(\epsilon) \cdot D$. In the end, we can transform f to a one-to-one embedding f' by picking an arbitrary copy $v' \in f(v)$ and set $f'(v) = v'$ for each vertex $v \in V(G)$.

Recall that T_r is a shortest path tree rooted at r of $G(V, E, w)$. Since $G(V, E, w)$ has diameter D, T_r has radius D. Let $\delta = \frac{\epsilon D}{\log \log n}$. We first define a set of vertices N_r called *δ-portals* of T_r as follows: initially, N_r only contains r; we then visit every vertex of T_r in the depth-first order, and we add a vertex v to N_r if the nearest ancestor of v in N_r is at a distance larger than δ from v. For each r-path Q in T_r, we denote by $N(Q, \delta) = N_r \cap V(Q)$ the set of δ-portals in Q. Note that for every vertex $v \in Q$, there is a δ-portal $u \in N(Q, \delta)$ at a distance at most δ. Furthermore, the distance between a pair of consecutive δ-portals of Q is greater than δ. As the length of Q is bounded by D, we conclude that

$$|N(Q, \delta)| \le \left\lfloor \frac{D}{\delta} \right\rfloor \le \frac{\log \log n}{\epsilon} \ .$$

The construction works as follows: let Φ be an $O(1)$-RSPD of $G(V, E, w)$ given by Lemma 2. For every node $\alpha \in \Phi$, let $Q_1, Q_2 \ldots,$ be the $\eta = O(1)$ r-paths constituting $\mathcal{Q}_\alpha$. Denote by $P_\alpha = \cup_{Q \in \mathcal{Q}_\alpha} N(Q, \delta)$ the union of all δ-portals from all the r-paths on the boundary of α. Note that $|P_\alpha| = O(\eta \cdot \frac{\log \log n}{\epsilon}) = O(\frac{\log \log n}{\epsilon})$. Using Theorem 10, we construct an $O(\log \log n)$-hop emulator K_Φ for Φ with treewidth $O(\log \log n)$. We add the following sets of edges to the host graph H:

1. For every node $\alpha \in \Phi$, and vertex $v \in P_\alpha$, create a copy $\widetilde{v_\alpha}$. Denote this set of copies by $\widetilde{P_\alpha}$. Add the edge set $\widetilde{P_\alpha} \times \widetilde{P_\alpha}$ to H (i.e. a clique on the copies).

2. For every edge $\{\alpha, \beta\} \in K_\Phi$, add the edge set $\widetilde{P_\alpha} \times \widetilde{P_\beta}$ to H (i.e. a bi-clique between the respective copies).

3. For every leaf node $\alpha \in \Phi$, let I_α be the set of internal vertices. We add to H the vertices in I_α and two edge sets $I_\alpha \times I_\alpha$ and $I_\alpha \times \widetilde{P_\alpha}$.

4. Denote by $I_\Phi = \cup_{\alpha \in \Phi : \alpha \text{ is leaf}} I_\alpha$ the set of all vertices that belong to the interior of leaves of Φ. For $v \notin I_\Phi$, by Observation 1, it necessarily belongs to some r-path Q_v on the boundary of some leaf node α_v. We add v and the edge set $v \times \widetilde{P_{\alpha_v}}$ to H.

In summary, the set of vertices of H is $V_H = V_G \cup \bigcup_{\alpha\in\Phi} \widetilde{P_\alpha}$, while the set of edges is

$$E_H = \bigcup_{\alpha\in\Phi} \{(\widetilde{v_\alpha}, \widetilde{u_\alpha}) \mid v, u \in P_\alpha\} \cup \bigcup_{(\alpha,\beta)\in\Phi} \{(\widetilde{v_\alpha}, \widetilde{u_\beta}) \mid v \in P_\alpha, u \in P_\beta\}$$

$$\cup \bigcup_{\alpha\in\Phi:\alpha \text{ is leaf}} \{(v, \widetilde{u_\alpha}) \mid v \in I_\alpha, u \in P_\alpha \cup I_\alpha\} \cup \bigcup_{v\notin I_\Phi} \{(v, \widetilde{u_{\alpha_v}}) \mid u \in P_{\alpha_v}\}$$

The image $f(v)$ of every vertex v consists of v itself (added to H in either Step 3 or Step 4), and all the respective copies (added in Step 1, one per each node α such that $v \in P_\alpha$). The weights of the edges in H are defined in a natural way: for every edge (u', v') in H, $w_H(u', v') = d_G(u, v)$ where $u = f^{-1}(u')$ and $v = f^{-1}(v')$.

Observation 2. *Foor any leaf α in Φ, the image of $I_\alpha \cup P_\alpha$ induces a clique in H.*

Next, we bound the treewidth of H.

Lemma 3. *H has treewidth $O(\epsilon^{-1}(\log\log n)^2)$.*

Proof. Let $\mathcal{T}_\Phi(\mathcal{X}_\Phi, \mathcal{E}_\Phi)$ be a tree decomposition of K_Φ of width $O(\log\log n)$. We create a tree-decomposition $\mathcal{T}$ for H as follows: for every node $\alpha \in \Phi$, and every bag in $\mathcal{T}_\Phi$ containing α, we replace α with $\widetilde{P_\alpha}$. Next, for leaf node α, let B be some arbitrary bag containing α. We create a new bag B_α, containing the vertices $I_\alpha \cup \widetilde{P_\alpha}$, with a single edge in $\mathcal{T}$ between B and B_α. Finally, for every vertex $v \notin I_\Phi$, let B_{α_v} be some bag containing α_v, the leaf node containing v. We create a new bag B_v containing the vertex v and the set $\widetilde{P_\alpha}$. We then add a single edge in $\mathcal{T}$ between B and B_{α_v}.

As each bag contains at most $O(\log\log n)$ nodes from Φ, and for every $\alpha \in \Phi$, $|\widetilde{P_\alpha}| = O(\epsilon^{-1} \cdot \log\log n)$, it follows that the width of $\mathcal{T}$ is bounded by $O(\frac{(\log\log n)^2}{\epsilon})$. Finally, it is straightforward to verify that $\mathcal{T}$ is a valid tree decomposition for H: every edge of H is contained in some bag, and as every copy of every vertex v is associated with a unique node $\alpha \in \Phi$, the connectivity condition is satisfied. □

We now bound the distortion of the embedding. Specifically, we will show that for every pair of vertices u, v, and every two copies $u' \in f(u)$, $v' \in f(v)$, $d_G(u, v) \le d_H(u', v') \le d_G(u, v) + O(\epsilon) \cdot D$. The lower bound follows directly from the way we assign weights to the edges of H. Indeed, every u'-v' path in H corresponds to a u-v walk in G of the same weight. Henceforth, we focus on proving the upper bound. First, we need the following lemma, which is the generalization of the separation property.

Lemma 4. *Let α and β be two nodes in Φ. Let P_{uv} be any path between two vertices u and v in G such that $u \in \mathcal{Q}_\alpha$ and $v \in \mathcal{Q}_\beta$. Let $(\alpha = \lambda_1, \lambda_2, \ldots, \lambda_k = \beta)$ be a set of nodes in $\Phi[\alpha, \beta]$ such that $\lambda_{i+1} \in \Phi[\lambda_i, \beta]$ for any $1 \le i \le k-1$. Then, there exists a sequence of vertices $(u = x_1, x_2, \ldots, x_k = v)$ such that $x_i \in P_{uv} \cap \mathcal{Q}_{\lambda_i}$ and $x_{i+1} \in P_{uv}[x_i, v]$ for any $1 \le i \le k-1$.*

Proof. The proof is by induction. If $k \le 2$, then the lemma trivially follows. Henceforth, we assume that $k \ge 3$. Observe that λ_2 is either an ancestor of λ_1 and/or λ_k. Assume first that λ_2 is an ancestor of λ_1. Then by property P2(b) in Definition 3, $u \in X_{\lambda_2}$ and v either belongs to $\mathcal{Q}_{\lambda_2}$ or $v \in V(G) \setminus V(X_{\lambda_2})$. In both cases, $P_{uv} \cap \mathcal{Q}_{\lambda_2} \neq \emptyset$. Let x_2 be the last vertex in the path P_{uv} going from u to v such that $x_2 \in P_{uv} \cap \mathcal{Q}_{\lambda_2}$. By applying the induction hypothesis to $P_{x_2 v} = P_{uv}[x_2, v]$, there exists a sequence of vertices $(x_2, \ldots, x_k = v)$ such that $x_i \in P_{x_2 v} \cap \mathcal{Q}_{\lambda_i}$ and $x_{i+1} \in P_{x_2 v}[x_i, v]$ for any $2 \le i \le k-1$. Thus, $(u = x_1, x_2, \ldots, x_k = v)$ is the sequence of vertices claimed by the lemma. The case where λ_2 is an ancestor of λ_k follows by the same argument. □

Next, we show that the distortion of portal vertices on the boundaries of nodes in Φ is in check.

Lemma 5. *Let α and β be two nodes in Φ. Let $\widetilde{u_\alpha} \in \widetilde{P_\alpha}$ and $\widetilde{v_\beta} \in \widetilde{P_\beta}$ be two vertices in H, and $u = f^{-1}(\widetilde{u_\alpha}), v = f^{-1}(\widetilde{v_\beta})$. Then, it holds that:*

$$d_H(\widetilde{u_\alpha}, \widetilde{v_\beta}) \leq d_G(u,v) + O(\epsilon) \cdot D.$$

Proof. Let Q_{uv} be a shortest path from u to v in G. Recall that K_Φ is an emulator of Φ with hop diameter $O(\log\log n)$. Let $P = (\alpha = \lambda_1, \lambda_2, \dots, \lambda_k = \beta)$ be a shortest path from α to β in K_Φ such that $k = O(\log\log n)$. Since K_Φ preserves distances between nodes of Φ, i.e, $w_{K_\Phi}(P) = w_\Phi(\Phi[\alpha,\beta])$, it must be that the nodes on P constitute a subsequence of nodes on $\Phi[\alpha,\beta]$. By Lemma 4, there exists a sequence of vertices $(u = x_1, x_2, \dots, x_k = v)$ such that $x_i \in \mathcal{Q}_{\lambda_i} \cap Q_{uv}$ for all $1 \leq i \leq k$ and:

$$d_G(u,v) = \sum_{i=1}^{k-1} d_G(x_i, x_{i+1}) \tag{2}$$

For every $i \in [k]$, let $y_i \in P_{\lambda_i}$ be the δ-portal closest to x_i; note that $y_1 = u$ and $y_k = v$ because u and v are δ-portals. By the definition of δ-portals, $d_G(x_i, y_i) \leq \delta$. Thus, by the triangle inequality, it holds that:

$$d_G(y_i, y_{i+1}) \leq d_G(x_i, x_{i+1}) + 2\delta \tag{3}$$

Recall that $\widetilde{(y_i)_{\lambda_i}}$ be the copy of y_i created for λ_i. Observe that, since $(\lambda_i, \lambda_{i+1})$ is an edge in K_Φ, by construction, there is an edge between $\widetilde{(y_i)_{\lambda_i}}$ and $\widetilde{(y_{i+1})_{\lambda_{i+1}}}$ of weight $w_H(\widetilde{(y_i)_{\lambda_i}}, \widetilde{(y_{i+1})_{\lambda_{i+1}}}) = d_G(y_i, y_{i+1})$. Thus, we have:

$$\begin{aligned}
d_H(\widetilde{u_\alpha}, \widetilde{v_\beta}) = d_H(\widetilde{(y_1)_{\lambda_1}}, \widetilde{(y_k)_{\lambda_k}}) &\leq \sum_{i=1}^{k-1} d_H(\widetilde{(y_i)_{\lambda_i}}, \widetilde{(y_{i+1})_{\lambda_{i+1}}}) = \sum_{i=1}^{k-1} d_G(y_i, y_{i+1}) \\
&\leq \sum_{i=1}^{k-1} (d_G(x_i, x_{i+1}) + 2\delta) \qquad \text{(by Equation (3))} \\
&= d_G(u,v) + 2(k-1)\delta \qquad \text{(by Equation (2))} \\
&= d_G(u,v) + O(\log\log n)\frac{\epsilon D}{\log\log n} = d_G(u,v) + O(\epsilon)D \ ,
\end{aligned}$$

as claimed. ☐

Lemma 6. *For any $u, v \in V(G)$ and any $u' \in f(u), v' \in f(v)$, it holds that:*

$$d_H(u', v') \leq d_G(u,v) + O(\epsilon)D.$$

Proof. The proof is by case analysis. If both u, v are δ-portals, then the lemma holds by Lemma 5. Next, suppose that both u, v are not δ-portals. By Observation 1, there are two leaf nodes $\alpha, \beta \in \Phi$, such that $u \in X_\alpha$, and $v \in X_\beta$. In particular, there are unique copies of both u and v, one for each vertex. If $\alpha = \beta$, then H contains the edge (u,v) and we are done. Otherwise, let P_{uv} be a shortest u-v path in G. Then it follows from Definition 3 that there are two vertices x_u and x_v in P_{uv} such that $x_u \in \mathcal{Q}_\alpha$ and $x_v \in \mathcal{Q}_\beta$ (it might be that $u = x_u$ or $v = x_v$). Furthermore, there are δ-portals $y_u \in P_\alpha$, and $y_v \in P_\beta$, at distances at most δ from x_u and x_v respectively. By construction in Step

2 and Step 4 (depending on whether u, v are leaves), H contains the edges $(u, \widetilde{(y_u)_\alpha})$, $(v, \widetilde{(y_v)_\beta})$. By Lemma 5, we have:

$$\begin{aligned}
d_H(u,v) &\leq d_H(u, \widetilde{(y_u)_\alpha}) + d_H(\widetilde{(y_u)_\alpha}, \widetilde{(y_v)_\beta}) + d_H(\widetilde{(y_v)_\beta}, v) \\
&\leq d_G(u, y_u) + d_G(y_u, y_v) + O(\epsilon) \cdot D + d_G(y_v, v) \\
&\leq d_G(u, x_u) + 2 \cdot d_G(x_u, y_u) + d_G(x_u, x_v) + 2 \cdot d_G(x_v, y_v) + d_G(x_v, v) + O(\epsilon) \cdot D \\
&\leq d_G(u,v) + 4\delta + O(\epsilon) \cdot D \leq d_G(u,v) + O(\epsilon) \cdot D \ .
\end{aligned}$$

Finally, the cases where either u or v is a δ-portal are simpler and can be proved by the same argument. □

We now compute the runtime of our algorithm. First, we compute a (compact representation of) $O(1)$-RSPD Φ using Lemma 2 in $O(n \log n)$ time. Next, we use Theorem 10 to compute an emulator K_Φ in $O(n \cdot \log\log n)$ time, with $O(|\Phi| \cdot \log\log|\Phi|) = O(n \cdot \log\log n)$ edges. Next, we compute the set of δ-portals. We will use the same shortest path tree T_r rooted at r from Lemma 2 (that can be found in $O(n)$ time [HKRS97]). For each vertex $v \in T_r$, we compute and store a set of δ-portals $N(T_r[r,v], \delta)$ of the r-rooted shortest path $T_r[r,v]$ by depth-first tree traversal in $O(n)$ time. Note that we can afford to store all the portal of the path $T_r[r,v]$ at v explicitly since $|N(T_r[r,v],\delta)| \leq \frac{\log\log n}{\epsilon}$. Given a vertex v with a parent x_v in T_r, and the closest δ-portal ancestor y_v, if $d_{T_r}(v, y_v) = d_{T_r}(r,v) - d_{T_r}(r, y_v) \leq \delta$ then we set $N(T_r[r,v],\delta) = N(T_r[r,x_v],\delta)$, and otherwise set $N(T_r[r,v],\delta) = N(T_r[r,x_v],\delta) \cup \{v\}$.

Next, we construct the one-to-many embedding f into H with tree-decomposition $\mathcal{T}$ following to Steps (1)-(4). The construction is straightforward and takes

$$|\Phi| \cdot O(1)^2 \cdot O\left(\frac{\log\log n}{\epsilon}\right)^2 + |E(K_\Phi)| \cdot O\left(\frac{\log\log n}{\epsilon}\right)^2 + O(n) \cdot O\left(\frac{\log\log n}{\epsilon}\right) = O\left(n \cdot \frac{(\log\log n)^3}{\epsilon^2}\right) \ ,$$

time. The last and most time-consuming step is to assign weights to the edges in H. Recall that the weight of an edge in $(u', v') \in H$ where $u = f^{-1}(u')$ and $v = f^{-1}(v')$, is defined to be $w_H(u',v') = d_G(u,v)$. Computing the shortest distance between u and v in G takes $\Omega(n)$ time, and using shortest path computation to find the weight of every edge in H incurs $\Omega(n^2)$ time. Our solution is to use approximate distances instead of the exact ones to assign to edges of H. Specifically, we will use the *approximate distance oracle* of Thorup[7] [Tho04] to query the approximate distance between u and v in $O(\frac{1}{\epsilon})$ time.

Lemma 7 (Thorup [Tho04], Theorem 3.19)**.** *Given an n-vertex planar graph $G(V,E,w)$ and a parameter $\epsilon < 1$, one can construct an $O(\frac{n \log n}{\epsilon})$-space data structure $\mathcal{O}_{G,\epsilon}$ in $O(\frac{n \log^3(n)}{\epsilon^2})$ time such that given any two vertices u, v in G, the data structure, in $O(\frac{1}{\epsilon})$ time, returns $\mathcal{O}_{G,\epsilon}(u,v)$ with*

$$d_G(u,v) \leq \mathcal{O}_{G,\epsilon}(u,v) \leq (1+\epsilon) d_G(u,v)$$

To assign weights in our graph H, we simply construct a $(1+\epsilon)$-distance oracle $\mathcal{O}_{G,\epsilon}(u,v)$ for G using Lemma 7, and then for every edge $(u',v') \in H$ where $u = f^{-1}(u')$ and $v = f^{-1}(v')$, we set $w_H(u',v') = \mathcal{O}_{G,\epsilon}(u,v)$. Thus the total time for assigning the weights is $O(\frac{n \log^3(n)}{\epsilon^2}) + |E(H)| \cdot O(\frac{1}{\epsilon}) = O(\frac{n \log^3(n)}{\epsilon^2})$, which is also the overall running time of our algorithm.

[7] A similar distance oracle was constructed independently by Klein [Kle02]. However, Klein did not specify the construction time explicitly (a crucial property in our context).

Denote the graph constructed using the approximated distances by $\hat{H}$. It remains to argue that $\hat{H}$ has additive distortion $+O(\epsilon)\cdot D$. Note that for every $(u',v')\in H$, $w_H(u',v')\le w_{\hat{H}}(u',v')\le(1+\epsilon)\cdot w_H(u',v')$. This implies that for every $u',v'\in V(H)$, $d_H(u',v')\le d_{\hat{H}}(u',v')\le(1+\epsilon)\cdot d_H(u',v')$. As H has additive distortion $+O(\epsilon)\cdot D$, it follows that, for every $u',v'\in V(H)$ where $u=f^{-1}(u')$ and $v=f^{-1}(v')$,

$$\begin{aligned} d_G(u,v)\le d_H(u',v')\le\quad d_{\hat{H}}(u',v') &\le(1+\epsilon)\cdot d_H(u',v')\\ &\le(1+\epsilon)\cdot(d_G(u,v)+O(\epsilon)\cdot D)=d_G(u,v)+O(\epsilon)\cdot D\ , \end{aligned}$$

where the last equality holds as $\epsilon\cdot d_G(u,v)\le\epsilon\cdot D$. Theorem 3 now follows.

4.2 Rooted Stochastic Embedding

The main tool we use to design a PTAS for the bounded-capacity VRP is a *rooted stochastic embedding*.

Definition 4 ((r,ρ)-rooted stochastic embedding)**.** *Given a vertex $r\in V(G)$ and a parameter $\rho>0$, we say that a dominating stochastic embedding $f:V(G)\to V(H)$ of a graph G into a distribution over graphs H is an (r,ρ)*-rooted stochastic embedding *if*

$$\forall u,v\in V,\qquad \mathbb{E}[d_H(f(u),f(v))]\le d_G(u,v)+\rho\cdot(d_G(r,u)+d_G(r,v))\ .$$

We call ρ the distortion parameter. Using Theorem 3, we obtain the following lemma.

Lemma 8. *Let $G(V,E,w)$ be an n-vertex planar graph and r be a distinguished vertex in V. For any given parameter $\epsilon\in(0,\frac{1}{4})$, we can construct in $O_\epsilon(n\log^3(n))$ time an (r,ϵ)-rooted stochastic embedding of G into graphs H of treewidth $O_\epsilon(\log\log n)^2$.*

Proof. We assume w.l.o.g. that the minimum distance in G is 1. We construct an embedding f into a graph H such that for every $u,v\in G$, $\mathbb{E}[d_H(f(u),f(v))]\le d_G(u,v)+O(\epsilon)\cdot(d_G(r,u)+d_G(r,v))$; we can scale ϵ to get back distortion $\epsilon\cdot(d_G(r,u)+d_G(r,v))$. We begin by randomly "slicing" G into a collection of *bands* $\mathcal{B}=\{B_0,B_1,\ldots\}$ as follows. Choose a random $x\in[0,1]$. For every $i\ge0$, let $U_i=(\frac{1}{\epsilon})^{\frac{i+x}{\epsilon}}$, and $L_i=(\frac{1}{\epsilon})^{\frac{i-1+x}{\epsilon}}$. We then define $B_i=\{u: L_i\le d_G(r,u)<U_i\}$ (note that $L_0<1$, and hence every vertex $v\ne r$ belong some B_i). Let $G_0=G[B_0]$, and for $i\ge1$, let G_i be the graph obtained from G by removing every vertex of distance at least U_i from r and contracting every vertex of distance strictly less than L_i from r into r. For every vertex v which is the neighbor of r in G_i, the weight of the edge is defined to be $d_G(r,v)$, while all the other weights are the same as in G. Note that for every $v\in B_i$, $d_{G_i}(v,r)=d_G(v,r)$. In particular, G_i has diameter at most $2\cdot U_i$. Observe that graphs $\{G_i\}_i$ can be constructed in total $O(n)$ time as we can construct them iteratively; each edge is contracted at most one time. All distances from r can be computed in $O(n)$ time by applying a linear time single source shortest path algorithm [HKRS97].

For every $i\ge0$, we use Theorem 3 with stretch parameter $\delta=(\frac{1}{\epsilon})^{-\frac{1}{\epsilon}}=\epsilon^{\frac{1}{\epsilon}}$ to embed G_i into a graph H_i with treewidth $O(\delta^{-1}(\log\log n)^2)=O_\epsilon(\log\log n)^2$ and additive distortion $\delta\cdot U_i$. Finally, we construct a single graph H by identifying the (image of the) vertex r in all the graphs H_i; the embedding f is defined accordingly. In addition, for every vertex v, we add an edge from $f(r)$ to $f(v)$ of weight $d_G(r,v)$. Note that this can increase the overall treewidth by at most 1. It follows that H has treewidth $O_\epsilon(\log\log n)^2$. The total running time is bounded by $\sum_{i\ge0}O(|B_i|\cdot\frac{\log^3 n}{\delta^2})=O_\epsilon(n\cdot\log^3 n)$.

Finally, we bound the expected distortion. We will actually show a stronger distortion guarantee: a "Ramsey-type" property (see e.g. [BLMN05, MN07, ACE+20, FL21, Fil21]). Specifically, we will show that with probability $1-2\epsilon$, a vertex v has small distortion w.r.t. all other vertices. We say that the embedding is *successful* for a vertex v if there is some index i such that $d(r,v) \in [\epsilon^{-1}\cdot L_i, \epsilon\cdot U_i]$. To analyze the probability that v is successful, let $a \in [0,1)$ be some number such that $d(r,v) = (\frac{1}{\epsilon})^{\frac{i_v+a}{\epsilon}}$ for some integer $i_v \in \mathbb{N}$. Then v is successful if $(\frac{1}{\epsilon})^{\frac{i_v+a}{\epsilon}} \in [(\frac{1}{\epsilon})^{\frac{i-1+x}{\epsilon}+1}, (\frac{1}{\epsilon})^{\frac{i+x}{\epsilon}-1}]$ or equivalently, $i_v + a \in [i-1+x+\epsilon, i+x-\epsilon]$ for some index i. Hence

$$\Pr[v \text{ is unsuccessful}] = \Pr[\nexists\, i \text{ s.t. } i+x \in [i_v+a+\epsilon, i_v+a+1-\epsilon]] = 2\epsilon \ .$$

Next, suppose that a vertex u is successful, and let v be any other vertex. Let i be the integer such that $u \in B_i$, and let $P_{u,v}$ be a shortest u-v path. We consider three cases:

- If P is fully contained in B_i, then $d_{G_i}(u,v) = d_G(u,v)$, implying that

$$\begin{aligned} d_H(f(u),f(v)) &\le d_{H_i}(f(u),f(v)) \le d_{G_i}(u,v) + \delta\cdot 2U_i \\ &\le d_G(u,v) + L_i \le d_G(u,v) + 2\epsilon\cdot d_G(r,u) \ . \end{aligned} \tag{4}$$

- Else, if P contains a vertex $z \in B_{i'}$ for $i' > i$, then

$$\begin{aligned} d_G(r,u) &\le \epsilon\cdot U_i \le \epsilon\cdot d_G(r,z) \le \epsilon\cdot(d_G(r,u) + d_G(u,z)) \\ &\le \epsilon\cdot(d_G(r,u) + d_G(u,v)) \le \epsilon\cdot(2\cdot d_G(r,u) + d_G(r,v)) \ , \end{aligned}$$

  which implies that $d_G(r,u) \le \frac{\epsilon}{1-2\epsilon}\cdot d_G(r,v)$. Thus,

$$d_H(v,u) \le d_G(v,r) + d_G(r,u) \le d_G(v,u) + 2\cdot d_G(r,u) \le d_G(v,u) + \frac{2\epsilon}{1-2\epsilon}\cdot d_G(r,v) \ .$$

- Else, P is contained in $\cup_{i'\le i} B_{i'}$, and it contains at least one vertex $z \in B_{i'}$ for $i' < i$. Then $d_G(r,z) \le L_i \le \epsilon\cdot d_G(r,u)$. In particular

$$d_H(v,u) \le d_G(v,r) + d_G(u,r) \le d_G(v,z) + d_G(z,u) + 2\cdot d_G(r,z) \le d_G(v,u) + 2\cdot\epsilon\cdot d_G(r,u) \ .$$

We conclude that if u is successful, then $d_H(v,u) \le d_G(v,u) + O(\epsilon)\cdot(d_G(r,v) + d_G(r,u))$. Otherwise, $d_H(v,u) \le d_G(v,r) + d_G(r,u)$. As the probability of being unsuccessful is bounded by 2ϵ, we conclude that:

$$\begin{aligned} \mathbb{E}[d_H(f(u),f(v))] &\le \Pr[v \text{ is successful}]\cdot(d_G(v,u) + O(\epsilon)\cdot(d_G(r,v) + d_G(r,u))) \\ &\quad + \Pr[v \text{ is unsuccessful}]\cdot(d_G(v,r) + d_G(r,u)) \\ &\le d_G(v,u) + O(\epsilon)\cdot(d_G(r,v) + d_G(r,u)) \ . \end{aligned}$$

□

4.3 Applications (Proofs of Theorems 5 and 6)

4.3.1 Metric Baker Problems

In this section, we prove Theorem 5 that we restate below.

Theorem 5 (PTAS for Metric Baker Problems in Planar Graphs)**.** *Given an n-vertex planar graph $G(V,E,w)$, two parameters $\epsilon \in (0,1)$ and $\rho > 0$, and a measure $\mu : V \to \mathbb{R}^+$, one can find in $2^{\tilde{O}(1/\epsilon^{2+\kappa})} \cdot n^{1+o(1)}$ time for* any fixed *$\kappa > 0$: (1) a $(1-\epsilon)\rho$-independent set I such that for every ρ-independent set $\tilde{I}$, $\mu(I) \geq (1-\epsilon)\mu(\tilde{I})$ and (2) a $(1+\epsilon)\rho$-dominating set S such that for every ρ-dominating set $\tilde{S}$, $\mu(S) \leq (1+\epsilon)\mu(\tilde{S})$.*

Katsikarelis, Lampis, and Paschos [KLP19, KLP20] showed how to solve metric Baker problems in bounded treewidth graphs. Here we consider a slightly more general version where we are given a set of terminal $K \subseteq V(G)$ and we want to find a ρ-independent/dominating set for K. That is, for the ρ-independent set problem, the solution I must be a subset of K, and for the ρ-dominating set problem, we only require that the vertices in K be dominated (that is every vertex in K will be a at distance at most ρ from some vertex in the dominating set).

Lemma 9 (Theorem 21 [KLP20] and Theorem 31 [KLP19])**.** *Given an n-vertex graph $G(V,E,w)$ of treewidth tw, a terminal set $K \subseteq V(G)$, two parameters $\epsilon \in (0,1)$ and $\rho > 0$, and measure $\mu : V \to \mathbb{R}^+$, one can find in $n \cdot (\frac{\log n}{\epsilon})^{O(\mathsf{tw})}$ time:*

- *A $(1-\epsilon)\rho$-independent set $I \subseteq K$ such that for every ρ-independent set $\tilde{I} \subseteq K$, $\mu(I) \geq \mu(\tilde{I})$.*
- *A $(1+\epsilon)\rho$-dominating set S for K such that for every ρ-dominating set $\tilde{S}$ for K, $\mu(S) \leq \mu(\tilde{S})$.*

We note that the running time of Lemma 9 in the work of Katsikarelis *et al.* [KLP19, KLP20] is stated as $O^*((\frac{\mathsf{tw}}{\epsilon})^{O(\mathsf{tw})})$ where O^* notation hides the polynomial dependency on n; the main focus of their paper is to minimize the exponential dependency on the treewidth. The precise running time of their algorithms is $O((\frac{\log n}{\epsilon})^{O(\mathsf{tw})} n)$ (M. Lampis, personal communication, 2021); see page 18 in [KLP20] and page 114 in [KLP19]. While the algorithm is for ρ-independent/dominating set with uniform measure, the same algorithm works for ρ-independent/dominating set with non-uniform measure by simply storing the measure of the ρ-independent set corresponding to each configuration of the dynamic programming table.

We now focus on proving Theorem 5. By scaling every edge of $G(V,E,w)$, we assume that $\rho = 1$. For a cleaner presentation, we assume that $\frac{1}{\epsilon}$ is an integer. Let T_r be a shortest path tree rooted at a vertex r; T_r can be found in $O(n)$ time [HKRS97].

Similar to Fox-Epstein *et al.* [FKS19], we follow the classical layering technique of Baker [Bak94] to reduce to the case where the planar graph has diameter $O(1/\epsilon)$. We then embed the planar graph to a graph with treewidth $O(\frac{(\log\log n)^2}{\epsilon^2})$ and additive distortion ϵ. The additive distortion ϵ translates into $(1 \pm \epsilon)$ distance separation in the bicriteria PTAS. Next, we apply Lemma 9 to find a bicriteria PTAS for metric Baker problems in treewidth-$O(\frac{(\log\log n)^2}{\epsilon^2})$ graphs. Finally, we lift the solution to the solution of the input planar graphs.

In the analysis below, we use Young's inequality [You12]:

Theorem 11 (Young's Inequality)**.** *Given four real numbers $a \geq 0, b \geq 0, p > 1, q > 1$ such that $\frac{1}{p} + \frac{1}{q} = 1$, the following inequality holds:*

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

The equality holds if and only if $a^p = b^q$.

ρ-Independent Set Problem. Let σ be a number in $\{0, \ldots, 2/\epsilon - 1\}$. Let $\mathcal{I}_{-1,\sigma} = [0, \sigma)$, $\mathcal{I}^-_{-1,\sigma} = [0, \sigma - 1)$, and for each $j \geq 0$, we define:

$$\mathcal{I}_{j,\sigma} = [\frac{2}{\epsilon} \cdot j + \sigma \ , \ \frac{2}{\epsilon} \cdot (j+1) + \sigma] \qquad \mathcal{I}^-_{j,\sigma} = [\frac{2}{\epsilon} \cdot j + \sigma + 1 \ , \ \frac{2}{\epsilon} \cdot (j+1) + \sigma - 1]$$

Let $U_{j,\sigma} = \{v \in V : d_G(r, v) \in \mathcal{I}_{j,\sigma}\}$ and $U^-_{j,\sigma} = \{v \in V : d_G(r, v) \in \mathcal{I}^-_{j,\sigma}\}$ for every $j \geq -1$. Let $V_\sigma = \cup_{j=-1}^{\infty} U_{j,\sigma}$ and $V^-_\sigma = \cup_{j=-1}^{\infty} U^-_{j,\sigma}$. We observe that $V_\sigma = V$ for every σ and that:

Observation 3. *Each vertex $v \in V$ is contained in at least $\frac{2}{\epsilon} - 2$ sets in the collection $\{V^-_\sigma\}_{\sigma=0}^{\frac{2}{\epsilon}-1}$.*

Fix an index $j \geq -1$. Let $G_{j,\sigma}$ be defined as follows. For $j = -1$ set $G_{-1,\sigma} = G[U_{-1,\sigma}]$. In general, $G_{j,\sigma}$ is the graph obtained from G by removing every vertex of distance at least $\frac{2(j+1)}{\epsilon} + \sigma$ from r and contracting every vertex of distance strictly less than $\frac{2j}{\epsilon} + \sigma$ to r; we set weight 1 to every edge incident to r. Observe that $G_{j,\sigma}$ is a planar graph with a diameter at most $\frac{4}{\epsilon} + 2$. Furthermore, $V(G_{j,\sigma}) = U_{j,\sigma} \cup \{r\}$ when $j \geq 0$ and $V(G_{-1,\sigma}) = U_{-1,\sigma}$. Note that $G_{j,\sigma}$ can be constructed for all j in total $O(n)$ time as we can construct them iteratively; each edge is contracted only once.

Next, we find a $(1-\epsilon)$-independent set $I_{j,\sigma}$ in $G_{j,\sigma}$. Fix $\delta = \frac{\epsilon^2}{12}$, using Theorem 3 we construct an embedding $f_{j,\sigma}$ of $G_{j,\sigma}$ into a graph $H_{j,\sigma}$ with treewidth $O(\frac{(\log\log|G_{j,\sigma}|)^2}{\delta}) = O(\frac{(\log\log n)^2}{\epsilon^2})$ and additive distortion $\delta \cdot \mathrm{diam}(G_{j,\sigma}) \leq \frac{\epsilon^2}{12}(\frac{4}{\epsilon} + 2) \leq \frac{\epsilon}{2}$ when $\epsilon \leq \frac{1}{2}$. The construction time is $O(|G_{j,\sigma}| \cdot \frac{\log^3|G_{j,\sigma}|}{\delta^2}) = O(|G_{j,\sigma}| \cdot \frac{\log^3 n}{\epsilon^4})$. For each vertex $u_{H_{j,\sigma}} \in H_{j,\sigma}$, we assign measure $\mu_{H_{j,\sigma}}(u_{H_{j,\sigma}}) = \mu(u)$ if $u_{H_{j,\sigma}} = f_{j,\sigma}(u)$ for some $u \in U^-_{j,\sigma}$, and $\mu_{H_{j,\sigma}}(u_{H_{j,\sigma}}) = 0$ otherwise. Let $K_{j,\sigma} = f(U^-_{j,\sigma})$ be a terminal set. Using Lemma 9, we find a $(1-\epsilon/2)$-independent set $I^H_{j,\sigma}$ for the terminal set $K_{j,\sigma}$ in $O\Big((\frac{\log n}{\epsilon})^{O(\epsilon^{-2}(\log\log n)^2)}|G_{j,\sigma}|\Big)$ time. Let $I_{j,\sigma} = f^{-1}(I^H_{j,\sigma})$ and $I_\sigma = \cup_{j\geq -1} I_{j,\sigma}$.

Claim 2. *I_σ is a $(1-\epsilon)$-independent set of G.*

Proof. Let u and v be any two vertices in I_σ such that $d_G(u, v) < 1$. Our goal is to show that $d_G(u, v) \geq 1 - \epsilon$. Suppose that there exists j such that $u \in G_{j,\sigma}$ and $v \notin G_{j,\sigma}$. By the construction of $I_{j,\sigma}$, $u \in U^-_{j,\sigma}$ and hence, the distance in G from u to any vertex not in $G_{j,\sigma}$ is at least 1. In particular, $d_G(u, v) \geq 1$, a contradiction. It follows that there exists j such that both u and v are in $G_{j,\sigma}$. By the construction of $I_{j,\sigma}$, both u and v are in $U^-_{j,\sigma}$. By the definition of $U^-_{j,\sigma}$, any vertex with distance at most 1 from u or v is contained in $U_{j,\sigma}$. Since $d_G(u, v) < 1$, any vertex in the shortest path from u to v in G is contained in $U_{j,\sigma}$. It follows that $d_G(u, v) = d_{G_{j,\sigma}}(u, v)$. By Theorem 3, and the fact that $I^H_{j,\sigma}$ is a $1 - \frac{\epsilon}{2}$ independent set, $d_{G_{j,\sigma}}(u, v) + \frac{\epsilon}{2} \geq d_{H_{j,\sigma}}(f_{j,\sigma}(u), f_{j,\sigma}(v)) \geq 1 - \epsilon/2$. It follows that $d_{G_{j,\sigma}}(u, v) \geq 1 - \epsilon$, which implies $d_G(u, v) \geq 1 - \epsilon$ as claimed. □

Next, we return the set $I = I_\sigma$ for the σ which maximizes $\mu(I_\sigma)$ over all $\sigma \in [0, \frac{2}{\epsilon} - 1]$. By Claim 2, I is a $(1-\epsilon)$-independent set of G. We now bound the cost of I_σ. Let $\tilde{I}$ be any 1-independent set in G, and $\sigma^* = \arg\max_\sigma \mu(\tilde{I} \cap V^-_\sigma)$. By Observation 3, we have:

$$\mu(\tilde{I} \cap V^-_{\sigma^*}) \geq \frac{\sum_{\sigma=0}^{\frac{2}{\epsilon}-1} \mu(\tilde{I} \cap V^-_\sigma)}{2/\epsilon} = \frac{2/\epsilon - 2}{2/\epsilon} \mu(\tilde{I}) = (1-\epsilon)\mu(\tilde{I}) \ .$$

For every σ and j, $f(\tilde{I} \cap U^-_{j,\sigma})$ is a 1-independent set in $H_{j,\sigma}$ w.r.t. the terminal set $K_{j,\sigma}$, and $\mu_{H_{j,\sigma}}(f(\tilde{I} \cap U^-_{j,\sigma})) = \mu(\tilde{I} \cap U^-_{j,\sigma})$. We conclude that

$$\mu(I) = \max_{\sigma} \mu(I_\sigma) \geq \mu(I_{\sigma^*}) = \sum_{j=-1}^{\infty} \mu(I_{j,\sigma^*}) = \sum_{j=-1}^{\infty} \mu(f^{-1}(I^H_{j,\sigma^*})) = \sum_{j=-1}^{\infty} \mu_{H_{j,\sigma^*}}(I^H_{j,\sigma^*})$$
$$\geq \sum_{j=-1}^{\infty} \mu_{H_{j,\sigma^*}}(f_{j,\sigma^*}(\tilde{I} \cap V^-_{\sigma^*})) = \mu(\tilde{I} \cap V^-_{\sigma}) \geq (1-\epsilon) \cdot \mu(\tilde{I}) \ .$$

The total running time to construct all the embeddings $f_{\sigma,j}$ into graphs $H_{\sigma,j}$, and to run the approximation algorithm from Lemma 9, denoted by $T_{IS}(n)$, is

$$T_{IS}(n) = \sum_{\sigma}\sum_{j} O(|G_{j,\sigma}| \cdot \frac{\log^3 |G_{j,\sigma}|}{\delta^2}) + O\left((\frac{\log n}{\epsilon})^{O(\epsilon^{-2}(\log\log n)^2)} |G_{j,\sigma}| \right) \tag{5}$$
$$= n \cdot \sum_{\sigma} \left(O(\frac{\log^3 n}{\epsilon^4}) + 2^{O(\epsilon^{-2} \cdot \log \frac{1}{\epsilon} \cdot (\log\log n)^3)} \right) \tag{6}$$
$$= n \cdot \left(O(\frac{\log^3 n}{\epsilon^5}) + (1/\epsilon) 2^{O(\epsilon^{-2} \cdot \log \frac{1}{\epsilon} \cdot (\log\log n)^3)} \right) \tag{7}$$
$$\tag{8}$$

For a given fixed $\kappa > 0$, by applying Young's inequality (Theorem 11) with $a = \epsilon^{-2} \cdot \log \frac{1}{\epsilon}$, $b = (\log\log n)^3$, $p = 1 + \kappa/2$ and $q = 1 + \frac{2}{\kappa}$, we have that:

$$\epsilon^{-2} \cdot \log \frac{1}{\epsilon} \cdot (\log\log n)^3 \leq \epsilon^{-(2+\kappa)} (\log \frac{1}{\epsilon})^{1+\kappa/2} \frac{2}{2+\kappa} + (\log\log n)^{3+6/\kappa} \frac{\kappa}{2+\kappa} \tag{9}$$
$$\tag{10}$$

Since κ is fixed, combining Equation (5) and Equation (9), we have that:

$$T_{IS}(n) = n \cdot \left(O(\frac{\log^3 n}{\epsilon^5}) + 2^{\tilde{O}(\epsilon^{-(2+\kappa)})} \cdot 2^{O((\log\log n)^{3+6/\kappa})} \right) = 2^{\tilde{O}(\epsilon^{-(2+\kappa)})} n^{1+o(1)} \ ,$$

as desired.

ρ-Dominating Set Problem. The algorithm for the ρ-dominating set problem is similar to the algorithm for ρ-independent set problem. For each $\sigma \in \{0, \dots, \frac{2}{\epsilon} - 1\}$, we define $\mathcal{I}_{-1,\sigma} = [0, \sigma]$ and $\mathcal{I}^+_{-1,\sigma} = [0, \sigma + 1]$, and for each $j \geq 0$:

$$\mathcal{I}_{j,\sigma} = [\frac{2}{\epsilon} \cdot j + \sigma \ , \ \frac{2}{\epsilon} \cdot (j+1) + \sigma] \qquad \mathcal{I}^+_{j,\sigma} = [\frac{2}{\epsilon} \cdot j + \sigma - 1 \ , \ \frac{2}{\epsilon} \cdot (j+1) + \sigma + 1]$$

Let $U_{j,\sigma} = \{v \in V : d_G(r,v) \in \mathcal{I}_{j,\sigma}\}$ and for $j \geq -1$, $U^+_{j,\sigma} = \{v \in V : d_G(r,v) \in \mathcal{I}^+_{j,\sigma}\}$. For every σ, set $G_{-1,\sigma} = G[U^+_{-1,\sigma}]$, and for $j \geq 0$, $G_{j,\sigma}$ is the graph obtained from G by removing every vertex of distance at least $\frac{2}{\epsilon} \cdot (j+1) + \sigma + 1$ from r, and contracting every vertex of distance strictly less than $\frac{2}{\epsilon} \cdot j + \sigma - 1$ to r; we set every edge incident to r the weight 1. Observe that $G_{j,\sigma}$ is a planar graph with diameter at most $\frac{4}{\epsilon}+6$. Furthermore, $V(G_{j,\sigma}) = U^+_{j,\sigma} \cup \{r\}$ when $j \geq 0$ and $V(G_{-1,\sigma}) = U^+_{-1,\sigma}$. As in the ρ-independent set problem, $G_{j,\sigma}$ can be constructed for all j in total $O(n)$ time.

Next, for every j and σ, we find a $(1+\epsilon)$-dominating set $S_{j,\sigma}$ in $G_{j,\sigma}$. Fix $\delta = \frac{\epsilon^2}{16}$. We construct an embedding $f_{j,\sigma}$ of $G_{j,\sigma}$ into a graph $H_{j,\sigma}$ with treewidth $O(\frac{(\log\log|G_{j,\sigma}|)^2}{\delta}) = O(\frac{(\log\log n)^2}{\epsilon^2})$ and additive distortion $\delta\cdot\mathrm{diam}(G_{j,\sigma}) \leq \frac{\epsilon^2}{16}(\frac{4}{\epsilon}+6) \leq \frac{\epsilon}{2}$ when $\epsilon \leq \frac{1}{2}$; by Theorem 3, $f_{j,\sigma}$ can be constructed in $O(|G_{j,\sigma}|\cdot\frac{\log^3|G_{j,\sigma}|}{\delta^2}) = O(|G_{j,\sigma}|\cdot\frac{\log^3 n}{\epsilon^4})$ time. For each vertex $u_{H_{j,\sigma}} \in H_{j,\sigma}$, we assign measure $\mu_{H_{j,\sigma}}(u_{H_{j,\sigma}}) = \mu(u)$ if $u_{H_{j,\sigma}} = f_{j,\sigma}(u)$ for some $u \in U^+_{j,\sigma}$, and $\mu_{H_{j,\sigma}}(u_{H_{j,\sigma}}) = \infty$ otherwise. Let $K_{j,\sigma} = f(U_{j,\sigma})$ be a terminal set. Using Lemma 9, we find a set $S^H_{j,\sigma}$ which $(1+\epsilon/2)$-dominates the terminal set $K_{j,\sigma}$ in $O(\frac{\log n}{\epsilon})^{O(\epsilon^{-2}(\log\log n)^2)}|G_{j,\sigma}|)$ time, and such that for every $(1+\frac{\epsilon}{2})$-dominating set $\hat{S}$, $\mu_{H_{j,\sigma}}(S^H_{j,\sigma}) \leq \mu_{H_{j,\sigma}}(\hat{S})$. Let $S_{j,\sigma} = f^{-1}(S^H_{j,\sigma})$. As $f(U_{j,\sigma})$ is a dominating set of finite measure, and all the vertices out of $f(U^+_{j,\sigma})$ have measure ∞, it holds that $S^H_{j,\sigma} \subseteq f(U^+_{j,\sigma})$. Observe that for every vertex $u \in U_{j,\sigma}$, $f(u) \in K_{j,\sigma}$, and hence there is a vertex $f(v) \in S^H_{j,\sigma}$ such that $d_{H_{j,\sigma}}(f(u), f(v)) \leq 1+\frac{\epsilon}{2}$. It follows that $v \in S_{j,\sigma}$, and $d_G(u,v) \leq d_{H_{j,\sigma}}(f(u),f(v)) \leq 1+\epsilon$. Set $S_\sigma = \cup_{j\geq -1} S_{j,\sigma}$, as $V = \cup_{j=-1}^{\infty} U_{j,\sigma}$, S_σ is a $1+\epsilon$ dominating set. We return the set $S = S_\sigma$ for the σ which minimizes $\mu(S_\sigma)$ for $\sigma \in [0, \frac{2}{\epsilon}-1]$.

Let $\tilde{S}$ be any 1-dominating in G, and σ^* the index that minimizes $\sum_{j=-1}^{\infty}\mu(\tilde{S}\cap U^+_{j,\sigma^*})$. Observe that each vertex $v \in V$ is contained in at most $\frac{2}{\epsilon}+2$ sets in the collection $\{U^+_{j,\sigma}\}_{\sigma\in\{0,\dots,\frac{2}{\epsilon}-1\}, j\geq -1}$. Thus

$$\sum_{j=-1}^{\infty}\mu(\tilde{S}\cap U^+_{j,\sigma^*}) \leq \frac{\epsilon}{2}\cdot\sum_{\sigma=0}^{\frac{2}{\epsilon}-1}\sum_{j=-1}^{\infty}\mu(\tilde{S}\cap U^+_{j,\sigma}) = \frac{\epsilon}{2}\cdot\sum_{v\in\tilde{S}}\mu(v)\sum_{\sigma=0}^{\frac{2}{\epsilon}-1}\sum_{j=-1}^{\infty}\cdot\mathbf{1}_{v\in U^+_{j,\sigma}} = \frac{\epsilon}{2}\cdot\sum_{v\in\tilde{S}}\mu(v)\cdot(\frac{2}{\epsilon}+2) = (1+\epsilon)\cdot\mu(\tilde{S})\ .$$

By Lemma 9, $\mu(S^H_{j,\sigma}) \leq \mu(\tilde{S}\cap U^+_{j,\sigma})$. We conclude,

$$\mu(S) = \min_{\sigma}\mu(S_\sigma) \leq \mu(S_{\sigma^*}) = \sum_{j=-1}^{\infty}\mu(S_{j,\sigma^*}) \leq \sum_{j=-1}^{\infty}\mu(\tilde{S}\cap U^+_{j,\sigma^*}) \leq (1+\epsilon)\cdot\mu(\tilde{S})\ .$$

The time analysis is exactly the same as the time analysis of the ρ-independent set problem.

4.3.2 Bounded-Capacity Vehicle Routing Problem

Using Lemma 8, we are able to reduce the problem to graphs with treewidth $O_\epsilon(\log\log n)^2$. Cohen-Addad *et al.* [CFKL20] designed an algorithm to solve the bounded-capacity VRP in $O(Q\cdot\epsilon^{-1}\log n)^{O(Q\cdot\frac{\mathsf{tw}}{\epsilon})}\cdot n^{O(1)}$ time for graphs with treewidth tw (Theorem 8 in [CFKL20], recall that Q is the capacity of the vehicle). Most of the time is spent on computing distances between $O(n\cdot\mathsf{tw}^2)$ pairs of vertices in G. An exact distance oracle is a data structure that, after some prepossessing stage, can answer distance queries (exactly) between every pair of vertices. By using a distance oracle instead of computing all distances directly, we have:

Lemma 10 (Theorem 8 [CFKL20], implicit)**.** *Let $(G(V,E,w),K,r,Q)$ be an instance of the VRP problem, where $G(V,E,w)$ has treewidth tw and n vertices. Suppose that we can construct an exact distance oracle for G with preprocessing time $P(n,\mathsf{tw})$ and query time $T(n,\mathsf{tw})$, then we can find $(1+\epsilon)$-approximate solution in time $(Q\epsilon^{-1}\log n)^{O(Q\cdot\mathsf{tw})}\cdot n + O(n\cdot\mathsf{tw}^2)\cdot T(n,\mathsf{tw}) + P(n,\mathsf{tw})$.*

Chaudhuri and Zaroliagis [CZ00] constructed a distance oracle with low preprocessing time and query time for directed graphs with small treewidth. Here, we only need an oracle for undirected graphs.

Lemma 11 (Theorem 4.2 (ii) in [CZ00]). *Let $G(V,E,w)$ be a directed graph with treewidth tw and n vertices. We can construct a distance oracle with prepprocessing time $P(n,\mathsf{tw}) = O(n\cdot\mathsf{tw}^3)$ and query time $T(n,\mathsf{tw}) = O(\mathsf{tw}^3\cdot\alpha(n))$ where $\alpha(n)$ is the inverse Ackermann function.*

By plugging Lemma 11 into Lemma 10, we get:

Lemma 12. *Given an instance of the VRP problem $(G(V,E,w),K,r,Q)$ where $G(V,E,w)$ has treewidth tw and n vertices, one can find a $(1+\epsilon)$-approximate solution in time $(Q\epsilon^{-1}\log n)^{O(Q\cdot\mathsf{tw})}\cdot n$.*

The final missing piece we need for our algorithm for the VRP is a result of Becker *et al.* [BKS19], who showed how to solve the VRP from an (r,ϵ)-rooted stochastic embedding and an efficient dynamic programming for VRP in small treewidth graphs.

Lemma 13 (Becker, Klein, and Schild [BKS19], implicit). *Let $\epsilon\in(0,1)$ be a parameter. Let $(G(V,E,w),K,r,Q)$ be an instance of the VRP problem with n vertices and m edges such that $G(V,E,w)$ admits an (r,ϵ)-rooted stochastic embedding into graphs with treewidth $\tau(n,\epsilon)$ that can be constructed in time $T(n,m,\epsilon)$. If we can find a $(1+\epsilon)$-approximate solution of the VRP in graphs with treewidth tw in time $D(n,Q,\mathsf{tw},\epsilon)$, then we can find a $(1+\epsilon)$-approximate solution for the VRP in $G(V,E,w)$ in expected time $D(n,Q,\tau(n,\epsilon/(9Q)),\epsilon/3)+T(n,m,\epsilon/(9Q))+O(m)$.*

We now show that there exists a PTAS for the bounded-capacity VRP in almost linear time. We restate Theorem 6 below for convenience.

Theorem 6 (PTAS for Bounded-Capacity VRP in Planar Graphs). *There is a randomized polynomial time approximation scheme for the bounded-capacity VRP in planar graphs that runs in time $O_\epsilon(1)\cdot n^{1+o(1)}$.*

Proof. By Lemma 8, we can construct an (r,ϵ)-rooted stochastic embedding into graphs with treewidth $\tau(n,\epsilon) = O_\epsilon(\log\log n)^2$ in $O_\epsilon(n\log^3 n)$ time. By Lemma 12, $D(n,Q,\mathsf{tw},\epsilon) = (Q\epsilon^{-1}\log n)^{O(Q\cdot\mathsf{tw})}n$. By Lemma 13, we can find a $(1+\epsilon)$-approximate solution of the VRP in time:

$$\begin{aligned}
(Q\epsilon^{-1}\log n)^{O_{\frac{\epsilon}{9Q}}(Q\cdot\log\log n)^2}\cdot n &= 2^{O_\epsilon(\log\log n)^3}\cdot n && (\text{since } Q = O(1))\\
&\le 2^{f(\epsilon)\cdot(\log\log n)^3}\cdot n && (\text{for some function } f(\epsilon))\\
&\le 2^{\frac{f(\epsilon)^2+(\log\log n)^6}{2}} n = O_\epsilon(1)\cdot 2^{\frac{(\log\log n)^6}{2}}\cdot n = O_\epsilon(n^{1+o(1)})\ . && (11)
\end{aligned}$$

□

5 Additional Notation and Preliminaries

5.1 Metric Embeddings

We will study a more permitting generalization of metric embedding introduced by Cohen-Addad *et al.* [CFKL20], which is called *one-to-many* embedding.

Definition 5 (One-to-many embedding)**.** *A* one-to-many embedding *is a function $f : X \to 2^Y$ from the points of a metric space (X, d_X) into non-empty subsets of points of a metric space (Y, d_Y), where the subsets $\{f(x)\}_{x \in X}$ are disjoint. $f^{-1}(x')$ denotes the unique point $x \in X$ such that $x' \in f(x)$. If no such point exists, $f^{-1}(x') = \emptyset$. A point $x' \in f(x)$ is called a* copy *of x, while $f(x)$ is called the* clan *of x. For a subset $A \subseteq X$ of vertices, denote $f(A) = \cup_{x \in A} f(x)$.*

We say that f is dominating *if for every pair of points $x, y \in X$, it holds that $d_X(x,y) \le \min_{x' \in f(x), y' \in f(y)} d_Y(x', y')$. We say that f has an additive distortion $+\epsilon D$ if f is dominating and $\forall x, y \in X$, $\max_{x' \in f(x), y' \in f(y)} d_Y(x', y') \le d_X(x,y) + \epsilon D$.*

A stochastic one-to-many embedding is a distribution $\mathcal{D}$ over dominating one-to-many embeddings. We say that a stochastic one-to-many embedding f has an expected additive distortion $+\epsilon D$, if $\forall x, y \in X$, $\mathbb{E}[\max_{x' \in f(x), y' \in f(y)} d_Y(x', y')] \le d_X(x,y) + \epsilon D$.

While the embedding in Theorem 1 is one-to-one, it is technically more convenient to construct a one-to-many embedding. This is because the embedding is constructed in multiple steps, where the "topological complexity" grows from one step to the next. To transition from one step to the next, we sometimes require the embeddings be clique-preserving (see Definition 8), a property for which one-to-many embeddings are required. Note that we can always make the final embedding f of G one-to-one by retaining exactly one copy in each set $f(v)$ for each $v \in V(G)$.

Next, we define clan embedding, which will be used in the approximation of metric Baker's problems. This notion was introduced by the authors[FL21].

Definition 6 (Clan Embedding)**.** *A clan embedding from metric space (X, d_X) into a metric space (Y, d_Y) is a pair (f, χ) where $f : X \to 2^Y$ is a dominating one-to-many embedding, and $\chi : X \to Y$ is a classic embedding. For every $x \in X$, we have that $\chi(x) \in f(x)$; here $f(x)$ called the clan of x, while $\chi(x)$ is referred to as the chief of the clan of x (or simply the chief of x).*

We say that a clan embedding f has an additive distortion $+\Delta$ if for every $x, y \in X$, $\min_{y' \in f(y)} d_Y(y', \chi(x)) \le d_X(x,y) + \Delta$.

A (Δ, δ)-clan embedding *is a distribution over clan embeddings with additive distortion $+\Delta$ such that the expected clan size has $\mathbb{E}[|f(x)|] \le (1 + \delta)$ for every $x \in X$.*

We will construct embeddings for minor-free graphs using a divide-and-concur approach. In order to combine different embeddings into a single one, it will be important that these embeddings are *clique-preserving.*

Definition 7 (Clique-copy)**.** *Consider a one-to-many embedding $f : G \to 2^H$, and a clique Q in G. A subset $Q' \subseteq f(Q)$ is called clique copy of Q if Q' is a clique in H, and for every vertex $v \in Q$, $Q' \cap f(v)$ is a singleton.*

Definition 8 (Clique-preserving embedding)**.** *A one-to-many embedding $f : G \to 2^H$ is called clique-preserving embedding if for every clique Q in G, $f(Q)$ contains a clique copy of Q. A clan embedding (f, χ) is clique-preserving if f is clique-preserving.*

Definition 9 (Ramsey Type Embedding)**.** *For a given parameter $\delta \in (0,1)$, a* (Δ, δ)-Ramsey type embedding *from metric space (X, d_X) into a metric space (Y, d_Y) with additive distortion $+\Delta$ is a distribution over dominating one-to-one embeddings $f : X \to Y$ such that there is a subset $M \subseteq X$ of vertices for which the following claims hold:*

1. *For every $u \in X$, $\Pr[u \in M] \ge 1 - \delta$.*

2. For every $u \in M$ and $v \in X$, $d_Y(f(u), f(v)) \le d_X(u,v) + \Delta$.

We note that the distortion in Definition 9 is worst-case.

5.2 Graph Minor Theory and Known Embeddings

We introduce the notation used in the graph minor theory Robertson and Seymour. A *vortex* is a graph W equipped with a path decomposition $\{X_1, X_2, \dots, X_t\}$ (see the discussion after Definition 1) and a sequence of t designated vertices $x_1, \dots, x_t$, called the *perimeter* of W, such that each $x_i \in X_i$ for all $1 \le i \le t$. The *width* of the vortex is the width of its path decomposition. We say that a vortex W is *glued* to a face F of a surface embedded graph G if $W \cap F$ is the perimeter of W whose vertices appear consecutively along the boundary of F.

If $G = G_\Sigma \cup W$ consists of planar graph G_Σ with a single vortex W of width h, we call G a *h-vortex planar graph.*

h-Multi-vortex-genus graph. a h-multi-vortex-genus graph is a graph $G = G_\Sigma \cup W_1 \cup \dots \cup W_h$, where G_Σ is (cellularly) embedded on a surface Σ of genus h, and each W_i is a vortex of width at most h glued to a face of G_Σ.

In constructing a low-treewidth embedding of K_r-minor-free graphs, Cohen-Addad *et al.* [CFKL20] provided a *deterministic* reduction from planar graphs with one vortex to surface-embedded graphs with many (but constant) vortices. The following lemma is a reinterpretation of Lemma 7 and Lemma 8 in the full version of their paper.

Lemma 14 (Multiple Vortices and Genus, Lemmas 7 and 8 [CFKL20], adapted)**.** *Suppose that every n-vertex h-vortex planar graph with diameter D can be deterministically embedded into a graph with treewidth $t(h, \epsilon, n)$ and additive distortion $+\epsilon D$ via a one-to-many clique-preserving embedding, then given an n-vertex h-multi-vortex-genus graph G of diameter D, G can be deterministically embedded into a graph with treewidth $O_h(t(O_h(1), O_h(\epsilon), O_h(n)))$ and additive distortion $+\epsilon D$ via a clique-preserving embedding.*

Nearly h-embeddability. A graph G is nearly h-embeddable if there is a set of at most h vertices A, called *apices*, such that $G \setminus A$ is an h-multi-vortex-genus graph $G_\Sigma \cup \{W_1, W_2, \dots, W_h\}$.

Cohen-Addad *et al.* [CFKL20] showed how to handle the apices by using randomness, effectively extending their Lemma 14 to obtain a reduction from planar graphs with one vortex to nearly h-embeddable graphs. The resulting embedding is stochastic. The stochasticity is necessary due to a lower bound for deterministic embedding of Cohen-Addad *et al.* [CFKL20] who showed that any (deterministic) embedding into graphs with treewidth $o(\sqrt{n})$ must incur an additive distortion at least $\frac{D}{20}$ (Theorem 3 in [CFKL20]). The following lemma is a reinterpretation of their Lemma 9.

Lemma 15 (Lemma 9 [CFKL20], adapted)**.** *If every h-vortex planar graphs of n vertices and diameter D can be (deterministically or stochastically) embedded into a graph with treewidth $t(h, \epsilon, n)$ and distortion $+\epsilon D$ via a one-to-many clique-preserving embedding, then nearly h-embeddable graphs with n vertices and diameter D can be* stochastically *embedded into a distribution over graphs with treewidth $O_h\left(t\left(O_h(1), O_h(\epsilon^2), n\right)\right) + O_h(\epsilon^{-2})$ and expected additive distortion $+\epsilon D$ via a clique-preserving embedding.*

h-Clique-sum. A graph G is an h-clique-sum of two graphs G_1 and G_2, denoted by $G = G_1 \oplus_h G_2$, if there are two cliques of size exactly h each such that G can be obtained by identifying vertices of the two cliques and removing some clique edges of the resulting identification. The shared clique between G_1 and G_2 is called the *adhesion* of the two graphs. The adhesion is also called the joint set of the two graphs in [CFKL20].

Note that clique-sum is not a well-defined operation since the clique-sum of two graphs is not unique due to the clique edge deletion step.

The celebrated theorem of Robertson and Seymour (Theorem 12, [RS03]) intuitively said that every K_r-minor-free graph can be decomposed into nearly embeddable graphs glued together following a tree-like structure by clique-sum operations.

Theorem 12 (Theorem 1.3 [RS03])**.** *There is a constant $h = O_r(1)$ such that any K_r-minor-free graph G can be decomposed into a tree $\mathbb{T}$ where each node of $\mathbb{T}$ corresponds to a nearly h-embeddable graph such that $G = \cup_{X_i X_j \in E(\mathbb{T})} X_i \oplus_h X_j$.*

The graphs corresponding to the nodes in the clique-sum decomposition above are referred to as *pieces*. A piece may not be a subgraph of G since some edge in a clique involving the clique-sum operation of the piece with another piece may not be present in G. However, we can modify the graph to guarantee that every piece is a subgraph of G by adding an edge (u, v) in the piece to G if the edge does not appear in G. The weight of (u, v) is set to be the shortest distance between u and v in G. This operation does not change the Robertson-Seymour decomposition of G nor its metric shortest path.

6 Stochastic Embedding of Minor-free Graphs and Applications

6.1 Stochastic Embedding of Minor-free Graphs (Proof of Theorem 1)

In this section, we construct a stochastic additive embedding for minor-free graphs as claimed in Theorem 1. We will construct a one-to-many embedding for G. Afterwards, the embedding can be converted into a one-to-one embedding by retaining a single copy of each vertex. We refer readers to Section 1.2 for a high-level overview of our construction. We start with an embedding for h-vortex planar graphs. In a nutshell, we improve the treewidth in the construction of Cohen-Addad *et al.* [CFKL20] in two places: (1) the embedding of h-vortex planar graphs and (2) gluing the embeddings of nearly embeddable graphs, a.k.a. pieces, via clique-sum operations. Our results are described in the following lemmas.

Lemma 16. *Given an n-vertex h-vortex planar graph $G_\Sigma \cup W$ of diameter D and a parameter $\epsilon < 1$, there is a polynomial time algorithm that constructs a one-to-many (deterministic) clique-preserving embedding $f : V(G) \to H$ into a graph H of treewidth at most $O(\frac{h(\log\log n)^2}{\epsilon})$ and additive distortion $+\epsilon D$.*

The proof of Lemma 16 is presented in Section 6.1.1.

Lemma 17. *If nearly h-embeddable graphs of n vertices and diameter D can be stochastically embedded into a graph with treewidth $t(h, \epsilon, n)$ by a clique-preserving embedding f of expected additive distortion $+\epsilon D$, then K_r-minor-free graphs with n vertices and diameter D can be stochastically embedded into a graph with treewidth $t(h(r), \epsilon, n) + h(r)$ and expected additive distortion $+\epsilon D$ for some function h depending on r only.*

The proof of Lemma 17 is presented in Section 6.1.2. We are now ready to prove Theorem 1, which we restate below for convenience.

Theorem 1 (Embedding Minor-free Graphs to Low Treewidth Graphs)**.** *Given an n-vertex K_r-minor-free graph $G(V,E,w)$ of diameter D, we can construct in polynomial time a stochastic additive embedding $f : V(G) \to H$ into a distribution over graphs H of treewidth at most $O_r(\epsilon^{-2}(\log\log n)^2)$ and expected additive distortion ϵD.*

Proof. By plugging Lemma 16 to Lemma 15 (here $\hat{t}(h,\epsilon,n) = O(\frac{h(\log\log n)^2}{\epsilon}))$, we obtain a stochastic, clique-preserving embedding of nearly h-embeddable graphs with additive distortion $+\epsilon D$ and treewidth:

$$O_h\left(\hat{t}\left(O_h(1), O_h(\epsilon^2), n\right)\right) + O_h(\epsilon^{-2}) = O_h\left(\frac{(\log\log n)^2}{\epsilon^2}\right) . \tag{12}$$

By applying Lemma 17 with $t(h,\epsilon,n) = O_h(\frac{(\log\log n)^2}{\epsilon^2})$ (due to Equation (12)), we get a stochastic embedding of K_r-minor-free graphs with additive distortion $+\epsilon D$ and treewidth:

$$t(h(r),\epsilon,n) + h(r) = O_r\left(\frac{(\log\log n)^2}{\epsilon^2}\right) + O_r(1) = O_r\left(\frac{(\log\log n)^2}{\epsilon^2}\right) ,$$

as desired. □

6.1.1 h-vortex planar graphs

In this section, we focus on constructing an additive embedding for h-vortex planar graphs as described in Lemma 16. We follow the same idea in the construction of an additive embedding for planar graphs in Section 4. Here we rely on (an implicit) construction of an $O(h)$-RSPD for h-vortex planar graphs by Cohen-Addad *et al.* [CFKL20] (see Section 5.1. in the full version).

Lemma 18 ([CFKL20])**.** *Given an edge-weighted h-vortex planar graph $G = G_\Sigma \cup W$ with n vertices, an $O(h)$-RSPD Φ of G can be constructed in polynomial time.*

Proof of Lemma 16. The deterministic embedding in Theorem 3 was constructed using RSPD and did not exploit any planar properties (other than the distance oracle Lemma 7 for obtaining an efficient implementation). Hence, we can use the exact same construction for h-vortex planar graph based on their $O(h)$-RSPD guaranteed in Lemma 18. The only difference is, since each node $\alpha \in \Phi$ can contain up to $O(h)$ shortest paths in its boundary $\mathcal{Q}_\alpha$, the width of the bags in the resulting tree decomposition might increase by up to an $O(h)$ factor. Thus other than the clique-preserving property, all the other properties guaranteed by the lemma hold.

In order to obtain the clique preserving property, we slightly modify the embedding by adding additional edges. We rely on the following structure of the RSPD , which is implicit in Cohen-Addad *et al.* [CFKL20]. We include the proof here for completeness.

Lemma 19 ([CFKL20])**.** *Let Q be any clique in an h-vortex planar graph $G_\Sigma \cup W$, and Φ be an* RSPD *of $G_\Sigma \cup W$. Then, there is a leaf $\alpha \in \Phi$ such that $Q \subseteq X_\alpha$.*

Proof. We prove by induction that at any level of Φ, there exists a node $\alpha \in \Phi$ such that $Q \subseteq X_\alpha$. The base case is when α is the root of Φ. The claim trivially holds since $X_\alpha = G_\Sigma \cup W$.

Let α is a node at level ℓ for some $\ell \geq 0$ such that $Q \subseteq X_\alpha$. Let β_1, β_2 be two children of α. Let $Y = V(X_{\beta_1}) \cap V(X_{\beta_2})$. Observe that there is no edge in G between $V(X_{\beta_1}) \setminus Y$ and $V(X_{\beta_2}) \setminus Y$ by property (2c) of RSPD (Definition 3). Thus, either $Q \cap (V(X_{\beta_1}) \setminus Y) = \varnothing$ or $Q \cap (V(X_{\beta_2}) \setminus Y) = \varnothing$ since it is a clique. In the former case, $Q \subseteq X_{\beta_2}$, and in the latter case, $Q \subseteq X_{\beta_1}$. □

Consider the $O(h)$-RSPD of $G_\Sigma \cup W$ given by Lemma 18. Φ has $O(n)$ nodes. Recall that we portalized all the shortest paths, and then applied the (4-step) construction in Section 4.1 to the $O(h)$-RSPD Φ to construct an embedding f, the host graph H, and a tree decomposition $\mathcal{T}$ of H. We remark that there are only $O_h(n)$ maximal cliques in $G_\Sigma \cup W$ [ELS10]. [8]

Let Q be a maximal clique in $G_\Sigma \cup W$, according to Lemma 19, there is a leaf $\alpha \in \Phi$ such that $Q \subseteq X_\alpha$. Recall that P_α is the set of portals of paths associated with α, and I_α is the set of internal vertices of α. Let $\widetilde{I_\alpha}$ and $\widetilde{P_\alpha}$ be the images of I_α and P_α in H, respectively. By Observation 2, $\widetilde{I_\alpha} \cup \widetilde{P_\alpha}$ induces a clique in H. Let B_α be the bag of $\mathcal{T}$ that contains $\widetilde{I_\alpha} \cup \widetilde{P_\alpha}$. We add a new copy $\widetilde{Q}$ of Q to H, add the edge set $\widetilde{Q} \times (\widetilde{I_\alpha} \cup \widetilde{P_\alpha})$ to $E(H)$, and update f accordingly. We then make a new bag $B_Q = \tilde{Q} \cup \widetilde{I_\alpha} \cup \widetilde{P_\alpha}$ and connect B_Q to B_α by adding an edge (B_Q, B_α) to $\mathcal{E}(\mathcal{T})$. Note that $|Q| = O(h)$ since $G_\Sigma \cup W$ is an h-vortex planar graph. Thus, the asymptotic bound on the treewidth of $\mathcal{T}$ remains unchanged. As usual, for every newly added edge (u', v') in H, set its weight to be $d_G(u,v)$ where $u = f^{-1}(u')$ and $v = f^{-1}(v')$. Denote by $\widetilde{H}$ the version of H after this last modification step. The clique preservation property holds by the construction. It is left to show that f has an additive distortion $O(\epsilon)D$; we can get additive distortion $+\epsilon D$ by scaling ϵ.

Claim 3. *For any $\tilde{u} \in \widetilde{Q}$, there is a vertex $\tilde{x} \in \widetilde{P_\alpha} \cup \widetilde{I_\alpha}$ such that $d_H(\tilde{u}, \tilde{x}) \leq \epsilon D$.*

Proof. Let $u = f^{-1}(\tilde{u})$. If $u \in I_\alpha$, then there is another copy $\tilde{u}'$ of u in $\widetilde{I_\alpha}$. By construction, there is an edge $(\tilde{u}, \tilde{u}')$ of weight 0. Thus, $\tilde{x}$ is $\tilde{u}'$ in this case. Otherwise, u must belong to some boundary path of X_α. Thus, there exists a portal $x \in P_\alpha$ such that $d_G(x,u) \leq \delta < \epsilon D$. By construction, the copy, say $\tilde{x}$, of x in $\widetilde{P_\alpha}$ has an edge to $\tilde{u}$ in H. The claim then follows from the fact that $w_H(\tilde{x}, \tilde{u}) = d_G(x,u) \leq \epsilon D$. □

We are now ready to bound the distortion of f. Let $\tilde{u}$ and $\tilde{v}$ be any two vertices in $\widetilde{H}$. Let $\tilde{x}$ and $\tilde{y}$ be closest vertices to $\tilde{u}$ and $\tilde{v}$, respectively, that belong to H (it is possible that $\tilde{x} = \tilde{u}$ or $\tilde{y} = \tilde{v}$). By Claim 3, $d_{\widetilde{H}}(\tilde{x}, \tilde{u}) \leq \epsilon D$ and $d_{\widetilde{H}}(\tilde{y}, \tilde{v}) \leq \epsilon D$. Let $x = f^{-1}(\tilde{x})$ and $y = f^{-1}(\tilde{y})$. Observe that $d_H(x,y) \leq d_G(x,y) + c\epsilon D$ for some constant c. By the triangle inequality, it holds that:

$$d_G(x,y) \leq d_G(x,u) + d_G(u,v) + d_G(v,y) \leq d_G(u,v) + 2\epsilon D \ .$$

Also by the triangle inequality, we have:

$$\begin{aligned} d_{\widetilde{H}}(\tilde{u}, \tilde{v}) &\leq d_{\widetilde{H}}(\tilde{u}, \tilde{x}) + d_H(\tilde{x}, \tilde{y}) + d_{\widetilde{H}}(\tilde{y}, \tilde{v}) \ \leq \ d_H(\tilde{x}, \tilde{y}) + 2\epsilon D \\ &\leq d_G(x,y) + (c+2)\epsilon D \ \leq \ d_G(u,v) + (c+4)\epsilon D, \end{aligned}$$

as desired. □

[8] The degeneracy of a graph is the minimal d such that there is an order $v_1 v_2 \dots v_n$ over the vertices where for every i, v_i has at most d neighbors among $v_{i+1} \dots v_n$. Eppstein, Löffler and Strash [ELS10] showed that every graph with degeneracy d has at most $(n-d)3^{\frac{d}{3}}$ maximal cliques. As every n-vertex K_r-minor-free graph has average degree at most $O(r \cdot \sqrt{\log r})$ [Kos82, Tho84], it has degeneracy at most $O(r \cdot \sqrt{\log r})$, and hence $O_r(n)$ maximal cliques.

6.1.2 From nearly embeddable graphs to general K_r-minor-free graphs

In this section, we prove a reduction from nearly embeddable graphs to general K_r-minor-free graphs as described in Lemma 17. Our reduction removes a loss of an additive $O(\log n)$ factor in the treewidth in the reduction of Cohen-Addad *et al.* [CFKL20]. To give context to our improvement, we briefly review their construction below.

By the Robertson-Seymour theorem (Theorem 12), a K_r-minor-free graphs G can be decomposed into a tree $\mathbb{T}$ of nearly $h(r)$-embeddable graphs, called *pieces*, that are glued together by taking clique-sums along the edges of $\mathbb{T}$. We then embed each piece into a graph with treewidth $t(h(r), \epsilon, n)$ by a clique-preserving embedding. The clique-preserving property allows us to take clique-sum of the host graphs of the pieces to obtain an embedding of G. The remaining problem is that the distortion of the embedding could be arbitrarily large due to long paths in $\mathbb{T}$. Cohen-Addad *et al.* [CFKL20] resolved this issue by "shortcutting" long paths via centroids[9]. Specifically, they apply the following steps: (a) find a centroid piece G_c of $\mathbb{T}$, (b) recursively construct the embedding for each subgraph K obtained by removing the centroid piece from $\mathbb{T}$, (c) glue the embedding of G_c and K via clique-sum, and (d) add (the image of) the adhesion J between K and G_c to every bag in the tree decomposition of the host graph of K. Step (d) helps "shortcut" the shortest path, say Q_{uv}, between any two vertices u and v between two different subgraphs K_i and K_j of $\mathbb{T} \setminus G_c$ in the following way: there is a pair of vertices $x \in Q_{uv} \cap J_i$ and $y \in Q_{uv} \cap J_j$ such that the distances between u and x, and between v and y are preserved *exactly*. Thus, the final embedding only incurs an additive distortion $+\epsilon D$ for the distance between x and y (in the embedding of G_c). However, step (d) increases the treewidth of the embedding additively by $O(h(r)\log n)$, since each adhesion has size $O(h(r))$ and the recursion has depth $O(\log n)$—by using the centroids for shortcutting—where each recursive step adds one adhesion to the embedding a piece in $\mathbb{T}$.

Our idea to remove the $\log n$ factor is as follows. We first root the tree $\mathbb{T}$. Then for each piece G_i, we simply add the adhesion J_0 between G_i and its parent bag G_0 to every bag in the tree decomposition of the host graph H_i in the embedding of G_i; this increases the treewidth of G_0 by only $+O(h(r))$, and the overall treewidth of the final embedding is also increased only by $+O(h(r))$. For the stretch, we show by induction that for every vertex $u \in G_a$ and $v \in G_b$ for any two pieces G_a, G_b such that G_i is their lowest common ancestor, there is a vertex $x \in Q_{uv} \cap G_i$ and $y \in Q_{uv} \cap G_i$, the distances between u and x, and between v and y are preserved *exactly*. Thus, the distortion for $d_G(u,v)$ is just $+\epsilon D$. Since we apply this construction to *every* piece of $\mathbb{T}$, $+\epsilon D$ is also the distortion of the final embedding.

We are now ready to prove Lemma 17, which we restate below.

Lemma 17. *If nearly h-embeddable graphs of n vertices and diameter D can be stochastically embedded into a graph with treewidth $t(h, \epsilon, n)$ by a clique-preserving embedding f of expected additive distortion $+\epsilon D$, then K_r-minor-free graphs with n vertices and diameter D can be stochastically embedded into a graph with treewidth $t(h(r), \epsilon, n) + h(r)$ and expected additive distortion $+\epsilon D$ for some function h depending on r only.*

Proof. We apply Theorem 12 to obtain a decomposition of G into a tree $\mathbb{T}$ of pieces $\{G_1, \ldots, G_\ell\}$. Since every adhesion involved in a clique-sum operation is a clique, G_i has a diameter at most D for every $i \in [\ell]$. Root $\mathbb{T}$ at a piece G_{rt} for some $rt \in [\ell]$. For each G_i, let f_i be the clique-preserving embedding G_i into H_i with treewidth $t(h(r), \epsilon, |V(G_i)|) \le t(h(r), \epsilon, n)$ and expected additive distortion $+\epsilon D$. Let $\mathcal{T}_i$ be the tree decomposition of H_i with width at most $t(h(r), \epsilon, n)$.

[9] A centroid of a tree T is a vertex v such that every subtree of $T \setminus v$ has at most $|V(T)|/2$ vertices

For each pair of pieces G_i and G_j where G_j is the parent piece of G_i in $\mathbb{T}$. Let J_{ij} be the adhesion involving the clique sum of G_i and G_j. Let Q^i_{ij} and Q^j_{ij} be the clique copies of J_{ij} in H_i and H_j. We apply the following two steps:

- **(Step 1).** We add to H_i the edge set $Q^i_{ij} \times V(H_i)$ and set the weight of each edge to be the distance between the pre-images of its endpoints in G; recall that $f_i(V(G_i)) = V(H_i)$. We then add Q^i_{ij} to every bag of the tree decomposition $\mathcal{T}_i$ so that it remains a valid tree decomposition of H_i after adding edges.
- **(Step 2).** We identify Q^i_{ij} and Q^j_{ij} into a single clique copy Q'_{ij}. This effectively connects H_i and H_j. We then replace the vertices of Q^i_{ij} and Q^j_{ij} in $\mathcal{T}_i$ and $\mathcal{T}_j$, respectively, by the vertices of Q'_{ij}. We then take a pair of arbitrary two bags B_i and B_j of $\mathcal{T}_i$ and $\mathcal{T}_j$, respectively, that contain Q'_{ij}, and add an edge (B_i, X_j) to connect $\mathcal{T}_i$ and $\mathcal{T}_j$ into a single tree decomposition. X_i and X_j exist since f_i and f_j are clique-preserving.

Let H and $\mathcal{T}$ be the host graph and the tree decomposition obtained by applying the above two steps to every pair of pieces G_i and G_j where G_j is the parent piece of G_i in $\mathbb{T}$. For each vertex $\tilde{v} \in V(H)$, we add $\tilde{v}$ to $f(v)$, which is initially empty, if $\tilde{v}$ is the image of v in the embedding of some piece in the decomposition of G. This completes our construction of the embedding of G.

Observe that $\mathcal{T}$ is a valid tree decomposition of H since every time we add an edge $(\tilde{u}, \tilde{v})$ to $E(H)$ that is not present in any $\{H_i\}_{i=1}^{\ell}$ (in Step 1), we guarantee that $\tilde{u}$ and $\tilde{v}$ are in the same bag. Additionally, the construction in Step 1 only increases the treewidth of H_i by at most $h(r)$, which is the upper bound on the size of J_{ij} (this happens only once per piece). The width of the resulting tree in Step 2 is bounded by the maximum width of $\mathcal{T}_i$ and $\mathcal{T}_j$. Thus, the width of $\mathcal{T}$ is $t(h(r), \epsilon, n) + h(r)$.

Observe that f is a dominating embedding since $\{f_i\}_i$ are dominating. Thus, it is left to show that the expected distortion of the embedding is $+\epsilon D$. To that end, we show the following claim.

Claim 4. *Let G_a and G_b be two pieces of $\mathbb{T}$ such that G_a is a proper ancestor of G_b. Let $J^{\downarrow}_{ab}$ be the adhesion in the clique-sum between G_a and its child on the path from G_a to G_b. For any two vertices $u \in J^{\downarrow}_{ab}$ and $v \in G_b$, there exists a copy $\tilde{u} \in f_a(u)$ such that for every $\tilde{v} \in f_b(v)$:*

$$d_H(\tilde{u}, \tilde{v}) = d_G(u, v) \ .$$

Proof. We prove by induction. The base case is when G_a is a parent of G_b. Then $J^{\downarrow}_{ab} = J_{ab}$. Let Q^a_{ab} and Q^b_{ab} be two clique copies of J^{ab} in H_a and H_b, respectively, that are identified into a single clique copy Q'_{ab} in the construction of Step 2. Let $\tilde{u}$ be the copy of u in Q'_{ab}. By the construction in Step 1, there is an edge from $\tilde{u}$ to every vertex in $f(V(G_b))$. Thus for every $\tilde{v} \in f_b(v)$, $d_H(\tilde{u}, \tilde{v}) = d_G(u, v)$.

For the inductive case, let G_c be the parent piece of G_b. Let S_{uv} be a shortest path between u and v in G. Since $G_c \in \mathbb{T}[G_a, G_b]$, there must be a vertex $x \in J_{cb} \cap S_{uv}$. By the induction hypothesis, there exists a copy $\tilde{x} \in f_c(x)$ such that for every $\tilde{v} \in f_b(v)$, $d_H(\tilde{x}, \tilde{v}) = d_G(x, v)$. Also by induction, there exists a copy $\tilde{u} \in f_a(u)$ such that $d_H(\tilde{u}, \tilde{x}) = d_G(u, x)$. By the triangle inequality, for every $\tilde{v} \in f_b(v)$:

$$d_H(\tilde{u}, \tilde{v}) \leq d_H(\tilde{u}, \tilde{x}) + d_H(\tilde{x}, \tilde{v}) = d_G(u, x) + d_G(x, v) = d_G(u, v) \ .$$

As f is dominating, we conclude that $d_H(\tilde{u}, \tilde{v}) = d_G(u, v)$. $\square$

We now show that for every u, v,

$$\mathbb{E}\left[\max_{\tilde{u}\in f(u),\tilde{v}\in f(v)} d_H(\tilde{u},\tilde{v})\right] \le d_G(u,v) + \epsilon D \ . \tag{13}$$

Let G_a and G_b be the two pieces closest to the root of $\mathbb{T}$, containing u and v respectively. If $G_a = G_b$, then $d_{G_a}(u,v) = d_G(u,v)$, and by the assumption that f_a has expected additive distortion $+\epsilon D$, we have that $\mathbb{E}[\max_{\tilde{u}\in f_a(u),\tilde{v}\in f_a(v)} d_H(\tilde{u},\tilde{v})] \le d_H(u,v) + \epsilon D$. By Claim 4, for any copy $\tilde{u} \in f(u)$ (resp. $\tilde{v} \in f(v)$) coming from a decedent piece of G_a (resp. G_b), there is a copy $\hat{u} \in f_a(u)$ (resp. $\hat{v} \in f_a(v)$) such that $d_H(\hat{u},\tilde{u}) = d_G(u,u) = 0$ (resp. $d_H(\hat{v},\tilde{v}) = 0$). It follows that

$$\begin{aligned}\mathbb{E}\left[\max_{\tilde{u}\in f(u),\tilde{v}\in f(v)} d_H(\tilde{u},\tilde{v})\right] &\le \mathbb{E}\left[\max_{\tilde{u}\in f(u),\tilde{v}\in f(v)} d_H(\tilde{u},\hat{u}) + d_H(\hat{u},\hat{v}) + d_H(\hat{v},\tilde{v})\right] \\ &\le \mathbb{E}\left[\max_{\hat{u}\in f_a(u),\hat{v}\in f_a(v)} d_H(\hat{u},\hat{v})\right] \quad \le \quad d_G(u,v) + \epsilon D \ . \end{aligned} \tag{14}$$

Otherwise ($G_a \ne G_b$), let G_c be the least common ancestor of G_a and G_b in $\mathbb{T}$. Let S_{uv} be the shortest path between u and v in G. Then, there must exist two vertices $x \in J^{\downarrow}_{ca} \cap S_{uv}$ and $y \in J^{\downarrow}_{cb} \cap S_{uv}$. By Claim 4, there exist $\hat{x} \in f_c(x)$ and $\hat{y} \in f_c(y)$ such that for every two copies $\tilde{u} \in f_a(u)$ and $\tilde{v} \in f_b(v)$,

$$d_H(\hat{x},\tilde{u}) = d_G(x,u) \qquad \& \qquad d_H(\hat{y},\tilde{v}) = d_G(y,v).$$

By the triangle inequality, linearity of expectation, the fact that f_c has expected distortion $+\epsilon D$, and the fact that x and y lie on a shortest u-v path, it holds that:

$$\begin{aligned}\mathbb{E}\left[\max_{\tilde{u}\in f_a(u),\tilde{v}\in f_b(v)} d_H(\tilde{u},\tilde{v})\right] &\le \mathbb{E}\left[\max_{\tilde{u}\in f_a(u),\tilde{v}\in f_b(v)} d_H(\tilde{u},\hat{x}) + d_H(\hat{x},\hat{y}) + d_H(\hat{y},\tilde{v})\right] \\ &\le d_G(u,x) + \mathbb{E}\left[\max_{\tilde{x}\in f_c(x),\tilde{y}\in f_c(y)} d_H(\tilde{x},\tilde{y})\right] + d_G(y,v) \\ &= d_G(u,x) + d_G(x,y) + d_G(y,v) + \epsilon D \quad = \quad d_G(u,v) + \epsilon D \ . \end{aligned}$$

For any copy $\tilde{u} \in f(u)$ (resp. $\tilde{v} \in f(v)$) coming from a decedent piece of G_a (resp. G_b), by Equation (14), we conclude that Equation (13) holds, as desired. □

6.2 An Efficient PTAS for the Bounded-capacity VRP (Proof of Theorem 2)

In this section, we show that the embedding in Theorem 1 can be used to obtain an efficient PTAS for the bounded-capacity VRP in K_r-minor-free graphs as claimed in Theorem 2. Our result significantly improves the previous result by Cohen-Addad *et al.* [CFKL20] who showed an approximation scheme with *quasi-polynomial time* for this problem.

The algorithm for minor-free graphs closely resembles the algorithm for planar graphs in Section 4.3.2. The following lemma is similar to Lemma 8.

Lemma 20. *Let $G(V,E,w)$ be an n-vertex K_r-minor-free graph and r be a distinguished vertex in V. For any given parameter $\epsilon \in (0, \frac{1}{4})$, we can construct in $O_{\epsilon,r}(1)\cdot n^{O(1)}$ time an (r,ϵ)-rooted stochastic embedding of G into graphs H of treewidth $O_{r,\epsilon}((\log\log n)^2)$.*

Proof. Our construction follows exactly the same steps as the construction of Lemma 8; the only difference is that we use Theorem 1 instead of Theorem 3 (and the standard shortest path algorithm instead of [HKRS97]). The running time is clearly polynomial.

For the distortion analysis, consider a pair of vertices $u, v \in V$, and let i be the index such that $u \in B_i$, and let f_i be the embedding of G_i to the low treewidth graph H_i. Recall that the vertex u is successful with probability $1 - 2\epsilon$. If u is successful and the shortest path P between u and v is not fully contained in B_i, then by the case analysis in Lemma 8, $d_H(v,u) \le d_G(v,u) + 2 \cdot \epsilon \cdot d_G(r,u)$ (as that argument used only edges incident to r, which are present in our embedding as well). Otherwise, $P \subseteq B_i$, and hence $d_{G_i}(u,v) = d_G(u,v)$. Following the calculation in Equation (4), we have

$$\mathbb{E}\left[d_H(f(u), f(v)) \mid d_{G_i}(u,v) = d_G(u,v)\right] \le d_G(u,v) + \delta \cdot 2U_i \le d_G(u,v) + 2\epsilon \cdot d_G(r,u) \ .$$

Thus, conditioning on u being successful, we have

$$\mathbb{E}[d_H(f(u), f(v)) \mid u \text{ is successful}] \le d_G(v,u) + O(\epsilon) \cdot (d_G(r,v) + d_G(r,u)) \ ,$$

and it follows that:

$$\begin{aligned}\mathbb{E}[d_H(f(u), f(v))] \le \Pr\left[v \text{ is successful}\right] \cdot \mathbb{E}[d_H(f(u), f(v)) \mid u \text{ is successful}] \\ + \Pr\left[v \text{ is unsuccessful}\right] \cdot (d_G(v,r) + d_G(r,u)) \\ \le d_G(v,u) + O(\epsilon) \cdot (d_G(r,v) + d_G(r,u)) \ .\end{aligned}$$

□

We are now ready to prove Theorem 2, which we restate below for convenience.

Theorem 2 (PTAS for Bounded-Capacity VRP in Minor-free Graphs)**.** *There is a randomized polynomial time approximation scheme for the bounded-capacity VRP in K_r-minor-free graphs that runs in $O_{\epsilon,r}(1) \cdot n^{O(1)}$ time.*

Proof. Following exactly the same steps in the proof of Theorem 6 in Section 4.3.2, we plug Lemma 20 and Lemma 12 into Lemma 13 to obtain a $(1 + \epsilon)$-approximate solution. Following Equation (11), the running time hence is:

$$(Q\epsilon^{-1} \log n)^{O_{\frac{\epsilon}{9Q}}(Q \cdot \log\log n)^2} \cdot n + O_{\epsilon,r}(1) \cdot n^{O(1)} = O_{\epsilon,r}(1) \cdot n^{O(1)} \ .$$

□

7 Ramsey Type and Clan Embeddings of Minor-free Graphs and their Applications

7.1 Ramsey Type and Clan Embeddings (Proofs of Theorems 7 and 8)

In this section, we prove Theorem 7 and Theorem 8. A key idea in the construction of the authors [FL21] is that one can construct a clan/Ramsey type embedding from a deterministic one-to-one embedding of h-multi-vortex-genus graphs with an $O(\log n)$ loss in the treewidth. Lemma 21 below is a reinterpretation of Lemma 15 and Lemma 16 for Ramsey type embedding and Lemma 18 and Lemma 19 for clan embedding in [FL21].

Lemma 21 (Lemmas 15,16, 18 and 19 [FL21], adapted)**.** *Suppose that every n-vertex h-multi-vortex-genus graph of diameter D can be deterministically embedded into a graph with treewidth $t(h,\epsilon,n)$ and additive distortion $+\epsilon D$ via a clique-preserving embedding. Then for every parameter $\delta \in (0,1)$, and every n-vertex K_r-free graph $G = (V,E,w)$ of diameter D, there exist an $(\epsilon D,\delta)$-Ramsey type embedding and an $(\epsilon D,\delta)$-clan embedding, both with treewidth $t(O_r(1), O(\frac{\delta\epsilon}{\log(n)}), n) + O_r(\log n)$.*

To obtain their treewidth bound $O_r(\frac{\log^2}{\delta\epsilon})$, the authors[FL21] used the embedding of h-multi-vortex-genus graphs with treewidth $O_r(\frac{\log(n)}{\epsilon})$ by Cohen-Addad *et al.* [CFKL20].

We show below that Lemma 21 and Lemma 14 provide a reduction from *h-vortex planar graphs* to K_r-minor-free graphs.

Lemma 22. *Suppose that every h-vortex planar graph of n vertices and diameter D can be deterministically embedded into a graph with treewidth $t(h,\epsilon,n)$ and additive distortion $+\epsilon D$ via a clique-preserving embedding. Then for every parameter $\delta \in (0,1)$ and every K_r-free graph $G = (V,E,w)$ of n vertices and diameter D, there exist an $(\epsilon D,\delta)$-Ramsey type embedding and an $(\epsilon D,\delta)$-clan embedding, both with treewidth $O_r(t(O_1(r), O_r(\frac{\delta\epsilon}{\log n}), n) + \log(n))$.*

Proof. By Lemma 14, any h-multi-vortex-genus graph G of n vertices and diameter D can be deterministically embedded into a graph with treewidth $\hat{t}(h,\epsilon,n) = O_h(t(O_h(1), O_h(\epsilon), n))$ via a clique-preserving embedding. This embedding and Lemma 21 give a $(\epsilon D,\delta)$-Ramsey type embedding and a $(\epsilon D,\delta)$-clan embedding with treewidth:

$$\begin{aligned}\hat{t}(O_r(1), O(\frac{\delta\epsilon}{\log(n)}), n) + O_r(\log n) &= O_r(t(O_r(1), O_r(\frac{\delta\epsilon}{\log n}), n) + O_r(\log(n))) \\ &= O_r(t(O_1(r), O_r(\frac{\delta\epsilon}{\log n}), n) + \log(n)) \ ,\end{aligned}$$

as claimed. □

We are now ready to prove Theorem 7 and Theorem 8.

Proof of Theorem 7 and Theorem 8. By Lemma 16, any h-vortex planar graph of n vertices and diameter D has a clique-preserving embedding with additive distortion $+\epsilon D$ and treewidth:

$$t(h,\epsilon,n) = O(\frac{h(\log\log n)^2}{\epsilon}).$$

It follows from Lemma 22 that K_r-free graphs of n vertices and diameter D has an $(\epsilon D,\delta)$-Ramsey type embedding and an $(\epsilon D,\delta)$-clan embedding, both with treewidth:

$$O_r(t(O_1(r), O_r(\frac{\delta\epsilon}{\log n}), n) + \log(n)) = O_r\left(O_r\left(\frac{(\log\log n)^2}{(\delta\epsilon)/\log(n)}\right) + \log(n)\right) = O_r\left(\frac{\log n(\log\log n)^2}{\epsilon\delta}\right) \ ,$$

as desired. □

7.2 Applications to Metric Baker's Problems (Proof of Theorem 9)

Given a clan embedding, one could obtain a bicriteria approximation scheme for the ρ-dominating set problem by dynamic programming on the host graph. Similarly, given a Ramsey type embedding, one could obtain a bicriteria approximation scheme for the ρ-independent set problem, also by dynamic programming. Lemma 9 provides efficient dynamic programs for both problems in graphs with low treewidth. The following lemmas specify the running time of the resulting algorithms.

Lemma 23 (Theorems 7 and 8 [FL21], adapted)**.** *Suppose that for every parameters $\epsilon,\delta\in(0,1)$, and every K_r-free graph $G=(V,E,w)$ of diameter D and n vertices, there are $(\epsilon D,\delta)$-clan and $(\epsilon D,\delta)$-Ramsey type embeddings, both with treewidth $\mathsf{tw}(\epsilon,\delta)$. Then given an n-vertex K_r-minor-free graph $G(V,E,w)$, two parameters $\epsilon\in(0,1)$ and $\rho>0$, and a vertex weight function $\mu:V\to\mathbb{R}^+$, one can find in $2^{O(\mathsf{tw}(\Theta(\epsilon),\Theta(\epsilon))\cdot\log\frac{\mathsf{tw}(\Theta(\epsilon),\Theta(\epsilon))}{\epsilon})}\cdot\mathrm{poly}(n)$ time:*
(1) a $(1-\epsilon)\rho$-independent set I such that for every ρ-independent set $\tilde{I}$, $\mu(I)\ge(1-\epsilon)\mu(\tilde{I})$, and
(2) a $(1+\epsilon)\rho$-dominating set S such that for every ρ-dominating set $\tilde{S}$, $\mu(S)\le(1+\epsilon)\mu(\tilde{S})$.

We are now ready to prove Theorem 9.

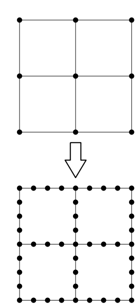
Proof of Theorem 9. By Theorem 7 and Theorem 8, K_r-minor-free graphs of diameter D and n vertices have $(\epsilon D,\delta)$-clan and $(\epsilon D,\delta)$-Ramsey type embeddings, both with treewidth $\mathsf{tw}(\epsilon,\delta)=O_r(\frac{\log n(\log\log n)^2}{\epsilon\delta})$. Thus, by Lemma 23, we obtain a bicriteria approximation scheme for ρ-independent set and ρ-dominating set problems in time:

$$\begin{aligned}2^{O(\mathsf{tw}(\Theta(\epsilon),\Theta(\epsilon))\cdot\log\frac{\mathsf{tw}(\Theta(\epsilon),\Theta(\epsilon))}{\epsilon})}\cdot\mathrm{poly}(n) &= 2^{O(O_r(\frac{\log n(\log\log n)^2}{\epsilon^2}\log(\frac{\log n(\log\log n)^2}{\epsilon^3})}\\ &= n^{\tilde{O}_r(\epsilon^{-2}(\log\log(n))^3)}~,\end{aligned}$$

as claimed. □

8 Treewidth Lower Bounds

In this section, we prove a lower bound on the treewidth of the embedding planar graphs with additive distortion. For a given graph $J=(V_J,E_J)$, we denoted by J_k a k-subdivision of J. That is, J_k is obtained from J by replacing each edge with a path of length k. See the illustration on the right for a 4-subdivision. Our lower bound uses the following lemma by Carrol and Goel [CG04].



Lemma 24 (Lemma 1 [CG04])**.** *Let H be a (possibly weighted) graph that excludes J as a minor. Every dominating embedding from J_k to H has a multiplicative distortion at least $\frac{k-3}{6}$.*

Theorem 4 (Embedding Planar Graphs Lower Bound)**.** *For any $\epsilon\in(0,\frac{1}{2})$ and any $n=\Omega(1/\epsilon^2)$, there exists an unweighted n-vertex planar graph $G(V,E)$ of diameter $D\le(1/\epsilon+2)$ such that for any deterministic dominating embedding $f:V(G)\to H$ into a graph H with additive distortion $+\epsilon D$, the treewidth of H is $\Omega(1/\epsilon)$.*

Proof. W.l.o.g, we assume that ϵ is sufficiently smaller than 1 since otherwise, the theorem trivially holds. Set $n'=\frac{1}{42\epsilon}+1$. Let Q be an $n'\times n'$-grid. The diameter of Q is $2\cdot\frac{1}{42\epsilon}=\frac{1}{21\epsilon}$. Let Q_k be a k-subdivision of Q for $k=21$. Observe that the diameter of Q_k is $21\cdot\frac{1}{21\epsilon}=\frac{1}{\epsilon}$. Furthermore, $|V(Q_k)|=$

$O(1/\epsilon^2)$. Finally, we add $n - |V(Q_k)|$ dummy vertices and connect all of them to some vertex of Q_k in a star-like way. Let G be the resulting graph; G is planar by construction. Furthermore, the diameter of G is $D = 1/\epsilon + 1$.

Suppose that there is a deterministic, dominating embedding f of G with an additive distortion at most ϵD into a treewidth $o(1/\epsilon)$ graph H. Observe that $\epsilon D = \epsilon(\frac{1}{\epsilon} + 2) = 1 + 2\epsilon \leq 2$ when $\epsilon \leq 1/2$. It follows that f has multiplicative distortion at most 2. As H has treewidth $o(1/\epsilon)$ (see e.g. [RST94]), it excludes Q as a minor. Thus by Lemma 24, any embedding of Q_k into H, including f_{Q_k}, must have a multiplicative distortion at least $\frac{k-3}{6} = 3$, a contradiction. Thus, the treewidth of H must be $\Omega(1/\epsilon)$ as claimed. □

Acknowledgments The authors are grateful to Philip Klein and Vincent Cohen-Addad for helpful conversations. We thank anonymous reviewers for their helpful comments. In particular, one reviewer pointed out the notion of least common ancestor closure and reference [FLSZ18] which help us simplify the proof of Theorem 10 considerably.

References

[ACE+20] I. Abraham, S. Chechik, M. Elkin, A. Filtser, and O. Neiman. Ramsey spanning trees and their applications. *ACM Trans. Algorithms*, 16(2):19:1–19:21, 2020. preliminary version published in SODA 2018. doi:10.1145/3371039.

[ACL09] A. Adamaszek, A. Czumaj, and A. Lingas. PTAS for k-tour cover problem on the plane for moderately large values of k. In *International Symposium on Algorithms and Computation*, ISAAC '09, pages 994–1003, 2009. doi:10.1007/978-3-642-10631-6_100.

[AFGN18] I. Abraham, A. Filtser, A. Gupta, and O. Neiman. Metric embedding via shortest path decompositions. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018, Los Angeles, CA, USA, June 25-29, 2018*, pages 952–963, 2018. full version: https://arxiv.org/abs/1708.04073. doi:10.1145/3188745.3188808.

[AGG+19] I. Abraham, C. Gavoille, A. Gupta, O. Neiman, and K. 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.

[AGK+98] S. Arora, M. Grigni, D. R. Karger, P. N. Klein, and A. Woloszyn. A polynomial-time approximation scheme for weighted planar graph TSP. In *Proceedings of the Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, 25-27 January 1998, San Francisco, California, USA*, pages 33–41, 1998. see here.

[AHLT05] S. Alstrup, J. Holm, K. D. Lichtenberg, and M. Thorup. Maintaining information in fully dynamic trees with top trees. *ACM Transactions on Algorithms*, 1(2):243–264, 2005. doi:10.1145/1103963.1103966.

[AKTT97] T. Asano, N. Katoh, H. Tamaki, and T. Tokuyama. Covering points in the plane by k-tours: Towards a polynomial time approximation scheme for general k. In *Proceedings of the 29thh Annual ACM Symposium on Theory of Computing*, STOC '97, page 275–283, 1997. `doi:10.1145/258533.258602`.

[AMS99] N. Alon, Y. Matias, and M. Szegedy. The space complexity of approximating the frequency moments. *J. Comput. Syst. Sci.*, 58(1):137–147, 1999. preliminary version published in STOC 1996. `doi:10.1006/jcss.1997.1545`.

[AS87] N. Alon and B. Schieber. *Optimal preprocessing for answering on-line product queries.* 1987. see here.

[AS03] V. Athitsos and S. Sclaroff. Database indexing methods for 3d hand pose estimation. In *Gesture-Based Communication in Human-Computer Interaction, 5th International Gesture Workshop, GW 2003, Genova, Italy, April 15-17, 2003, Selected Revised Papers*, pages 288–299, 2003. `doi:10.1007/978-3-540-24598-8\_27`.

[Bak94] B. S. Baker. Approximation algorithms for NP-complete problems on planar graphs. *Journal of the ACM*, 41(1):153–180, 1994. preliminary version published in FOCS 1983. `doi:10.1145/174644.174650`.

[Bar96] Y. 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`.

[Bar98] Y. Bartal. On approximating arbitrary metrices by tree metrics. In *Proceedings of the Thirtieth Annual ACM Symposium on the Theory of Computing, Dallas, Texas, USA, May 23-26, 1998*, pages 161–168, 1998. `doi:10.1145/276698.276725`.

[Bar04] Y. Bartal. Graph decomposition lemmas and their role in metric embedding methods. In *Algorithms - ESA 2004, 12th Annual European Symposium, Bergen, Norway, September 14-17, 2004, Proceedings*, pages 89–97, 2004. `doi:10.1007/978-3-540-30140-0\_10`.

[Bar21] Y. Bartal. Advances in metric ramsey theory and its applications. *CoRR*, abs/2104.03484, 2021. `arXiv:2104.03484`.

[BBM06] Y. Bartal, B. Bollobás, and M. Mendel. Ramsey-type theorems for metric spaces with applications to online problems. *J. Comput. Syst. Sci.*, 72(5):890–921, 2006. Special Issue on FOCS 2001. `doi:10.1016/j.jcss.2005.05.008`.

[BCL+18] S. Bubeck, M. B. Cohen, Y. T. Lee, J. R. Lee, and A. Madry. k-server via multiscale entropic regularization. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018, Los Angeles, CA, USA, June 25-29, 2018*, pages 3–16, 2018. `doi:10.1145/3188745.3188798`.

[Bec18] A. Becker. A tight 4/3 approximation for capacitated vehicle routing in trees. In 118, editor, *Approximation, Randomization, and Combinatorial Optimization. Algorithms* *and Techniques (APPROX/RANDOM 2018)*, pages 3:1–3:15, 2018. `doi:10.4230/LIPIcs.APPROX-RANDOM.2018.3`.

[BFC00] M. A. Bender and M. Farach-Colton. The LCA problem revisited. In *Latin American Symposium on Theoretical Informatics*, LATIN '00, pages 88–94, 2000. `doi:10.1007/10719839_9`.

[BFM86] J. Bourgain, T. Figiel, and V. Milman. On Hilbertian subsets of finite metric spaces. *Israel J. Math.*, 55(2):147–152, 1986. `doi:10.1007/BF02801990`.

[BGS16] G. E. Blelloch, Y. Gu, and Y. Sun. A new efficient construction on probabilistic tree embeddings. *CoRR*, abs/1605.04651, 2016. `https://arxiv.org/abs/1605.04651`. `arXiv:1605.04651`.

[BKS17] A. Becker, P. N. Klein, and D. Saulpic. A quasi-polynomial-time approximation scheme for vehicle routing on planar and bounded-genus graphs. In *25th Annual European Symposium on Algorithms*, volume 87 of *ESA '17*, pages 12:1–12:15, 2017. `doi:10.4230/LIPIcs.ESA.2017.12`.

[BKS19] A. Becker, P. N. Klein, and A. Schild. A PTAS for bounded-capacity vehicle routing in planar graphs. In *Algorithms and Data Structures*, WADS '19, pages 99–111, 2019. `doi:10.1007/978-3-030-24766-9_8`.

[BL16] G. Borradaile and H. Le. Optimal dynamic program for r-domination problems over tree decompositions. In *11th International Symposium on Parameterized and Exact Computation, IPEC 2016, August 24-26, 2016, Aarhus, Denmark*, pages 8:1–8:23, 2016. `doi:10.4230/LIPIcs.IPEC.2016.8`.

[BLMN05] Y. Bartal, N. Linial, M. Mendel, and A. Naor. On metric ramsey-type phenomena. *Annals of Mathematics*, 162(2):643–709, 2005. `doi:10.4007/annals.2005.162.643`.

[BLW17] G. Borradaile, H. Le, and C. Wulff-Nilsen. Minor-free graphs have light spanners. In *2017 IEEE 58th Annual Symposium on Foundations of Computer Science*, FOCS '17, pages 767–778, 2017. `doi:10.1109/FOCS.2017.76`.

[BM04] J. M. Boyer and W. J. Myrvold. On the cutting edge: Simplified o(n) planarity by edge addition. *J. Graph Algorithms Appl.*, 8(3):241–273, 2004. `doi:10.7155/jgaa.00091`.

[BTS94] H. L. Bodlaender, G. Tel, and N. Santoro. Trade-offs in non-reversing diameter. *Nord. J. Comput.*, 1(1):111–134, 1994. See: here.

[CAdVK+16] V. Cohen-Addad, É. C. de Verdière, P. N. Klein, C. Mathieu, and D. Meierfrankenfeld. Approximating connectivity domination in weighted bounded-genus graphs. In *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing*, pages 584–597. ACM, 2016.

[CFKL20] V. Cohen-Addad, A. Filtser, P. N. Klein, and H. Le. On light spanners, low-treewidth embeddings and efficient traversing in minor-free graphs. In *61th Annual IEEE Symposium on Foundations of Computer Science, FOCS*, pages 589–600, 2020. See: conference version, arXiv version,.

[CG04] D. E. Carroll and A. Goel. Lower bounds for embedding into distributions over excluded minor graph families. In *Algorithms - ESA 2004, 12th Annual European Symposium, Bergen, Norway, September 14-17, 2004, Proceedings*, pages 146–156, 2004. `doi:10.1007/978-3-540-30140-0\_15`.

[CG08] T.-H. H. Chan and A. Gupta. Small hop-diameter sparse spanners for doubling metrics. *Discrete & Computational Geometry*, 41(1):28–44, 2008. `doi:10.1007/s00454-008-9115-5`.

[CG12] T. H. Chan and A. Gupta. Approximating TSP on metrics with bounded global growth. *SIAM J. Comput.*, 41(3):587–617, 2012. preliminary version published in SODA 2008. `doi:10.1137/090749396`.

[Cha87] B. Chazelle. Computing on a free tree via complexity-preserving mappings. *Algorithmica*, 2(1-4):337–361, 1987. `doi:10.1007/bf01840366`.

[CJLV08] A. Chakrabarti, A. Jaffe, J. R. Lee, and J. Vincent. Embeddings of topological graphs: Lossy invariants, linearization, and 2-sums. In *49th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2008, October 25-28, 2008, Philadelphia, PA, USA*, pages 761–770, 2008. `doi:10.1109/FOCS.2008.79`.

[CZ00] S. Chaudhuri and C. D. Zaroliagis. Shortest paths in digraphs of small treewidth. part I: sequential algorithms. *Algorithmica*, 27(3):212–226, 2000. `doi:10.1007/s004530010016`.

[DFHT05] E. D. Demaine, F. V. Fomin, M. T. Hajiaghayi, and D. M. Thilikos. Fixed-parameter algorithms for (k, r)-center in planar graphs and map graphs. *ACM Trans. Algorithms*, 1(1):33–47, 2005. preliminary version published in ICALP 2003. `doi:10.1145/1077464.1077468`.

[DH05] E. D. Demaine and M. Hajiaghayi. Bidimensionality: New connections between FPT algorithms and PTASs. In *Proceedings of the Sixteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA'05, pages 590–601, 2005.

[DHK05] E. D. Demaine, M. Hajiaghayi, and K. Kawarabayashi. Algorithmic graph minor theory: Decomposition, approximation, and coloring. In *Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science*, pages 637–646, 2005. `doi:10.1109/SFCS.2005.14`.

[DHK11] E. D. Demaine, M. Hajiaghayi, and K. Kawarabayashi. Contraction decomposition in h-minor-free graphs and algorithmic applications. In *Proceedings of the 43rd ACM Symposium on Theory of Computing, STOC 2011, San Jose, CA, USA, 6-8 June 2011*, pages 441–450, 2011. `doi:10.1145/1993636.1993696`.

[DHM07] E. D. Demaine, M. Hajiaghayi, and B. Mohar. Approximation algorithms via contraction decomposition. In *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '07, pages 278–287, 2007.

[DM15] A. Das and C. Mathieu. A quasipolynomial time approximation scheme for euclidean capacitated vehicle routing. *Algorithmica*, 73(1):115–142, 2015. `doi:10.1007/s00453-014-9906-4`.

[DR59] G. B. Dantzig and J. H. Ramser. The truck dispatching problem. *Management science*, 6(1):80–91, 1959. `doi:10.1287/mnsc.6.1.80`.

[EKM14] D. Eisenstat, P. N. Klein, and C. Mathieu. Approximating k-center in planar graphs. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2014, Portland, Oregon, USA, January 5-7, 2014*, pages 617–627, 2014. `doi:10.1137/1.9781611973402.47`.

[ELS10] D. Eppstein, M. Löffler, and D. Strash. Listing all maximal cliques in sparse graphs in near-optimal time. In *International Symposium on Algorithms and Computation*, pages 403–414, 2010. `doi:10.1007/978-3-642-17517-6_36`.

[Epp00] D. Eppstein. Diameter and treewidth in minor-closed graph families. *Algorithmica*, 27(3):275–291, 2000. `doi:10.1007/s004530010020`.

[FFKP18] A. E. Feldmann, W. S. Fung, J. Könemann, and I. Post. A $(1+\epsilon)$-embedding of low highway dimension graphs into bounded treewidth graphs. *SIAM J. Comput.*, 47(4):1667–1704, 2018. preliminary version published in ICALP 2015. `doi:10.1137/16M1067196`.

[Fil20a] A. Filtser. A face cover perspective to ℓ_1 embeddings of planar graphs. In S. 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`.

[Fil20b] A. Filtser. Scattering and sparse partitions, and their applications. In A. Czumaj, A. Dawar, and E. 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`.

[Fil21] A. Filtser. Hop-constrained metric embeddings and their applications. In *62nd IEEE Annual Symposium on Foundations of Computer Science, FOCS 2021, Denver, CO, USA, February 7-10, 2022*, pages 492–503. IEEE, 2021. `doi:10.1109/FOCS52979.2021.00056`.

[Fis95] M. Fisher. Chapter 1 vehicle routing. In *Handbooks in Operations Research and Management Science*, pages 1–33. Elsevier, 1995. `doi:10.1016/s0927-0507(05)80105-7`.

[FKM15] K. Fox, P. N. Klein, and S. Mozes. A polynomial-time bicriteria approximation scheme for planar bisection. In *Proceedings of the 47th annual ACM symposium on Theory of Computing*, 2015. `doi:10.1145/2746539.2746564`.

[FKS19] 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, page 1069–1088, 2019. doi:10.1137/1.9781611975482.66.

[FL21] A. Filtser and H. Le. Clan embeddings into trees, and low treewidth graphs. In S. Khuller and V. V. 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. Full version at https://arxiv.org/abs/2101.01146. doi:10.1145/3406325.3451043.

[FL22] A. Filtser and H. Le. Reliable spanners: Locality-sensitive orderings strike back. In A. Gupta, editor, *STOC '22: 54rd Annual ACM SIGACT Symposium on Theory of Computing, June 20-24, 2022*. ACM, 2022. to appear, see full version.

[FLRS11] F. V. Fomin, D. Lokshtanov, V. Raman, and S. Saurabh. Bidimensionality and eptas. In *Proceedings of the 22nd Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '11, page 748–759, 2011.

[FLSZ18] F. V. Fomin, D. Lokshtanov, S. Saurabh, and M. Zehavi. *Kernelization*. Cambridge University Press, December 2018. doi:10.1017/9781107415157.

[FRT04] J. Fakcharoenphol, S. Rao, and K. 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.

[GKK17] L. Gottlieb, A. Kontorovich, and R. Krauthgamer. Efficient regression in metric spaces via approximate lipschitz extension. *IEEE Trans. Inf. Theory*, 63(8):4838–4849, 2017. preliminary version published in SIMBAD 2013. doi:10.1109/TIT.2017.2713820.

[GKP95] M. Grigni, E. Koutsoupias, and C. H. Papadimitriou. An approximation scheme for planar graph TSP. In *36th Annual Symposium on Foundations of Computer Science, Milwaukee, Wisconsin, USA, 23-25 October 1995*, pages 640–645, 1995. doi:10.1109/SFCS.1995.492665.

[GNRS04] A. Gupta, I. Newman, Y. Rabinovich, and A. Sinclair. Cuts, trees and l_1-embeddings of graphs. *Comb.*, 24(2):233–269, 2004. doi:10.1007/s00493-004-0015-x.

[HBK+03] E. Halperin, J. Buhler, R. M. Karp, R. Krauthgamer, and B. Westover. Detecting protein sequence conservation via metric embeddings. *Bioinformatics*, 19(suppl 1):i122–i129, 07 2003. doi:10.1093/bioinformatics/btg1016.

[HHZ21a] B. Haeupler, D. E. Hershkowitz, and G. Zuzic. Deterministic tree embeddings with copies for algorithms against adaptive adversaries. *CoRR*, abs/2102.05168, 2021. arXiv:2102.05168.

[HHZ21b] B. Haeupler, D. E. Hershkowitz, and G. Zuzic. Tree embeddings for hop-constrained network design. In S. Khuller and V. V. Williams, editors, *STOC '21: 53rd Annual*

ACM SIGACT Symposium on Theory of Computing, Virtual Event, Italy, June 21-25, 2021, pages 356–369. ACM, 2021. doi:10.1145/3406325.3451053.

[HKRS97] M. R. Henzinger, P. Klein, S. Rao, and S. Subramanian. Faster shortest-path algorithms for planar graphs. *Journal of Computer and System Sciences*, 55(1):3–23, 1997. doi:10.1006/jcss.1997.1493.

[HR85] M. Haimovich and A. H. G. Rinnooy Kan. Bounds and heuristics for capacitated routing problems. *Math. Oper. Res.*, 10(4):527–542, 1985. doi:10.1287/moor.10.4.527.

[HT84] D. Harel and R. E. Tarjan. Fast algorithms for finding nearest common ancestors. *SIAM J. Comput.*, 13(2):338–355, 1984. doi:10.1137/0213024.

[JS22] A. Jayaprakash and M. R. Salavatipour. Approximation schemes for capacitated vehicle routing on graphs of bounded treewidth, bounded doubling, or highway dimension. In *Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '22, pages 877–893. 2022. doi:10.1137/1.9781611977073.37.

[KKM+12] M. Khan, F. Kuhn, D. Malkhi, G. Pandurangan, and K. Talwar. Efficient distributed approximation algorithms via probabilistic tree embeddings. *Distributed Comput.*, 25(3):189–205, 2012. preliminary version published in PODC 2008. doi:10.1007/s00446-012-0157-9.

[KKR12] K. Kawarabayashi, Y. Kobayashi, and B. Reed. The disjoint paths problem in quadratic time. *Journal of Combinatorial Theory, Series B*, 102(2):424–435, 2012. doi:10.1016/j.jctb.2011.07.004.

[KKS11] K. Kawarabayashi, P. N. Klein, and C. Sommer. Linear-space approximate distance oracles for planar, bounded-genus and minor-free graphs. In *The 38th International Colloquium on Automata, Languages and Programming*, ICALP '11, pages 135–146, 2011. doi:10.1007/978-3-642-22006-7_12.

[Kle02] P. 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.

[Kle05] P. N. Klein. A linear-time approximation scheme for planar weighted TSP. In *Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '05, pages 647–657, 2005. doi:10.1109/SFCS.2005.7.

[KLMN05] R. Krauthgamer, J. R. Lee, M. Mendel, and A. 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.

[KLP19] I. Katsikarelis, M. Lampis, and V. T. Paschos. Structural parameters, tight bounds, and approximation for (k, r)-center. *Discret. Appl. Math.*, 264:90–117, 2019. preliminary version published in ISAAC 2017. doi:10.1016/j.dam.2018.11.002.

[KLP20] I. Katsikarelis, M. Lampis, and V. T. Paschos. Structurally parameterized d-scattered set. *Discrete Applied Mathematics*, 2020. preliminary version published in WG 2018. doi:10.1016/j.dam.2020.03.052.

[Kos82] A. V. Kostochka. The minimum Hadwiger number for graphs with a given mean degree of vertices. *Metody Diskret. Analiz.*, (38):37–58, 1982. doi:10.1090/trans2/132.

[LLR95] N. Linial, E. London, and Y. 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.

[MN07] M. Mendel and A. Naor. Ramsey partitions and proximity data structures. *Journal of the European Mathematical Society*, 9(2):253–275, 2007. preliminary version published in FOCS 2006. doi:10.4171/JEMS/79.

[MP15] D. Marx and M. Pilipczuk. Optimal parameterized algorithms for planar facility location problems using voronoi diagrams. In *Algorithms - ESA 2015 - 23rd Annual European Symposium, Patras, Greece, September 14-16, 2015, Proceedings*, pages 865–877, 2015. doi:10.1007/978-3-662-48350-3_72.

[MZ21] C. Mathieu and H. Zhou. A PTAS for capacitated vehicle routing on trees, 2021. doi:10.48550/ARXIV.2111.03735.

[NPSW22] J. Nederlof, M. Pilipczuk, C. M. F. Swennenhuis, and K. Wegrzycki. Isolation Schemes for Problems on Decomposable Graphs. In *39th International Symposium on Theoretical Aspects of Computer Science (STACS 2022)*, Leibniz International Proceedings in Informatics (LIPIcs), pages 50:1–50:20, 2022. Full version at https://arxiv.org/abs/2105.01465.

[NS07] G. Narasimhan and M. H. M. Smid. *Geometric spanner networks*. Cambridge University Press, 2007. doi:10.1017/CBO9780511546884.

[NT12] A. Naor and T. Tao. Scale-oblivious metric fragmentation and the nonlinear dvoretzky theorem. *Israel Journal of Mathematics*, 192(1):489–504, 2012. doi:10.1007/s11856-012-0039-7.

[PY93] C. H. Papadimitriou and M. Yannakakis. The Traveling Salesman Problem with distances one and two. *Mathematics of Operations Research*, 18(1):1–11, 1993. doi:10.2307/3690150.

[Rao99] S. 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.

[RS86] N. Robertson and P. Seymour. Graph minors. v. excluding a planar graph. *Journal of Combinatorial Theory, Series B*, 41(1):92–114, 1986. doi:https://doi.org/10.1016/0095-8956(86)90030-4.

[RS03] 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`.

[RST94] N. Robertson, P. Seymour, and R. Thomas. Quickly excluding a planar graph. 62(2):323–348, 1994. `doi:10.1006/jctb.1994.1073`.

[Sol11] S. Solomon. An optimal-time construction of sparse euclidean spanners with tiny diameter. In D. Randall, editor, *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2011, San Francisco, California, USA, January 23-25, 2011*, pages 820–839. SIAM, 2011. `doi:10.1137/1.9781611973082.65`.

[SV88] B. Schieber and U. Vishkin. On finding lowest common ancestors: Simplification and parallelization. *SIAM Journal on Computing*, 17(6):1253–1262, 1988. `doi:10.1137/0217079`.

[Tal04] K. Talwar. Bypassing the embedding: algorithms for low dimensional metrics. In *Proceedings of the 36th Annual ACM Symposium on Theory of Computing, Chicago, IL, USA, June 13-16, 2004*, pages 281–290, 2004. `doi:10.1145/1007352.1007399`.

[Tho84] A. Thomason. An extremal function for contractions of graphs. *Mathematical Proceedings of the Cambridge Philosophical Society*, 95(2):261–265, 1984. `doi:10.1017/s0305004100061521`.

[Tho97] M. Thorup. Parallel shortcutting of rooted trees. *Journal of Algorithms*, 23(1):139–159, 1997. `doi:10.1006/jagm.1996.0829`.

[Tho04] M. Thorup. Compact oracles for reachability and approximate distances in planar digraphs. *Journal of the ACM*, 51(6):993–1024, 2004. `doi:10.1145/1039488.1039493`.

[You12] On classes of summable functions and their fourier series. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 87(594):225–229, 1912. `doi:10.1098/rspa.1912.0076`.