On Approximability of the Minimum-Cost k-Connected Spanning Subgraph Problem

Artur Czumaj* Andrzej Lingas†

Abstract

We present the first truly polynomial-time approximation scheme (PTAS) for the minimum-cost k-vertex- (or, k-edge-) connected spanning subgraph problem for complete Euclidean graphs in $\mathbb{R}^d$. Previously it was known for every positive constant ε how to construct in a polynomial time a graph on a *superset* of the input points which is k-vertex connected with respect to the input points, and whose cost is within $(1+\varepsilon)$ of the minimum-cost of a k-vertex connected graph spanning the input points. We subsume that result by showing for every positive constant ε how to construct in a polynomial-time a k-connected subgraph spanning the input points without any Steiner points and having the cost within $(1+\varepsilon)$ of the minimum.

We also study hardness of approximations for the minimum-cost k-vertex- and k-edge-connected spanning subgraph problems. The only inapproximability result known so far for the minimum-cost k-vertex- and k-edge-connected spanning subgraph problems states that the k-edge-connectivity problem in unweighted graphs does not have a PTAS unless $\mathsf{P} = \mathsf{NP}$, even for $k = 2$. We present a simpler proof of this result that holds even for graphs of bounded degree, and provide the first proof that finding a PTAS for the k-vertex-connectivity problem in unweighted graphs is NP-hard even for $k = 2$ and for graphs of bounded degree. We further show that our algorithmic results for Euclidean graphs cannot be extended to arbitrarily high dimensions. We prove that for weighted graphs there is no PTAS for the k-vertex- and the k-edge-connectivity problem unless $\mathsf{P} = \mathsf{NP}$, even for Euclidean graphs in $\mathbb{R}^{\log n}$ and $k = 2$.

1 Introduction

Connectivity problems are fundamental in graph theory and they have many important applications in computer science and operation research. In this paper we study the central algorithmic problem of finding a *minimum-cost k-connected spanning subgraph* (this may be either k-vertex- or k-edge-connected spanning subgraph) of an undirected graph. Since this problem is known to be NP-hard even for $k = 2$ [7] (and even for Euclidean graphs on the plane [10]), polynomial-time approximation algorithms are the main object of the algorithmic research on minimum-cost k-connectivity.

In this paper we present new approximability and inapproximability results for the minimum-cost k-connectivity problems.

PTAS for minimum-cost k-connected spanning subgraph of a complete Euclidean graph. A *polynomial-time approximation scheme* (PTAS) for an optimization problem is a family of algorithms $\{\mathcal{A}_\varepsilon\}$ such that for each fixed $\varepsilon > 0$, $\mathcal{A}_\varepsilon$ runs in time polynomial in the size of the input and produces a $(1+\varepsilon)$-approximate solution [8].

We consider the problem of finding a minimum-cost k-vertex-connected spanning subgraph of a complete Euclidean graph in $\mathbb{R}^d$. The cost of an edge in such a graph is equal to the Euclidean distance between its endpoints. Our main result is the design of the first truly PTAS for the Euclidean minimum-cost k-vertex-connectivity and k-edge-connectivity problem in $\mathbb{R}^d$. For every given c, we design a randomized algorithm running in time $n \cdot 2^{(\log\log n)^{\binom{2d}{2}+2d} \cdot ((\mathcal{O}(cdk^2))^d \log c)!}$ and producing an Euclidean k-vertex (or, k-edge, respectively) connected spanning subgraph whose cost is within $(1+\frac{1}{c})$ of the minimum. We can further derandomize the algorithm by increasing the running time by a factor $\mathcal{O}(n)$.

Our PTAS is based on a subtle decomposition of k-connected Euclidean graphs combined with the general framework proposed recently by Arora [1] for designing PTAS for Euclidean versions of TSP, Minimum Steiner Tree, Min-Cost Perfect Matching, k-TSP, and k-MST. It significantly subsumes our earlier work presented in [5], where approximation algorithms for a restricted version of the problem were obtained. In [5] we presented an approximation scheme for the Euclidean minimum-cost k-vertex- (and k-edge-) connectivity in $\mathbb{R}^d$ that in time $n \cdot (\log n)^{(\mathcal{O}(c\sqrt{d}\,k))^{d-1}} \cdot 2^{((\mathcal{O}(c\sqrt{d}\,k))^{d-1})!}$ constructs an Euclidean graph on a *superset* of the input points, which is k-vertex-connected (k-edge-connected, respectively) with respect to the input point set and with probability $\frac{1}{2}$ has cost within $(1+\frac{1}{c})$ of the minimum-cost of a k-vertex-connected (k-edge-connected) Euclidean graph on the input point set. (In [5], the cost of the output graph using Steiner points is compared to the cost of the best graph with no Steiner points instead of the best graph that may use Steiner points.)

The use of Steiner points seems to be crucial in all

*Department of Mathematics & Computer Science and Heinz Nixdorf Institute, University of Paderborn, D-33095 Paderborn, Germany. E-mail: `artur@uni-paderborn.de`. Research supported in part by ALCOM EU ESPRIT Long Term Research Project 20244 (ALCOM-IT) and by DFG Research Cluster "Efficient Algorithms for Discrete Problems and their Applications."

†Department of Computer Science, Lund University, Box 118, S-22100 Lund, Sweden. E-mail: `Andrzej.Lingas@dna.lth.se`. Research supported in part by TFR grant 96-278.

previous applications of Arora's framework [1, 5]. (In all applications considered by Arora [1] and Rao and Smith [13] it was relatively easy to deal with Steiner points because of the low connectivity of the output graphs.) Following the general approach of Arora's framework, in the k-connectivity problem it is very natural to introduce (in the analogue of the so-called patching procedure) Steiner points of degree at least 3 (cf. [5]). However, it is difficult to remove these points cheaply, because their high degree is critical to ensure k-connectivity with respect to the input points. Our approach is different: We disallow introducing any Steiner points (with the exception of those of degree 2 that are used to bend the edges and which are easily removable). This forces us to define the main procedure of modifying the input graph (related to the patching procedure in [1, 5]) to work only on the input points. To achieve this, we limit the application domain of our patchings solely to edges whose length is within $2\sqrt{d}$ of the size of the crossed facet. Only with such a restriction we can obtain a decomposition procedure that reduces the number of facets' crossings and does not increase the cost of the obtained graph too significantly (compare with [1, 5]). This restriction requires an additional special care in the proof of the Structure Theorem and in the dynamic programming. In particular, the proof of our Structure Theorem has an interesting multiple level structure in a nontrivial way relying on the techniques used in [1, 5, 13].

We apply also our arguments to design a PTAS for the Euclidean k-edge-connectivity problem.

Inapproximability results for the connectivity problems. Our further results concern hardness of approximation for the minimum-cost k-vertex- and k-edge-connected spanning subgraph problem. To the best of our knowledge (cf. also [4]), the only known inapproximability result for the connectivity problems is due to Fernandes [6]. She shows that the k-edge-connectivity problem in unweighted graphs is Max-SNP-hard even for $k = 2$. By the characterization of Max-SNP provided in [2], this result implies that the minimum-cost k-edge-connected spanning subgraph problem does not have PTAS unless P = NP. In this paper we add new basic results to the collection of problems hard to approximate. We present a general method of reducing approximation of TSP in the so called 1-2-Δ graphs to approximation of the connectivity problems. Since TSP in 1-2-Δ graphs is Max-SNP-hard, this reduction implies inapproximability results for the connectivity problems (unless P = NP). We apply this technique to prove the non-existence of a PTAS (unless P = NP) for the k-vertex- and k-edge-connectivity problem in unweighted graphs of bounded degree even for $k = 2$. We also prove that our algorithmic results for Euclidean graphs cannot be extended to high dimensions. Similarly as it is known already for Euclidean TSP [14], we prove that there is no PTAS for the k-vertex- and the k-edge-connectivity problem (unless P = NP) for Euclidean graphs in $\mathbb{R}^{\log n}$ even for $k = 2$.

Related works. For a very extensive presentation of results concerning problems of finding minimum-cost k-vertex- and k-edge-connected spanning subgraphs, non-uniform connectivity, connectivity augmentation problems, and geometric problems, we refer the reader to various chapters of [8], especially to [9].

Organization of the paper. In Sections 2 and 3 we provide basic terminology and present structural lemmas used in our PTAS in case $\mathbb{R}^2$ (the case $\mathbb{R}^d$ with $d > 2$ is deferred to the full version of the paper). Section 4 contains a PTAS for the minimum-cost k-connected spanning subgraph problem in Euclidean graphs. Finally, in Section 5 we present our inapproximability results.

2 Decomposition Lemmata and Structure Theorem

We shall adhere to standard notation on approximation graph algorithms [8] and follow in general terminology used in [1, 5]. We need introduce more specific notation on geometric graphs and partitioning.

DEFINITION 1. *Let S be a set of points in $\mathbb{R}^2$. A* straight-line graph *(SG) on S is a graph whose vertices are in one-to-one correspondence with the points in S, and whose edges are in one-to-one correspondence with the straight-line segments connecting the points corresponding to the incident vertices. The* cost *of an edge is the length of the corresponding segment, and the cost of the graph is the sum of the costs of its edges.*

In the paper we shall assume a more general form of geometric graphs. In our construction we allow the edges to be *straight-line paths* (i.e., paths consisting of straight-line segments) connecting the endpoints. This relaxation enables the edges to pass through some prespecified points (called portals) by bending at them.

DEFINITION 2. *A* straight-line path graph *(SPG) on a multiset of points in $\mathbb{R}^2$ is a graph whose vertices one-to-one correspond to the elements of this multiset and whose edges one-to-one correspond to straight-line paths connecting the points corresponding to the incident vertices. The cost of a SPG is the total length of the paths.*

DEFINITION 3. *An SP (SPG) on S is* k-vertex-connected *if for each pair of points in S there are at least k internally vertex-disjoint paths connecting the corresponding vertices in the SP (SPG, respectively).*

Let the *bounding box* of the input multiset of points in $\mathbb{R}^2$ be any axis-aligned square W^2 in $\mathbb{R}^2$ that contains it. Following [1, 5], the geometric partitioning of the bounding box can be defined as follows.

DEFINITION 4. *A* (4-ary) dissection *of the bounding box W^2 of a multiset of points in $\mathbb{R}^2$ is its recursive partitioning into smaller squares, called* regions. *Each* region U^2 *of volume > 1 is recursively partitioned into 4 regions $(U/2)^2$.* *A* quadtree *is a tree whose root corresponds to the bounding box, and whose other non-leaf nodes correspond to the regions containing at least two points from the input multiset. For a non-leaf node v of the tree, the nodes corresponding to the 4 regions partitioning the region corresponding to v are the children of v in the tree.*

Note that the dissection has $\Theta(W^2)$ regions and its recursion depth is logarithmic in W.

DEFINITION 5. *For any a, where $0 \le a < W$, the a-*shifted *dissection of the bounding box W^2 is defined by shifting all the coordinates of the hyperplanes forming the dissection by a, and then reducing them modulo W.*

The *a-shifted quadtree* is defined analogously. [1]

To each region of a dissection we assign a *level* in the natural way: the bounding box is at level zero, its 4 sub-regions are at level one, and in general, a region $(W/2^i)^2$ is at level i. We assign also a level to each facet of a region to be the level of that region. Similarly, we say a hyperplane $\mathcal{H}$ has level i in the dissection if it contains a facet at level i. Since if $\mathcal{H}$ has level i in the dissection then it has also level j for all $j > i$, we define the *minimal level* of $\mathcal{H}$ to be the minimal level $\mathcal{H}$ has (it was called "maximal" in [1, 3, 13]).

DEFINITION 6. *A crossing of an edge (a straight-line path) with a region facet of side length W at level j in a dissection of the bounding box is called* relevant *if the length of the edge is at most $\sqrt{2}\,W \,/\, 2^{j-1}$.*

DEFINITION 7. [1] *An m-*regular *set of* portals *in a segment of length N is a set of m points in the segment where the spacing between the portals is $(N+1)\cdot m^{-1}$.*

The restricted forms of SPG considered in the Structure Theorem can be formalized as follows.

DEFINITION 8. *An SPG is* r-gray *with respect to a shifted dissection if each facet of each region in the dissection has at most r relevant crossings. An SPG is* r-light *with respect to a shifted dissection if each facet of each region in the dissection is crossed at most r times. If an SPG is r-light and each crossing is through one of the m portals within the facet then the SPG is called* (m,r)-light. *(The paths corresponding to the edges can cross a portal up to m times, possibly bending at it.)*

[1]Some of the regions of the shifted dissection are wrapped around. Assume a to be only in $\{0,\ldots,\frac{W}{2}\}$, and define the bounding box such that all input points have all coordinates in $(\frac{W}{2}, W)$. Then the regions that are wrapped around are always pointless and as such can be ignored.

We shall call a multiset of n points in $\mathbb{R}^2$ *well-rounded* if it contains $\mathbf{0}$, all its points have non-negative integral coordinates, the volume of their bounding box is at most $\mathcal{O}(n^2)$, and the distance between any two points is at least 8 or is zero.

The proof of the following lemma is similar to the proof of the Perturbation Lemma presented by Arora [1] (see also Lemma 3 in [13]). Additionally, it relies on the fact that any minimum-cost SG in $\mathbb{R}^2$ which is k-vertex-connected with respect to n points (after possible removing zero-length edges) contains $\mathcal{O}(k \cdot n)$ edges.

LEMMA 2.1. **(Perturbation Lemma)** [2] *Suppose there is a PTAS for well-rounded instances of the minimum-cost k-connectivity problem for SGs in $\mathbb{R}^2$. Then there is a PTAS for all instances of the minimum-cost k-connectivity problem for SGs in $\mathbb{R}^2$.*

To state the Local Decomposition Lemma we need the following definition.

DEFINITION 9. *Let X be a set of points in $\mathbb{R}^2$ and let k be any positive integer. A* k-augmented traveling salesman tour *on X is either a clique on X if $|X| \le 2k$, or the kth power [3] of some traveling salesman tour on X if $|X| \ge 2k+1$.*

The following two lemmas play the crucial role in our construction.

LEMMA 2.2. **(Local Decomposition Lemma)** *Let G be a k-vertex-connected SPG in $\mathbb{R}^2$ on a set of points S. Let $\mathcal{F}$ be a 1-dimensional facet of side length W in a dissection of the bounding box of S. If the edges of G form l relevant crossings of $\mathcal{F}$, then there exist a constant α, a subgraph G^* of G, and two disjoint subsets S_1 and S_2 of S such that*

- *G^* crosses $\mathcal{F}$ at most k^2 times,*
- *S_1 and S_2 are contained in the opposite half-spaces induced by $\mathcal{F}$, respectively,*
- *each of S_1 and S_2 consists of at most l vertices, and thus there are a traveling salesman tour on S_1 and a traveling salesman tour on S_2 such that cost of each is bounded by $\frac{\alpha}{2}\cdot W \cdot \sqrt{l}$, and*
- *the graph resulting from G^* by adding any k-augmented travelings salesman tour on S_1 and any k-augmented travelings salesman tour on S_2 is a k-vertex-connected SPG on S.*

[2]The lemma assumes that ε, the quality of approximation, is a constant. If we consider ε as a parameter of the problem, then we must change the definition of the well-rounded multiset of n points in $\mathbb{R}^2$ so that the volume of their bounding box is at most $\mathcal{O}(n^2(1+\frac{1}{\varepsilon})^2)$, instead of $\mathcal{O}(n^2)$.

[3]A *kth power* of a simple graph G is obtained from G by connecting by the edge each pair of vertices in G for which there is a path consisting of at most k edges in G.

Proof. Since the lemma is trivial for $l \leq k^2$, we shall assume that $l > k^2$.

Construction of G^*, S_1 and S_2. G^* is a subgraph of G containing all the vertices of G and two edge sets, E_G and $E_{\mathcal{F}}$. To define E_G, consider the set E_1^2 of the l edges of G forming the l relevant crossings with $\mathcal{F}$. E_G is the set of the edges of G resulting from removing the edges in E_1^2. To define $E_{\mathcal{F}}$, consider a maximum cardinality subset $\mathbb{M}$ of E_1^2 such that no two edges in the subset are incident. Let $q = \min\{k, |\mathbb{M}|\}$. Then $E_{\mathcal{F}}$ consists of

- the first q edges of $\mathbb{M}$, and
- if $q < k$, then, additionally, for each endpoint v of any edge from $\mathbb{M}$ we add to $E_{\mathcal{F}}$ $\min\{k-1, deg_{E_1^2}(v) - 1\}$ edges in $E_1^2 - \mathbb{M}$ incident to v, where $deg_{E_1^2}(v)$ is the number of edges in E_1^2 incident to v.

We define $S_1 = \{x_1, \dots, x_l\}$ as the set of endpoints of the edges in E_1^2 in the first half-space induced by $\mathcal{F}$ and $S_2 = \{y_1, \dots, y_l\}$ as the corresponding set of endpoints of these edges in the other half-plane.

Properties of G^*, S_1 and S_2. Clearly, $E_{\mathcal{F}}$ is of size at most k^2, and therefore G^* crosses $\mathcal{F}$ at most k^2 times. Furthermore, each of S_1 and S_2 consists of at most l vertices that are included in some bounding box of size $\mathcal{O}(W)$. Therefore, there is a traveling salesman tour on each of S_1 and S_2 of total length smaller than $\frac{\alpha}{2} \cdot W \cdot \sqrt{l}$ (see, e.g., Section 6 in [11]), where α is a positive constant [4].

It remains to prove the last property of the decomposition. Let X and Y be any k-augmented travelings salesman tours on S_1 and S_2, respectively. Let G^+ be the graph resulting from augmenting G^* by the edges of X and Y (it is clearly enough to add only the edges in X and Y that do not appear in G^*). To prove the k-vertex-connectivity requirement of the augmented graph G^+, we show that if we pick any set U of $k-1$ vertices in G^+ then the graph H^+ obtained by removing U from G^+ is connected.

Let H be the graph resulting from G by removing the vertices in U. Let v and u be any pair of vertices in H^+. It is sufficient to show that there exists a path in H^+ connecting v and u. This clearly would imply that H^+ is connected, and hence, that G^+ is k-vertex-connected.

Since G is k-connected, there is a path $\mathcal{P}$ connecting v and u in H. If $\mathcal{P}$ contains at most one vertex from $S_1 \cup S_2$ then $\mathcal{P}$ exists also in H^+, and therefore v and u are connected in H^+. Therefore let us consider the case that $\mathcal{P}$ contains at least two vertices from $S_1 \cup S_2$, and let s and t be the first and the last such a vertex. If s and t both are either in S_1 or in S_2, then there is a path connecting them in H^+, since the graph X or Y is either k-connected or it is a clique. This implies the existence of a path connecting v and u in H^+. Suppose now that one of s, t, say s, is in S_1 and the other is in S_2. If at least one of the edges in $E_{\mathcal{F}}$ is present in H^+, then it is sufficient to connect s to the endpoint of this edge in S_1 and t to the other endpoint in S_2 using the edges in X and Y, respectively. These two additional connections exist, again, by the k-connectivity or completeness of the graphs X or Y, respectively. Thus let us consider the case that no edge in $E_{\mathcal{F}}$, and in particular no edge in $\mathbb{M}$, is present in H^+. It follows also that $q < k$. Note that if $\mathcal{P}$ does not traverse any edge in E_1^2 that is not in H^+, then $\mathcal{P}$ clearly exists in H^+. Therefore we suppose now that $\mathcal{P}$ uses some edge $e = (x, y)$ in E_1^2 that does not exist in H^+. Clearly, $x, y \notin U$, and therefore $e \notin E_{\mathcal{F}}$. Observe also that by the maximality of $\mathbb{M}$ at least one of x, y, say x, must be incident to an edge in $\mathbb{M}$. Consequently, by the construction of $E_{\mathcal{F}}$ and since e has not been chosen to $E_{\mathcal{F}}$, there must be at least $k-1$ other edges in $E_{\mathcal{F}}$ outside of M incident to x. Therefore x is adjacent in G^+ to at least k vertices through the edges from $E_{\mathcal{F}}$. This implies that at least one of these edges must exist in H^+. Thus we have proven that v and u must be connected in H^+.

[4] Note that we need here the assumption that each of S_1 and S_2 is included in a bounding box of size $\mathcal{O}(W)$. In contrast, in the Patching Lemmas used in [1, 5, 13] the points could be arbitrarily far away from each other and thus, for example, there could be no TST on S_1 of length $o(n)$.

LEMMA 2.3. **(Global Decomposition Lemma)** *Let $c > 0$ and let W be the size of the smallest bounding box for a well-rounded point multiset S in $\mathbb{R}^2$. Let G be any k-vertex-connected SG on S. Pick $a \in \{0, \dots, W-1\}$ uniformly at random. Then there is a subgraph G^* of G and a collection $\mathbb{S}$ of (possible intersecting) subsets of S such that:*

- *G^* is r-gray with respect to the a-shifted dissection, with $r = \mathcal{O}(c^2\, k^4)$,*
- *the graph resulting from G^* by adding any k-augmented traveling salesman tours (in the form of an SPG) on each set $X \in \mathbb{S}$ is a k-vertex-connected SPG on S, and*
- *there is a graph H consisting of (possible non-disjoint) traveling salesman tours on every set $X \in \mathbb{S}$ whose expected (over the choice of a) total cost is $\frac{1}{8\, c\, k^2}$ times the cost of G.*

Proof. Our approach is similar to that introduced by Arora [1]. However in a few crucial aspects our proof is different. Firstly, the Local Decomposition Lemma is applied to the 0-shifted dissections of hyperplanes whereas in [1] the corresponding Patching Lemma is applied to the a-shifted dissection of the hyperplanes. The other differences follow from the restricted form of the Local Decomposition Lemma applicable solely to relevant crossings and causing only edge removal. For the latter reason, we need to perform a separate "charging" analysis for the traveling salesman tours on the sets in $\mathbb{S}$ ensuring k-connectivity.

Let $\mathcal{H}$ be a unit hyperplane perpendicular to some coordinate axis that has the minimal level i in the a-shifted dissection. Within the bounding box $\mathcal{H}$ forms a 1-dimensional segment of size W. By the *standard dissection* of $\mathcal{H}$ (within the bounding box) we shall mean the 0-shifted dissection of the aforementioned segment. For $0 \leq j \leq \lceil \log W \rceil$, let $\mathbb{F}_j(\mathcal{H})$ denote the set of the 4^j cubes at level j in the standard dissection of $\mathcal{H}$.

Let $r^* = \lfloor \frac{r}{2} \rfloor$. We incorporate the Local Decomposition Lemma to modify G by decreasing the number of relevant crossings of $\mathcal{H}$ using the following bottom-up procedure $\mathbb{MODIFY}(i, \mathcal{H})$ (We emphasize here that G is modified with respect to the *standard dissection* of $\mathcal{H}$.).

$\mathbb{MODIFY}(i, \mathcal{H})$:

Sequentially, for $j = \lceil \log W \rceil$ down to i do

for every $\mathcal{F} \in \mathbb{F}_j(\mathcal{H})$:
 if there are more than r^* relevant crossings
 (at level j) of $\mathcal{F}$ in G then apply
 the Local Decomposition Lemma to $\mathcal{F}$.

It is easy to see that after applying $\mathbb{MODIFY}(i, \mathcal{H})$ no facet in $\mathbb{F}_i(\mathcal{H})$ is crossed by a relevant edge remaining in the SPG more than r^* times. Therefore, since each facet at level i of a region lying in $\mathcal{H}$ in the a-shifted dissection is covered by at most 2 facets in $\mathbb{F}_i(\mathcal{H})$, after applying $\mathbb{MODIFY}(i, \mathcal{H})$ no such a facet has more than $2 \cdot r^* \leq r$ relevant crossings.

We transform G into G^* by repeatedly modifying G to make it r-gray with respect to the a-shifted dissection. For that we apply procedure $\mathbb{MODIFY}(i, \mathcal{H})$ independently to every unit hyperplane $\mathcal{H}$ intersecting the bounding box, where i is the minimal level of $\mathcal{H}$.

Each invocation of the Local Decomposition Lemma deletes some relevant crossings of a region facet and creates two subsets of points in S for which any k-augmented traveling salesman tour added to the modified graph would make it k-vertex-connected. All these subsets introduced by the invocations of the Local Decomposition Lemma form the collection $\mathbb{S}$. By the Local Decomposition Lemma, this yields the second claim in the lemma. To prove the third claim, we bound the cost of the traveling salesman tours promised by choosing a bottom up order of applications of the Local Decomposition Lemma to the facets of each hyperplane, and by charging TST of each $X \in \mathbb{S}$ to some unit grid hyperplane. Using arguments similar to those in [1, 13] and in [5], we can show that the expected (over the choice of a) total cost of the TSTs on all the sets in $\mathbb{S}$ introduced by the invocation of $\mathbb{MODIFY}$ is at most

$$\beta \cdot \frac{\sqrt{r^*+1}}{r^*+1-k^2} \cdot \mathrm{cost}(G) \ ,$$

for some positive constant β. We set $r^* = \mathcal{O}(c^2\, k^4)$ to bound this cost by $\frac{1}{8\,c\,k^2} \cdot \mathrm{cost}(G)$.

The Global Decomposition Lemma will be used to transform G into an r-gray graph. Even if in an r-gray graph each facet has only $\mathcal{O}(r)$ relevant crossings, many other crossings are possible.

Consider a k-vertex-connected SPG $\tilde{G}$ on a multiset S in $\mathbb{R}^2$ that is r-gray with respect to a shifted dissection. Let $\mathcal{F}$ be a facet of a region Q in the dissection. Let E_i be the set of edges of length in interval $[2^i,\ 2^{i+1})$ that cross $\mathcal{F}$ and have an endpoint in Q. Finally, let $s \in \mathbb{N}$ and $t_s = \sum_{i=0}^{s} |E_i|$. We can apply the Local Decomposition Lemma to remove all but at most k^2 edges in $E_0 \cup \cdots \cup E_s$ and obtain two disjoint sets of nodes X and Y of size at most t_s each, such that the resulting graph with added any k-augmented TSP on X and k-augmented TSP on Y is k-vertex-connected, and for each of X, Y, there exists a TSP of length $\mathcal{O}(2^s \cdot \sqrt{t_s})$. Using this, we can derive the following important lemma.

LEMMA 2.4. **(Lightening Lemma)** *Let $\tilde{G}$ be a k-vertex-connected SPG on a multiset S in $\mathbb{R}^2$ that is r-gray with respect to a shifted dissection. Let $\mathcal{F}$ be a facet of a region Q in the dissection. Let E_i be the set of edges of length in interval $[2^i,\ 2^{i+1})$ that cross $\mathcal{F}$ and have an endpoint in Q. If $\sum_{i \geq 0} |E_i| \geq 3\,r\,\log\log W$, then we can reduce the number of the crossings in $\bigcup_{i\geq 0} E_i$ to at most $3\,r\,\log\log W$ such that each of the obtained sets X and Y has a TSP of length bounded by $\frac{1}{32\,c\,k^2\,\log W}$ times the total length of the original edges in $\bigcup_{i\geq 0} E_i$.*

Moreover, the total TSP cost incurred by such reductions over all region facets in the dissection is bounded by $\frac{1}{8\,c\,k^2}$ times the cost of $\tilde{G}$.

To formulate our Structure Theorem we need also the following technical definition.

DEFINITION 10. *Let G be an SPG whose vertex set is included in the bounding box B of size W. We call G (m, r)-*blue *with respect to a shifted dissection of B if G is r-gray with respect to the dissection and for each facet $\mathcal{F}$ of each region Q in the dissection, the set of the edges crossing $\mathcal{F}$ and having an endpoint in Q can be partitioned into two sets $E_1,\ E_2$ such that* (i) *each edge in E_1 crosses $\mathcal{F}$ through a portal of $\mathcal{F}$,* (ii) $|E_1| = \mathcal{O}(r\,\log c)$, (iii) *all edges in E_2 cross $\mathcal{F}$ through the same middle portal of $\mathcal{F}$, and* (iv) $|E_2| = \mathcal{O}(r\,\log\log W)$.

Now we are ready to prove our Structure Theorem.

THEOREM 2.1. **(Structure Theorem)** *Let $c > 0$. Let W be the size of the bounding box for a well-rounded point multiset S in $\mathbb{R}^2$. Pick $a \in \{0, \ldots, W-1\}$ uniformly at random. With probability at least $\frac{1}{2}$, there is a k-vertex-connected SPG on S that is (m, r)-blue with respect to the a-shifted dissection and has cost within $1 + \frac{1}{c}$ of the minimum cost of a k-vertex-connected SG on S, where $m = \mathcal{O}(c \log W)$ and $r = \mathcal{O}(c^2\, k^4)$.*

Proof. Let G be a minimum-cost k-vertex-connected SG on the well-rounded input multiset. We shall modify G to an SPG satisfying the theorem requirements in five main stages.

Stage 1: We move each relevant crossing of a facet in the a-shifted dissection to its nearest portal. If we set $m = \mathcal{O}(c \cdot \log W)$, then we can show that the expected cost of this modification is at most $\frac{1}{8\,c}$ times the cost of the SG G.

Stage 2: We apply the Global Decomposition Lemma to the SG G. Let G^* be the SG and $\mathbb{S}$ be the collection of subsets of S guaranteed by the lemma. Note that G^* is r-gray and since it is a subgraph of G, its cost does not exceed the cost of G.

Stage 3: We modify G^* by moving some of the non-relevant crossings of the facets in the a-shifted dissection to (not necessarily closest) portals in order to make the resulting SPG (m, r)-blue. (Let us remind that the relevant crossings were moved to the portals in Stage 1.)

For any region Q in the dissection, let P^Q denote the set of portals on the facets of Q. Additionally, let P_0^Q be the subset of P^Q consisting of each middle portal on each facet of Q.

Since G^* is r-gray with respect to the dissection, for any region Q in the dissection of side length L and any non-negative integer j, there are $\mathcal{O}(r)$ edges of length in the range $[2^j\,L,\ 2^{j+1}\,L)$ having an endpoint in the region and crossing the facets of Q. Observe that if we move such a crossing by bending the edge at a portal on the facet of Q at a distance bounded by $\frac{2^{j+1}L}{\alpha}$, then the cost of the bent edge increases at most by $\mathcal{O}(\frac{1}{\alpha})$ of its initial cost. We modify the SPG G^* by moving such a (non-relevant) crossing to the nearest portal either in P_0^Q if $\frac{2^{j+1}L}{\gamma\,c} \geq L$, or in P^Q otherwise. By choosing the constant γ appropriately large and straightforward geometric argumentation, we can ensure that all such crossing movements increase the cost of G^* by at most $\mathcal{O}(\frac{\log W}{m}) \leq \frac{1}{8\,c}$ of its initial cost.

Stage 4: The graph resulting from Stage 3 is (m, r)-blue apart from the fact that $|E_2| = \mathcal{O}(r\,\log W)$. Now, we apply the Lightening Lemma to reduce the size of E_2 to $\mathcal{O}(r\,\log\log W)$. The resulting graph is (m, r)-blue. All the sets X and Y obtained by applying the Lightening Lemma are added to the set $\mathbb{S}$.

Stage 5: We construct an SPG H^* on S that is (m, r)-light, whose expected cost is at most $\frac{1}{4\,c}$ times the cost of G, and such that $G^* \cup H^*$ is a k-vertex-connected SPG on S. This clearly yields the theorem.

Let T be the set of TSTs for the subset of S in $\mathbb{S}$ guaranteed by the Global Decomposition Lemma in Stage 2 and the Lightening Lemma in Stage 4. Observe that if SPG $\overline{H}$ was defined as the union of the k-augmented TSTs in T, then the cost of $\overline{H}$ would be at most $k^2 \cdot \frac{1}{8\,c\,k^2} = \frac{1}{8\,c}$ times the cost of G and $G^* \cup \overline{H}$ would be k-vertex-connected on S. This construction, however, would not guarantee the (m, r)-lightness of $\overline{H}$.

Therefore our approach is to "approximate" $\overline{H}$ defined above. Our aim is to find a set T^* of TSTs for a partition $\mathbb{S}^*$ of $\bigcup_{X \in \mathbb{S}} X$ such that $\mathbb{S}^*$ is a cover of $\mathbb{S}$ (i.e., if $X \in \mathbb{S}$ then there is a $Y \in \mathbb{S}^*$ such that $X \subseteq Y$), the union of the TSTs in T^* is $(m, \frac{r}{k^2})$-light with respect to the a-shifted dissection, and the expected sum of the costs of the TSTs in T^* is at most $\frac{1}{4\,c\,k^2}$ times the cost of G. Note that the graph H^* defined as the union of the k-augmented TSTs in T^* satisfies the required properties (in particular, since $\mathbb{S}^*$ is a cover of $\mathbb{S}$, $G^* \cup H^*$ is k-vertex-connected).

We apply a modification of the approach of Rao and Smith [13]. We first construct a sparse γ-spanner Q on S whose cost is $\mathcal{O}(\mathrm{MST})$, where MST denotes the cost of minimum spanning tree of S and γ is a constant to be chosen at the end of the proof. Next we transform the spanner (by breaking some of its edges and adding some new line segments, see [13]) so the resulting graph F on S with Steiner points is $(m, \mathcal{O}(c\,k^2))$-light with respect to the a-shifted dissection and the expected total length of the line segments in F added to Q is at most $\frac{\mathrm{MST}}{c\,k^2}$. (For this we use the analysis of Rao and Smith [13] modified by our analysis — similar to that given in the proof of the Global Decomposition Lemma — that enables us to use one shift for all directions.)

Now we emulate in F each TST in T by applying the arguments of Rao and Smith [13], so that each line segment of F is used at most twice. Whenever two or more of such emulated TSTs overlap we connect them together into a single one inductively using Arora's TSP patching (of zero cost in this case, cf. p. 15 in [1]) and Eulerian tours. Hence each line segment of F is still used at most twice. Thus, we obtain a new set T^* of TSTs such that the vertices of each TST from T are covered by exactly one TST in T^*. By our arguments given above, the total cost of the TSTs in T^* is at most $2\,\gamma\,\frac{1}{c\,k^2}$ times the cost of G plus $\frac{2\,\mathrm{MST}}{c\,k^2}$. Therefore, if we define H^* as the union of the k-augmented TSTs in T^*, then the cost of H^* is at most $\frac{2\,\gamma}{c}$ times the cost of G plus $\frac{2\,\mathrm{MST}}{c}$. Now we set the parameter γ so that the expected cost of H^* is at most $\frac{1}{4\,c}$ times the cost of G.

3 Computing a minimum-cost (m, r)-blue SPG

Let us fix a random shift a and set m to $\mathcal{O}(c\,\log W)$. By our Structure Theorem, it is sufficient to find a minimum-cost k-vertex-connected SPG on the input multiset S that is (m, r)-blue with respect to the shifted dissection. The Structure Theorem ensures that for most of the choices of the shift a the obtained SPG is a $1 + \frac{1}{c}$ approximation of the minimum-cost k-vertex-connected SG on S.

An *SPG within a region* is an SPG on the part of the multiset S contained in the region of the shifted binary dissection and on a multiset of some portals on the facets of the region such that the straight-line paths corresponding to its edges lie within the region. It is (m, r)-*blue* if it does not include any edge whose both endpoints correspond to the portals and the multiset of its vertices corresponding to the portals can be partitioned into two multisets V_1, V_2 such that $|V_1| = \mathcal{O}(r \log c)$, $|V_2| = \mathcal{O}(r \log\log W)$ and V_2 is a multiset on a subset of the set of vertices corresponding to the middle portals on the region facets.

Our dynamic programming procedure will combine "optimal" (m, r)-blue SPGs within sibling regions on the same level in a bottom-up manner. *Importantly, while combing two such graphs we shall alternately consider their extensions by matchings of portal vertices in order to allow the possibility of edges incident to a vertex in S within the union of the graphs that cross one of the two adjacent regions.* This is the main difference between our dynamic procedure and that for (m, r)-light SPGs considered in [5] (as well as in [1, 13]).

Our procedure will determine for each region, for each multiset P of at most $\mathcal{O}(r \log\log W)$ portals chosen from a subset of $\mathcal{O}(r \log c)$ portals of the region, and for each possible *connectivity characteristic*[5] of an (m, r)-blue SPG within the region, the cost of an optimal (m, r)-blue SPG within the region using P and having these characteristics. The efficiency of the procedure relies on the requirement of the (m, r)-blueness and our theorem on efficiently computing the connectivity characteristic of an SPG from the connectivity characteristic of two SPGs within adjacent regions.

[5] Generally, the SPGs within regions do not have to be k-vertex-connected (with respect to the included subset of S) as the "missing" connectivity can be added from the complementary SPG outside a given region.

We abstractly model the SPGs within regions and their fragments as subgraphs of a graph G on an extension of S by a distinguished set of point vertices of even degree called *portal vertices.* For a subgraph H of G, let $P(H)$ be the set of portal vertices in H that have positive degree in H equal to half of that in G. A *portal completion of* H is an augmentation of H by the edges in an f-matching of the complete multigraph (including self-loops) on $P(H)$, where each edge has multiplicity $\mathcal{O}(r \log\log W)$ and $f(v)$ is the degree of v in $P(H)$. Throughout this section, internally vertex disjoint paths will mean vertex disjoint, with the exception of the endpoints and portal vertices, paths. Now, similarly as in [5], we can assume the following definition.

Definition 11. *For each pair of distinct vertices v, u of H, let $Com_H(v,u)$ be the set of portal completions which augment H to a graph with k internally vertex-disjoint paths between v and u. For each vertex v of H, let $Path_H(v)$ be the set of pairs (U,Q) where U is a multiset on $P(H)$ with vertex multiplicities bounded from above by the vertex degrees in H and Q is a partial g-matching on the complete multigraph on $P(H)$ with $g(w)$ equal to the difference between the degree of w in H and the multiplicity of w in U such that there are $|U|+|Q|$ internally vertex-disjoint paths in H connecting v with each copy of a vertex in U, and connecting each pair of endpoints of each edge copy in Q, respectively. $Path_H$ is the set of all Q for which $(\emptyset, Q)$ is in $Path_H(v)$ for an arbitrary vertex v in $P(H)$. We call the set of values of $Com_H(v,u)$, $Path_H(u)$ and $Path_H$ where u and v range over the set of non-portal vertices of H, the* connectivity characteristic *of H, and denote it by $Char(H)$.*

Let F be the family of all subgraphs H of G such that: (i) after augmenting H with the edges of the complete multigraph on $P(H)$, for any pair of non-portal vertices there are k internally vertex-disjoint (with the exception of portals) paths connecting this pair in the augmented graph; (ii) 4 vertices in $P(H)$ have degree $\mathcal{O}(r \log\log W)$ and the total degree of the remaining vertices in $P(H)$ is $\mathcal{O}(r \log c)$. Note that the members of F model SPGs within regions that can be extended to a k-vertex-connected (m,r)-blue SPG on S.

For $H \in F$, let $M(H)$ denote the number of f-matchings in the complete multigraph on $P(H)$. It follows that there are at most $2^{M(H)}$ different values of $Com_H(v,u)$, at most $2^{2M(H)}$ different values of $Path_H(v)$, and at most $2^{M(H)}$ different values of $Path_H$ for $v,\ u\ \in P(H)$. Note that $M(H) < (\log\log W)^{10} \cdot (\mathcal{O}(r \log c))!$. Similarly as in [5], we obtain:

Lemma 3.1. *Let H, H' be subgraphs in F with disjoint non-empty sets of non-portal vertices. The connectivity characteristic of the graph $H \cup H'$ can be determined on the basis of those for H and H' in time $2^{(\log\log W)^{10} \cdot (\mathcal{O}(r \log c))!}$.*

We shall need also the following lemma.

Lemma 3.2. *For any portal completion B of a subgraph H in F, the connectivity characteristic of $H \cup B$ can be determined on the basis of that for H in time $\mathcal{O}(M(H))$.*

Proof. To compute $Comp_{H\cup B}(v,u)$, where u, u are non-portal vertices in H, for each completion D in $Comp_H(v,u)$, we insert D and the set difference $D \setminus B$ into $Comp_{H\cup C}(v,u)$. To compute $Path_H(u)$, where u is a non-portal vertex of H, for each (U,P) in $Path_H(u)$, and each $P' \subset B \setminus U \times U$, we insert P' into $Path_H(u)$. Analogously, to determine $Path_{H\cup B}$, for each P in $Path_{H\cup B}$ and each $P' \cup B$, we insert $P \cup P'$ into $Path_{H\cup B}$.

Our dynamic programming procedure will generate solely subgraphs without direct edges between outer portals. Lemma 3.1 combined with Lemma 3.2 yields the following key theorem for our dynamic programming procedure.

Theorem 3.1. *Let H, H' be subgraphs in F with disjoint non-empty sets of non-portal vertices, and let B, B' be their portal completions. The connectivity characteristic of the graph $(H \cup B) \cup (H' \cup B')$ restricted to edges with at least one endpoint outside $P(H) \cup P(H')$ can be determined on the basis of those for H and H' in time $2^{(\log\log W)^{10} \cdot (\mathcal{O}(r \log c))!}$.*

Corollary 3.1. *Let Q be a non-leaf region of the shifted quadtree. For all multisets P of portals satisfying the requirements from the definition of an (m,r)-blue SPG within a region, and all possible connectivity characteristics C for P, we can compute minimum-costs of (m,r)-blue SPGs within Q using P and having the characteristics C, on the basis of the minimum-costs of (m,r)-blue SPGs within the child regions of Q (for all possible portal multisets and connectivity characteristics) in total time $m^{\mathcal{O}(r \log c)} \cdot 2^{(\log\log W)^{10} \cdot (\mathcal{O}(r \log c))!}$.*

The quadtree for the input point multiset S has $\mathcal{O}(n \log n)$ regions, and it can be constructed in time $n(\log n)^{\mathcal{O}(1)}$. For each leaf region of the tree, each possible multiset of its portals satisfying the requirements from the definition of an (m,r)-blue SPG within a region, each possible connectivity characteristic, the minimum-cost of a (m,r)-blue SPG within the region (containing only a one point in S) can be easily computed in time $m^{\mathcal{O}(r \log c)} 2^{(\log\log W)^{10} \cdot (\mathcal{O}(r))!}$. These facts combined with Corollary 3.1 yield the following theorem by straightforward calculations.

Theorem 3.2. *Let $n = |S|$, $m = \mathcal{O}(c \log W)$ and $r = \mathcal{O}(c^2 k^4)$. A minimum-cost k-vertex-connected SPG on a well rounded S which is (m,r)-blue with respect to the shifted dissection can be computed in time $n (\log n)^{\mathcal{O}(1)}\, 2^{(\log\log W)^{10} \cdot (\mathcal{O}(c^2 k^4 \log c))!}$.*

4 PTAS for Euclidean k-connectivity

We combine our results presented in two previous sections to obtain a polynomial-time approximation scheme for the minimum-cost k-vertex-connectivity problem in $\mathbb{R}^2$.

Let S be a set of n points in $\mathbb{R}^2$. We first apply the Perturbation Theorem to reduce our problem to the one on a well-rounded set S' of at most n points in $\mathbb{R}^2$. This transformation ensures that a $(1+\Omega(\frac{1}{c}))$-approximation for S' in time T can be used to obtain an $(1+\frac{1}{c})$-approximation for S in time $\mathcal{O}(T+n)$. Then we pick at random an integer $a \in \{0,1,\ldots,W-1\}$, where W is the size of the bounding box of S'. Now, the Structure Theorem ensures that with probability at least $\frac{1}{2}$, there is an SPG that is k-vertex-connected with respect to S', is $\mathcal{O}(c^2\,k^4)$-gray with respect to the a-shifted dissection, and has the cost within $(1+\Omega(\frac{1}{c}))$ of the minimum-cost of a k-connected SG on S'. If there is such an SPG, then by Theorem 3.2, we can find a not worse k-vertex-connected and $(\mathcal{O}(c\log W), \mathcal{O}(c^2\,k^4))$-blue SPG on S' in time $n\,(\log n)^{\mathcal{O}(1)}\,2^{(\log\log W)^{10}\cdot(\mathcal{O}(c^2\,k^4\,\log c))!}$. Then we replace each path corresponding to the edges of the SPG by straight-line segments in order to obtain not worse (by triangle inequality) approximation in the form of a straight-line graph. We can also derandomize our PTAS by trying all possible a in the range $\{0,\ldots,W-1\}$ which increases the running time by a factor $\mathcal{O}(n)$. Putting everything together, we obtain our main result.

THEOREM 4.1. *There is polynomial-time approximation scheme for the minimum-cost k-vertex-connectivity problem in $\mathbb{R}^2$. For a set of n-points in $\mathbb{R}^2$, and $c>1$, the randomized version of the scheme finds a k-connected straight-line graph on S which is within $(1+\frac{1}{c})$ of the minimum in time $n\,2^{(\log\log n)^{10}\cdot(\mathcal{O}(c^2 k^4\,\log c))!}$. The algorithm can be derandomized by increasing the running time by a factor $\mathcal{O}(n)$.*

Our PTAS from Theorem 4.1 can be easily extended to higher dimensions. Also, it can be modified to include the minimum-cost k-edge-connectivity problem.

THEOREM 4.2. *There is polynomial-time approximation scheme for the minimum-cost k-vertex (or, k-edge) connectivity problem in $\mathbb{R}^d$. For a set of n-points in $\mathbb{R}^d$, and $c>1$, the randomized version of the scheme finds a k-edge (or, k-edge, respectively) connected straight-line graph on S which is within $(1+\frac{1}{c})$ of the minimum in time $n\,2^{(\log\log n)^{\binom{2d}{2}+2d}\cdot((\mathcal{O}(cdk^2))^d\,\log c)!}$. The algorithm can be derandomized by increasing the running time by a factor $\mathcal{O}(n)$.*

5 Hardness of approximations

In this section we prove that it is NP-hard to find a PTAS for the problem of finding the cost of minimum-cost biconnected graph spanning a set of n points in metric ℓ_p in $\mathbb{R}^{\log n}$ for any constant $p \geq 1$. We show our hardness result in two steps. First we prove that the approximation problem is NP-hard for complete undirected graphs whose edges have weights in $\{1,2\}$ and such that each vertex is incident to a constant number of edges with cost 1. For that we apply a reduction from TSP in such graphs, the problem known to be Max-SNP-hard [12]. Then we map weighted graphs into geometric graphs in ℓ_p metric in $\mathbb{R}^{\log n}$ such that a PTAS for the biconnectivity problem in geometric graphs would imply a PTAS for that problem in graphs whose edges have weights in $\{1,2\}$.

We apply also similar arguments to provide the first proof of non-existence of a PTAS (unless P = NP) for the problem of finding a k-vertex-connected spanning subgraph with the minimum number of edges of an unweighted graph. Finally, we extend our claims to the edge-connectivity problems.

5.1 Biconnectivity in 1-2-Δ graphs

DEFINITION 12. *A weighted undirected complete graph G is a* 1-2-Δ graph *if each of its edges has weight either 1 or 2 and each vertex is incident to at most Δ edges of weight 1.*

The following lemma plays the crucial role in our inapproximability results.

LEMMA 5.1. *Let $G=(V,E)$ be a 1-2-Δ graph on n vertices for some $\Delta, 0 \leq \Delta \leq n$. Let $H=(V,E^*)$ be a biconnected spanning subgraph of G of weight $w(H)$. Then one can construct in a polynomial time a Hamiltonian cycle in G of weight upper bounded by $2\cdot w(H)-n$.*

Proof. Let $E_i = \{e \in E^* : w(e)=i\}$ and $H_i=(V,E_i)$ for $i \in \{1,2\}$. Let $q(H_1)$ be the number of connected components of H_1. Let $\mathcal{F}$ be an arbitrary maximum-size spanning forest of H_1. Let L be the set of vertices of degree 1 in $\mathcal{F}$ (i.e., *leaves*) and let I be the set of isolated vertices in $\mathcal{F}$.

Observe that since H is biconnected, each connected component of H_1 must be incident in H to at least 2 edges of weight 2. Therefore $w(H) \geq |E^*| + q(H_1)$. On the other hand, since H is biconnected, the degree of each vertex in H is at least 2, and therefore $|E^*| \geq n - q(H_1) + \lceil |L|/2 \rceil + |I|$. Hence we get $w(H) \geq |E^*| + q(H_1) \geq n + \lceil |L|/2 \rceil + |I|$.

We construct a Hamiltonian cycle HC in G in two steps. We first set a Hamiltonian path in each connected component in H_1 and then connect these paths into one cycle in G.

Let C be a connected component in H_1 with t vertices and let $T=\mathcal{F}\cap C$. Let s be the number of leaves in T. We construct P, a Hamiltonian path in C, by connecting the vertices in C according to their preorder traversal in T. Note that if v is not a leaf in T then its leftmost child in T is its successor in P. Thus the edge in P from v to its ancestor in T is of weight 1. Therefore P contains at least $t-s$ edges of weight 1 and the other $s-1$ edges are of weight at most 2 each. Hence $w(P) \le t - s + 2(s-1) = t - 2 + s$. We perform the construction above for every connected component of H_1 containing more than one vertex. Then we can easily connect the obtained paths into a Hamiltonian cycle HC using $q(H_1)$ edges. Therefore, if for a connected component C, $L(C)$ denotes the number of leaves in $C \cap \mathcal{F}$, then HC satisfies the following

$$\begin{aligned} w(\mathsf{HC}) &\le 2q(H_1) + \\ &\quad \left(\sum_{\substack{\text{a non-trivial connected} \\ \text{component } C \text{ of } H_1}} |C| - 2 + L(C) \right) \\ &= n - |I| - 2(q(H_1) - |I|) + |L| + 2q(H_1) \\ &= n + |I| + |L| \\ &\le n + |I| + \lceil |L|/2 \rceil + \lceil |L|/2 \rceil \\ &\le w(H) + \lceil |L|/2 \rceil \\ &\le w(H) + (w(H) - n - |I|) \\ &\le 2 \cdot w(H) - n \ . \end{aligned}$$

THEOREM 5.1. *There exist constants $\Delta_0 > 0$ and $\varepsilon > 0$ such that, given an 1-2-Δ_0 graph G on n vertices, and given the promise that either its minimum-weight biconnected spanning subgraph H is of weight n, or $w(H) \ge (1+\varepsilon)n$, it is* NP*-hard to distinguish which of the two cases holds. In particular, it is* NP*-hard to approximate within $(1+\varepsilon)$ the cost of a minimum weight biconnected spanning subgraph of an 1-2-Δ_0 graph.*

Proof. Since the second claim follows trivially from the first one, we prove only the first claim. We show that there is a polynomial time reduction from that problem to the problem of finding an approximate solution for TSP in 1-2-Δ graphs, which is known to be NP-hard.

From the results of Papadimitriou and Yannakakis [12] and Arora et al. [2], we know that there exist positive constants Δ_0 and ε such that, given an 1-2-Δ_0 graph G on n vertices, and given the promise that either its minimum-weight traveling tour MIN-TST is of weight n, or $w(\text{MIN-TST}) \ge (1+2\varepsilon)n$, it is NP-hard to distinguish which of the two cases holds.

Let G be an 1-2-Δ_0 graph on n vertices for which we know that either its minimum-weight traveling tour MIN-TST is of weight n, or $w(\text{MIN-TST}) \ge (1+2\varepsilon)n$. Then Lemma 5.1 implies that the minimum-cost biconnected spanning subgraph H of G either is of weight of n or $w(H) \ge (1+\varepsilon)n$. Therefore, by the results of Papadimitriou and Yannakakis [12] and of Arora et al. [2] the problem of distinguishing whether $w(H) = n$ or $w(H) \ge (1+\varepsilon)n$ is NP-hard.

5.2 Inapproximability in Euclidean graphs

THEOREM 5.2. *For any fixed $p \ge 1$ there exists a constant $\epsilon > 0$ such that it is* NP*-hard to approximate within $1+\epsilon$ the minimum-cost biconnected graph spanning a set of n points in the ℓ_p metric in $\mathbb{R}^{\log n}$.*

Proof. We proceed along the lines of the proof of Theorem 9 in [14]. Fix p and let Δ_0 and ε be positive constants from Lemma 5.1. Let G be an arbitrary 1-2-Δ_0 graph for which we have the promise that either its minimum-cost biconnected spanning subgraph H is of weight n or $w(H) \ge (1+\varepsilon)n$. We show that distinguishing whether $w(H) = n$ or $w(H) \ge (1+\varepsilon)n$ can be reduced in a polynomial time to the problem of approximating within $1+\epsilon_p$ the minimum-cost biconnected graph spanning a set of n points in metric ℓ_p in $\mathbb{R}^{\log n}$.

Let $\epsilon = \frac{\varepsilon \cdot \delta}{4 + 2\delta + \varepsilon\delta}$, where δ will be set momentarily. As observed by Trevisan (cf. the proof of Theorem 9 in [14]) it is enough to consider ℓ_p metric in $\mathbb{R}^{\mathcal{O}(\log n/\epsilon^2)}$. Let $d(\cdot,\cdot)$ be the edge-cost in G and let $d_p(\cdot,\cdot)$ denote the ℓ_p distance between pairs of points in the Euclidean space $\mathbb{R}^{\mathcal{O}(\log n/\epsilon^2)}$. Trevisan (Corollary 8 in [14]) showed that one can find in a polynomial time a positive constant δ (depending only on Δ_0 and p) and an embedding Φ of G into $\mathbb{R}^{\mathcal{O}(\log n/\epsilon^2)}$ such that for any two vertices v, u in G

- if $d(v,u) = 1$ then $1 - \epsilon \le d_p(\Phi(v), \Phi(u)) \le 1$, and
- if $d(v,u) = 2$ then $1 + \delta - \epsilon \le d_p(\Phi(v), \Phi(u)) \le 1 + \delta$.

Let us apply this mapping to G and let S be the resulted set of n points in $\mathbb{R}^{\mathcal{O}(\log n/\epsilon^2)}$. Let H_S be a biconnected graph spanning S of the minimum cost. From our definition of S it is easy to see that if $w(H) = n$ then $w(H_S) \le n$.

Now suppose that $w(H) \ge (1+\varepsilon)n$. Let $H_S = \Phi(H^*)$ and let α and β be the numbers of edges in H^* of weight 1 and 2, respectively. Clearly $\alpha + \beta \ge n$. Since $w(H^*) \ge w(H) \ge (1+\varepsilon)n$, we also obtain $\alpha + 2\beta \ge (1+\varepsilon)n$. Now note that the definition of Φ implies that $w(H_S) \ge \alpha(1-\epsilon) + \beta(1+\delta-\epsilon) = (\alpha+\beta)(1-\epsilon) + \beta\delta$.

We consider two cases, depending on whether $\beta \ge \frac{2\epsilon}{\delta}n$ or not. If $\beta \ge \frac{2\epsilon}{\delta}n$, then $w(H_S) \ge (\alpha+\beta)(1-\epsilon) + \beta\delta \ge n(1-\epsilon) + \frac{2\epsilon}{\delta}n \cdot \delta = (1+\epsilon)n$. If $\beta < \frac{2\epsilon}{\delta}n$, then $w(H_S) \ge (\alpha+\beta)(1-\epsilon) + \beta\delta \ge \alpha(1-\epsilon) \ge ((1+\varepsilon)n - 2\beta)(1-\epsilon) > ((1+\varepsilon)n - \frac{4\epsilon}{\delta}n)(1-\epsilon) \ge (1+\epsilon)n$. We conclude that in order to achieve an approximation better than $(1+\epsilon)$ for the minimum-cost biconnected graph spanning a set of n points in metric ℓ_p in $\mathbb{R}^{\mathcal{O}(\log n/\epsilon^2)}$, it is sufficient to distinguish whether $w(H) = n$ or $w(H) \ge (1+\varepsilon)n$. This yields the theorem.

5.3 Inapproximability in unweighted graphs

One can further extend our approach and show that it is NP-hard to provide a PTAS for the problem of finding a *minimum* (i.e., having the smallest number of edges) biconnected spanning subgraph of a biconnected unweighted and undirected graph.

THEOREM 5.3. *There exist constants $\varepsilon > 0$ and δ such that it is* NP*-hard to approximate within $1+\varepsilon$ the minimum biconnected spanning subgraph of a biconnected unweighted undirected graph of maximum degree bounded by δ. In particular, the minimum k-vertex-connected spanning subgraph problem in graphs of the maximum degree bounded by some constant does not have a* PTAS *unless $P = NP$.*

Proof. For any graph G let $c(G)$ denote the number of edges in G. Let G be an 1-2-Δ graph on n vertices for some Δ, for which we know that either its minimum-cost salesman tour MIN-TST is of weight n, or $w(\text{MIN-TST}) \geq (1+3\varepsilon)n$. By the results of Papadimitriou and Yannakakis [12] and Arora et al. [2], we know that (for suitable chosen constants $\varepsilon > 0$ and Δ) it is NP-hard to distinguish which of the two cases holds. We show now that if we could approximate in a polynomial time the minimum cost of a biconnected spanning subgraph within $(1+\varepsilon)$, then we could distinguish in P which of the above cases holds.

Let G_1 be the spanning subgraph of G induced by the edges of weight 1 in G. Clearly, if G_1 is not biconnected then $w(\text{MIN-TST}) > n$, and hence, by our assumption, $w(\text{MIN-TST}) \geq (1+3\varepsilon)n$. Thus let us consider only the case that G_1 is biconnected. Let H_1 be any biconnected spanning subgraph of G_1. Let $\mathcal{T}$ be any spanning tree of H_1 and let $\mathcal{T}$ have s leaves (i.e., vertices of degree 1). Then, since every vertex must be of degree at least 2 in H_1, we obtain $c(H_1) \geq n-1+\lceil \frac{1}{2}s \rceil$. Given $\mathcal{T}$, we can construct a salesman tour TST in G analogous as in the proof of Lemma 5.1. Basic calculations show that $w(\text{TST}) \leq (n-s)+2s = 2(n-1+\frac{s}{2})-(n-2) \leq 2 \cdot c(H_1) - n + 2$. Therefore, by our discussion above, any biconnected spanning subgraph H_1 of G_1 with the minimum number of edges either consists of n edges or $c(H_1) \geq (1+\frac{3}{2}\varepsilon)n - 1 \geq (1+\varepsilon)n$. Hence, the results of Papadimitriou and Yannakakis [12] and of Arora et al. [2] imply that the problem of distinguishing whether $c(H_1) = n$ or $w(H_1) \geq (1+\varepsilon)n$ is NP-hard.

Since in our construction the degree of G_1 is bounded by Δ and the result of Papadimitriou and Yannakakis [12] holds for 1-2-Δ_0 graphs for some constant Δ_0, our claim above is valid even for graphs with the maximum degree bounded by the constant $\delta = \Delta_0$.

5.4 Inapproximability for k-edge connectivity

One can modify all the preceding proofs to obtain similar inapproximability results for the problem of finding a minimum-cost k-edge-connected subgraph of a k-vertex-connected graph. (A theorem similar to Theorem 5.6 has been already proven by Fernandes [6], but we believe that our proof is more intuitive and significantly simpler.)

THEOREM 5.4. *There exist constants $\Delta_0 > 0$ and $\varepsilon > 0$ such that, given an 1-2-Δ_0 graph G on n vertices, and given the promise that either its minimum-weight 2-edge-connected spanning subgraph H is of weight n, or $w(H) \geq (1+\varepsilon)n$, it is NP-hard to distinguish which of the two cases holds. In particular, it is NP-hard to approximate within $(1+\varepsilon)$ the cost of a minimum weight 2-edge-connected spanning subgraph of an 1-2-Δ_0 graph.*

THEOREM 5.5. *For any fixed $p \geq 1$ there exists a constant $\epsilon > 0$ such that it is NP-hard to approximate within $1+\epsilon$ the minimum 2-edge-connected spanning graph spanning a set of n points in the ℓ_p metric in $\mathbb{R}^{\log n}$.*

THEOREM 5.6. *There exist constants $\varepsilon > 0$ and δ such that it is NP-hard to approximate within $1+\varepsilon$ the minimum 2-edge-connected spanning subgraph of a 2-edge-connected unweighted undirected graph of maximum degree bounded by δ. In particular, the minimum k-edge-connected spanning subgraph problem in graphs of the maximum degree bounded by some constant does not have a PTAS unless P = NP.*

Acknowledgments We thank Luca Trevisan for his explanations of the inapproximability results for TSP in 1-2-Δ-graphs and in Euclidean graphs. We thank also the SODA'99 PC-chairs and SIAM for the possibility of late correction of the final version.

References

[1] S. Arora. Polynomial time approximation schemes for Euclidean TSP and other geometric problems. To appear in *J. Assoc. Comput. Mach.*. A preliminary versions of this paper appeared in Proc. 37th FOCS, pp. 2–11, 1996, and in Proc. 38th FOCS, pp. 554–563, 1997.

[2] S. Arora, C. Lund, R. Motwani, M. Sudan, and M. Szegedy. Proof verification and hardness of approximation problems. In Proc. 33rd FOCS, pp. 14–23, 1992.

[3] S. Arora, P. Raghavan, and S. Rao. Approximation schemes for Euclidean k-medians and related problems. In Proc. 30th STOC, pp. 106–113, 1998.

[4] P. Crescenzi and V. Kann. A compendium of $\mathcal{NP}$ optimization problems. April 21, 1997. Available at http://www.nada.kth.se/theory/problemist.html.

[5] A. Czumaj and A. Lingas. A polynomial time approximation scheme for Euclidean minimum cost k-connectivity. In Proc. 25th ICALP, pp. 682–694, 1998.

[6] C. G. Fernandes. A better approximation ratio for the minimum size k-edge-connected spanning subgraph problem. *J. Algorithms*, 28(1):105–124, July 1998.

[7] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-completeness.* Freeman, 1979.

[8] D. S. Hochbaum, ed. *Approximation Algorithms for $\mathcal{NP}$-Hard Problems.* PWS Publishing Company, 1996.

[9] S. Khuller. Approximation algorithms for finding highly connected subgraphs. In D. S. Hochbaum, ed, *Approximation Algorithms for $\mathcal{NP}$-Hard Problems*, chapter 6. PWS Publishing Company, 1996.

[10] D. Krznaric, C. Levcopoulos, G. Narasimhan, B. J. Nilsson, and M. Smid. The minimum 2-connected subgraph of a set of points. Manuscript, Sept. 15, 1997.

[11] E. L. Lawler, J. K. Lenstra, A. H. G. Rinnooy Kan, and D. B. Shmoys, editors. *The Traveling Salesman Problem.* John Wiley & Sons, 1985.

[12] C. H. Papadimitriou and M. Yannakakis. The Traveling Salesman problem with distances one and two. *Math. Oper. Res.*, 18(1):1–11, 1993.

[13] S. B. Rao and W. D. Smith. Approximating geometrical graphs via "spanners" and "banyans"". In Proc. 30th STOC, pp. 540–550, 1998.

[14] L. Trevisan. When Hamming meets Euclid: the approximability of geometric TSP and MST. In Proc. 29th STOC, pp. 21–29, 1997.