CrossMark

Approximability of the minimum-weight k-size cycle cover problem

Michael Khachay[1] · **Katherine Neznakhina[1]**

Abstract The cycle cover problem is a combinatorial optimization problem, which is to find a minimum cost cover of a given weighted digraph by a family of vertex-disjoint cycles. We consider a special case of this problem, where, for a fixed number k, all feasible cycle covers are restricted to be of the size k. We call this case the minimum weight k-size cycle cover problem (Min-k-SCCP). Since each cycle in a cover can be treated as a tour of some vehicle visiting an appropriate set of clients, the problem in question is closely related to the vehicle routing problem. Moreover, the studied problem is a natural generalization of the well-known traveling salesman problem (TSP), since the Min-1-SCCP is equivalent to the TSP. We show that, for any fixed $k > 1$, the Min-k-SCCP is strongly NP-hard in the general setting. The Metric and Euclidean special cases of the problem are intractable as well. Also, we prove that the Metric Min-k-SCCP belongs to APX class and has a 2-approximation polynomial-time algorithm, even if k is not fixed. For the Euclidean Min-2-SCCP in the plane, we present a polynomial-time approximation scheme extending the famous result obtained by S. Arora for the Euclidean TSP. Actually, for any fixed $c > 1$, the scheme finds a $(1 + 1/c)$-approximate solution of the Euclidean Min-2-SCCP in $O(n^3(\log n)^{O(c)})$ time.

Keywords Cycle cover problem (CCP) · Traveling salesman problem (TSP) · *NP*-hard problem · Polynomial-time approximation scheme (PTAS)

1 Introduction

The cycle cover problem (CCP) is a combinatorial optimization problem, which is to find an optimal cover of a given graph by a set of vertex-disjoint cycles. To the best of our knowledge, for the first time, this problem was introduced in the seminal paper by Sahni and Gonzales [22]. Since that time, the CCP and its various modifications were extensively

✉ Michael Khachay
   mkhachay@imm.uran.ru

[1] Krasovsky Institute of Mathematics and Mechanics, Ural Federal University, Ekaterinburg, Russia

studied in numerous publications (see, e.g. [4–7,19,20,23]). Since each cycle in a cover can be considered as a tour of some vehicle visiting an appropriate set of clients, the CCP is closely related to the vehicle routing problem (VRP). Moreover, the studied problem is a natural generalization of the well-known traveling salesman problem (TSP), since the Min-1-SCCP is equivalent to the TSP. Below, we propose a brief overview of the related problems and previous work.

The traveling salesman problem (TSP) [13] is a classic combinatorial optimization problem, which deals with finding the minimum-cost salesman tour (Hamiltonian cycle) in a given complete weighted graph. There are several special cases of the problem, and some of them, such as the Metric TSP and the Euclidean TSP, are of particular interest. An instance of the Metric TSP is defined by an undirected weighted graph so that edge weights satisfy the triangle inequality. Furthermore, in Euclidean TSP the nodes of the given graph are points in d-dimensional space (for some $d > 1$), and edge weights are Euclidean distances between the incident nodes. It is known [21] that the TSP is NP-hard even in the Euclidean case; i.e., the optimal solution can not be found in polynomial time unless $P = NP$. Although the TSP is hardly approximable [22] in the general setting, there are polynomial-time approximation algorithms for some special cases. For instance, the Metric TSP [8] can be approximated in polynomial time with a ratio of 3/2, and, for the Euclidean TSP, a polynomial-time approximation scheme [1] and a randomized asymptotically correct algorithm [14] are developed.

The vehicle routing problem (VRP) [10] deals with servicing a number of clients (customers) with a fleet of vehicles. In the simplest case (which is also known as the Multiple Traveling Salesmen Problem or the mTSP [3]), an instance of the VRP is defined by n client locations and one dedicated location (depot). The goal is to find a minimum-cost set of vehicle routes visiting every client only once so that any route starts and finishes at the depot. Surveys of the recent results concerning polynomial-time approximation algorithms and heuristics for several modifications of this problem can be found in [15,18,24].

The m-peripatetic salesmen problem (m-PSP) [11,17] is related to searching for several edge-disjoint salesmen routes optimizing a given objective function (e.g. the total weight of routes, maximum weight, etc.) As for the TSP, so the m-PSP is intractable and hardly approximable in the general case, but, for the Euclidean case, polynomial-time asymptotically correct algorithms are known [2,14].

Cycle covers of graphs are spanning subgraphs consisting of vertex-disjoint simple cycles and maybe isolated vertices. In [5,7], it is shown that cycle covers provide an efficient tool for approximating the TSP, the VRP, and their modifications. Among others, L-cycle covers are mostly discovered [4–7,20]. For a given subset $L \subset \mathbb{N}$, a cycle cover is called L-cycle cover if every containing cycle has the number of edges, which belongs to L. While computing L-covers of the minimum (or maximum) weight is NP-hard [19], approximation results for multiple restricted cases of the problem are known. For instance, the Min-k-UCC(1, 2) problem where $L = \{k, k+1, \ldots\}$ and edge weights restricted to 1 and 2, can be approximated polynomially within a factor 7/6 for any k [6]. Further, in [20] it is shown that the Metric Min-L-UCC problem, for any fixed L, is polynomial-time approximable within a constant factor (depending on L).

We introduce the minimum-weight k-size cycle cover problem (Min-k-SCCP), which is closely related to the Traveling Salesman, the Vehicle Routing, and the minimum L-cycle cover problems. In contrast to the Min-L-UCC problem, we restrict not the length of cycles covering the graph but the number of cycles (size of the cover) itself. Actually, for a fixed natural number k and a given complete weighted digraph (with loops) $G = (V, E, w)$, it is required to find a minimum-weight cover of the set V by k vertex-disjoint cycles.

We show that the Min-k-SCCP is strongly NP-hard and hardly approximable in general case and remains intractable in the metric and Euclidean special cases. For the Metric Min-k-SCCP, we propose a 2-approximation polynomial-time algorithm, whose approximation ratio and running time do not depend on k. For the Euclidean Min-2-SCCP in the plane, we present a polynomial-time approximation scheme, extending the result of Arora [1] obtained for the Euclidean TSP.

2 Problem statement

We consider the following combinatorial optimization problem. For a fixed natural number k and a given complete weighted digraph (with loops) $G = (V, E, w)$, it is required to find a minimum-weight cover of the set V by k vertex-disjoint cycles.

Let C be an arbitrary directed cycle in the digraph G. We denote sets of its nodes and arcs by $V(C)$ and $E(C)$, respectively. The arcs of G are weighted by some weighting (cost) function $w : E \to \mathbb{R}$. Since the set E is finite, the weight function w is defined by the matrix

$$W = (w_{ij}), \quad (1 \leqslant i, j \leqslant n)$$

uniquely, so $w(e) = w_{ij}$ for any arc $e = (i, j)$. The weight (cost) of a cycle C is defined by the equation $W(C) = \sum_{e \in E(C)} w(e)$.

Definition 1 Let $C_1, \ldots, C_k$ be vertex-disjoint simple (with no repeated vertices) directed cycles in the graph G such that $V(C_1) \cup \cdots \cup V(C_k) = V$. The family $\mathcal{C} = \{C_1, \ldots, C_k\}$ is called a k-Size Cycle Cover (k-SCC) of the graph G. The weight $\mathcal{W}(\mathcal{C})$ of the cover $\mathcal{C}$ is defined by equality $\mathcal{W}(\mathcal{C}) = \sum_{i=1}^{k} W(C_i)$.

The Minimum-weight k-size cycle cover problem (Min-k-SCCP) For a given complete weighted digraph $G = (V, E, w)$ with loops, it is required to find a cycle cover of the graph G consisting of k cycles and having the minimum possible weight.

The Min-k-SCCP can be stated in the optimization form

$$\mathcal{W}(\mathcal{C}) = \min \sum_{i=1}^{k} \sum_{e \in E(C_i)} w(e)$$
$$\text{s.t.}$$
$$\bigcup_{i=1}^{k} V(C_i) = V,$$
$$V(C_i) \cap V(C_j) = \varnothing, \quad (\{i, j\} \subset \{1, \ldots, k\} = \mathbb{N}_k).$$

Similarly to the Metric and Euclidean TSP, we consider special cases of the minimum-weight k-size cycle cover problem, namely, the Metric Min-k-SCCP and the Euclidean Min-k-SCCP.

In the Metric Min-k-SCCP, a weight function w satisfies the following constraints for any $1 \leq i, j, l \leq n$:

  (i) $w_{ij} \geq 0$,
  (ii) $w_{ij} = w_{ji}$,
  (iii) $w_{ii} = 0$,
  (iv) $w_{il} + w_{lj} \geq w_{ij}$.

Further, we denote by $\bar{G}$ the undirected graph induced by G. It should be noted that the weight of an arbitrary cycle cover does not depend on the orientation of the constituent

cycles. Therefore, in the following sections, we carry out all the reasonings related to the graph G in terms of cycle covers for the corresponding undirected graph $\bar{G}$.

The Euclidean Min-k-SCCP is just a subclass of the Metric Min-k-SCCP, in which nodes of the given graph G are points in d-dimensional space (for some $d > 1$), and edge weights are Euclidean distances between the incident nodes.

3 Computational complexity of the Min-k-SCCP

This section contains the intractability results for both the general case of the Min-k-SCCP and the the Metric and Euclidean special cases of the problem in question.

Theorem 1 *The Min-k-SCCP, the Metric Min-k-SCCP and Euclidean Min-k-SCCP (in d-dimensional space for some $d > 1$) are strongly NP-hard for any fixed $k \geqslant 1$.*

Proof Obviously, Theorem 1 is valid for $k = 1$, since the Min-1-SCCP is equivalent to the TSP, which is strongly NP-hard.

For an arbitrary $k > 1$, the main idea of the proof is as follows. We reduce the TSP to the Min-k-SCCP by cloning the given TSP instance k times and spreading the obtained clones apart. For the constructed instance of the Min-k-SCCP, we show that any its optimal solution consists of copies of optimal solutions (minimum-weight Hamiltonian cycles) of the initial TSP instance.

Indeed, let the TSP instance be given by a complete weighted undirected graph $G = (V, E, w)$, $|V| = n$, and

$$q = \sum_{1 \leqslant i < j \leqslant n} w_{ij} + 1.$$

Consider the Min-k-SCCP instance defined by the complete weighted undirected graph $G' = (V', E', w')$, which satisfies the following conditions

$$V' = V_1 \cup \cdots \cup V_k, \quad V_l = \{n(l-1)+1, n(l-1)+2, \ldots, nl\}, \quad (1 \leqslant l \leqslant k).$$

For an arbitrary $x \in V_{k_1}$, $y \in V_{k_2}$ and $i \equiv x(\mod n)$, $j \equiv y(\mod n)$, we define the weight $w'(e)$ of the edge $e = \{x, y\}$ by the equation

$$w'(e) = w_{ij} + q \operatorname{sign} |k_1 - k_2|.$$

Hence, if $k_1 = k_2$, $w'(e) = w_{ij}$; otherwise, $w'(e) > w_{ij}$. The reduction can be made in polynomial time.

Let $\mathcal{C}^* = \{C_1^*, \ldots, C_k^*\}$ be an optimal solution of the Min-k-SCCP (i.e., the minimum-weight k-size cycle cover of the graph G). Each number $l \in \mathbb{N}_k$ corresponds to a number $k(l) \in \mathbb{N}_k$ so that $V_{C_i^*} = V_{k(l)}$. Since the criterion of the Min-k-SCCP is additive, cycle C_l is the minimum-weight Hamiltonian cycle of the subgraph $G'(V_{k(l)})$. Each subgraph of this kind is isomorphic to the graph G according to the construction of the graph G'. So an arbitrary cycle C_l corresponds to an optimal solution of the initial TSP.

In order to reduce the Metric TSP to the Metric Min-k-SCCP or the Euclidean TSP to the Euclidean Min-k-SCCP in polynomial time, one can easily modify the argument above. Therefore, the metric and Euclidean special cases of the Min-k-SCCP are intractable as well. □

Adapting the technique of [22] to the case of k-size cycle covers, it is easy to show that the Min-k-SCCP has no polynomial approximation algorithms with any ratio of $O(2^n)$. In

subsequent sections, we show that the Metric Min-k-SCCP and the Euclidean Min-k-SCCP can be approximated with much higher accuracy.

4 Minimum spanning forest

We need the following standard definitions.

Definition 2 An acyclic graph consisting of k connected components is called a k-forest.

Definition 3 Any spanning subgraph F of the graph G, being a k-forest, is called a spanning k-forest of the graph G. A spanning k-forest having the minimum weight is called a k-minimum spanning forest (k-MSF) of the graph G.

Using a modification of Kruskal's algorithm [9], we construct a k-minimum spanning forest of the graph G.

Algorithm 1 Modification of Kruskal's algorithm

1: Start from the n-forest F consisting of isolated nodes;
2: For each $i \in \mathbb{N}_{n-k}$ the edge

$$e_i = \arg\min\{w(e) \colon F \cup \{e\} \text{ is a forest}\}$$

  add to F;
3: The output is the constructed k-forest $F^* = F$.

Assertion 1 *Algorithm 1 is correct, and the resulting k-forest F^* is a k-minimum spanning forest (of the graph G)*

Proof To prove that Algorithm 1 is valid, we need to show that the forest F^* is k-MSF. Evidently, F^* is a k-forest by construction. To complete the proof, it is necessary to show that the weight $W(F^*)$ takes the minimum value among weights of k-forests.

Assume the contrary. Then, let $(e_1, \ldots, e_{n-k})$ be a sequence of the edges of F^* ordered according to their addition to the forest by Algorithm 1. For any k-MSF F, we denote by $h(F)$ the largest index such that $e_1, \ldots, e_h \in F$. By construction,

$$w(e_{h+1}) \leq w(f) \quad (f \in F \backslash F^*). \tag{1}$$

Suppose

$$F' = \arg\max\{h(F) \colon F \text{ is a } k\text{-MSF}\}. \tag{2}$$

By our assumption, $h(F') < n - k$. Consider the graph $F' \cup \{e_{h+1}\}$. There are two options. Either this graph is acyclic or it has one simple cycle. We prove the second option, the first one can be treated by analogy.

Let C be a simple cycle in the graph $F' \cup \{e_{h+1}\}$. Since F^* is a forest, C has an edge $f \in F' \backslash F^*$. Consider the graph $\tilde{F} = F' \cup \{e_{h+1}\} \backslash \{f\}$. By construction, $\tilde{F}$ is a k-forest, and

$$W(\tilde{F}) = W(F') + w(e_{h+1}) - w(f) \leq W(F'),$$

by Eq. (1). Therefore, $\tilde{F}$ is a k-MSF. Further, $h(\tilde{F}) > h(F')$ that violates the optimality of F', see (2). Consequently, our initial assumption that F^* is not a k-MSF is not valid. The contradiction obtained completes the proof. □

5 2-approximation algorithm for the Metric Min-k-SCCP

Definition 4 An r-approximation algorithm for an optimization problem is a polynomial-time algorithm producing, for any instance I of the problem, a solution of value APP, which is at most r times the optimum value OPT. The parameter r is called an *approximation guarantee* and it is defined by the formula

$$r = \sup_{I} \frac{APP(I)}{OPT(I)},$$

where I is an instance of the optimization problem.

For the Metric Min-k-SCCP, we propose algorithm, which extends the 2-approximation algorithm for the Metric TSP based on the preliminary construction of minimum spanning tree for the given graph. W.o.l.g., we assume that $n > k$.

Algorithm 2 Polynomial-time 2-approximation algorithm

1: For the given graph G, construct a k-minimum spanning forest F applying algorithm 1;
2: Take all edges in F twice to transform all nonempty trees of F into Eulerian subgraphs;
3: For any obtained Eulerian subgraph, find a Eulerian cycle; then, transform all of them into Hamiltonian cycles (using the standard procedure);
4: Output the constructed set of cycles augmented by the necessary number of loops.

Assertion 2 *Algorithm 2 has a running time of $O(n^2 \log n)$ and an approximation guarantee r satisfying the inequality*

$$2(1 - 2/n) \leq r \leq 2(1 - 1/n). \tag{3}$$

Proof The running time of Algorithm 1 is upper-bounded by the running time of Kruskal's algorithm, which is $O(|E| \log |E|)$ [9]. In Algorithm 2, the running time of steps 2 and 3 is $O(|E|)$. Hence, the running time of Algorithm 2 is upper-bounded by $O(n^2 \log n)$ for any given complete graph G.

To prove the inequality in the right-hand side of (3), we consider an arbitrary minimum-weight k-size cycle cover $\mathcal{C}$ of the graph G. Since $k < n$, $\mathcal{C}$ contains at least one nonempty cycle[1]. We transform $\mathcal{C}$ into a spanning k-forest F by removing the most heavy edge from any nonempty cycle. In what follows, we denote the weights of F and k-minimum spanning forest by SF and by MSF, respectively. Then,

$$MSF \leqslant SF \leqslant OPT(1 - 1/n),$$

consequently,

$$APP \leqslant 2MSF \leqslant 2(1 - 1/n)OPT$$

and

$$r \leq 2(1 - 1/n).$$

To prove the inequality in the left-hand side of (3), we consider the following instance of the Min-2-SCCP (Fig. 1). Let G be a complete weighted graph without loops and $n = 4p + 2$

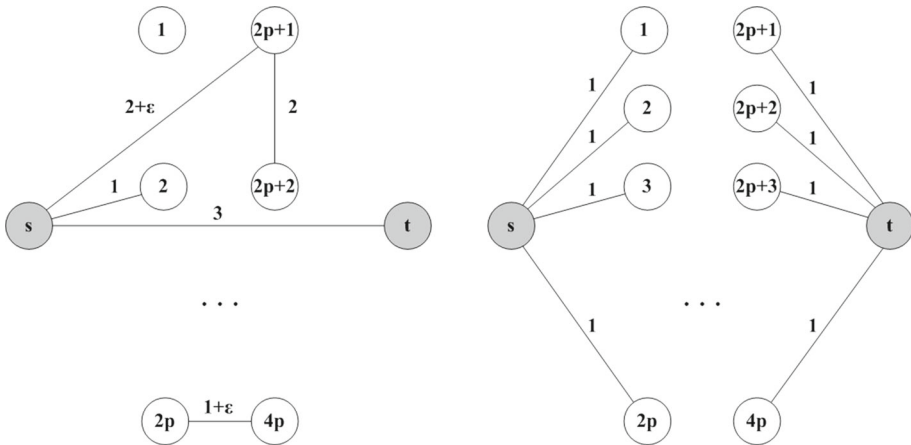[1] We call a cycle C empty if C is a loop.

Fig. 1 The instance of the Min-k-SCCP and 2-minimum spanning forest
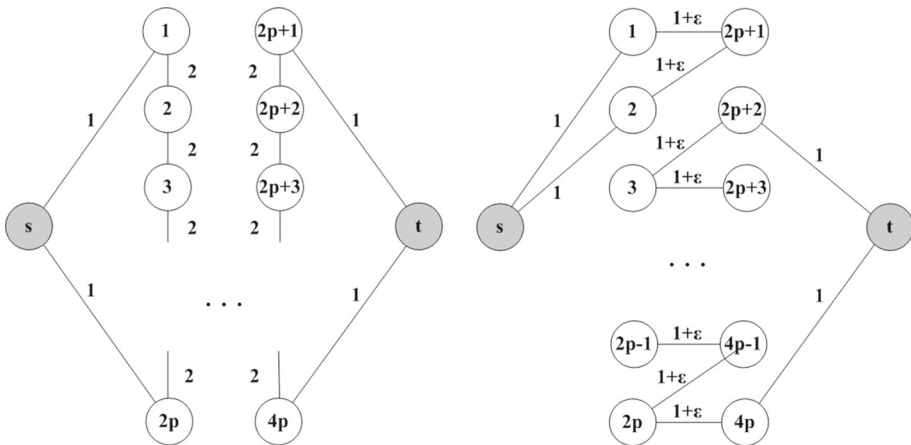

Fig. 2 2-Approximation constructed by Algorithm 2 and the better approximation of the Min-2-SCCP

(for some $p > 1$). Among n vertices of the graph, there are two distinguished vertices s and t. For any $1 \le i, l \le 2p$ and $2p + 1 \le j, m \le 4p$ such that $i \neq l$ and $j \neq m$, the weighting (cost) function w is defined by the equations:

(i) $w_{si} = w_{tj} = 1$,
(ii) $w_{sj} = w_{ti} = 2 + \epsilon$,
(iii) $w_{ij} = 1 + \epsilon$,
(iv) $w_{il} = w_{jm} = 2$,
(v) $w_{st} = 3$,

where $\epsilon \in (0, 1)$.

For this instance, the weight of any 2-minimum spanning forest F equals $4p$ (Fig. 1); therefore, $APP = 8p$ (Fig. 2). It is easy to verify that there are better ways to approximate the given instance. One of such approximate solutions of weight $4 + 2(1 + \epsilon)(2p - 1)$ is presented in Fig. 2. Therefore,

$$OPT \leq 2 + 4p + 2\epsilon(2p - 1).$$

Hence,

$$r \geq \sup_{\epsilon \in (0,1)} \frac{8p}{2 + 4p + 2\epsilon(2p - 1)} = \frac{4p}{1 + 2p},$$

and, finally,

$$r \geq 2(1 - 2/n).$$

$\square$

It should be noted that approximation ratio bounds (3) and the running time of Algorithm 2 are independent of k. Therefore, Algorithm 2 can be considered as a 2-approximation algorithm for the case of the Min-k-SCCP for which parameter k is given as a part of the input.

6 Polynomial-time approximation scheme

It is generally believed that a combinatorial optimization problem has a polynomial-time approximation scheme (PTAS) if, for any fixed $c > 1$, there exists an algorithm finding a $(1 + 1/c)$-approximate solution of the problem in time bounded by some polynomial of the instance length. Generally speaking, the order and coefficients of this polynomial can depend on c.

Further, we present a polynomial-time approximation scheme for the Euclidean Min-2-SCCP in the plane.

6.1 Preliminary preprocessing of the instance considered

We start from the claim that, for any instance of the Euclidean Min-2-SCCP in the plane, one of the following alternatives is valid; their verification can be conducted in polynomial time:

(i) the instance in question is decomposable into a pair of independent smaller instances of the Euclidean TSP;
(ii) the maximum inter-node distance (diameter) of the instance has an upper bound depending on OPT linearly.

Our considerations are based on the well-known geometric Jung's inequality [16] establishing the dependence between the diameter D of a bounded set in $\mathbb{R}^d$ and the radius R of its circumscribed sphere:

$$\frac{1}{2}D \leqslant R \leqslant \left(\frac{d}{2d + 2}\right)^{\frac{1}{2}} D.$$

Consequently, in the plane we have:

$$\frac{1}{2}D \leqslant R \leqslant \frac{\sqrt{3}}{3}D. \tag{4}$$

For the instance given by the graph $\bar{G}$ (Fig. 3), we construct a 2-minimum spanning forest $F = \{T_1, T_2\}$. We denote the weight of the forest F by MSF; the diameters of the trees T_1 and T_2 by D_1 and D_2. Also, we denote the radii of the circumscribed circles and the distance between their centers by R_1, R_2, and $\rho(T_1, T_2)$, respectively. Let $D = \max\{D_1, D_2\}$ and $R = \max\{R_1, R_2\}$.

Fig. 3 Minimum spanning forest $\{T_1, T_2\}$ and circumscribed circles
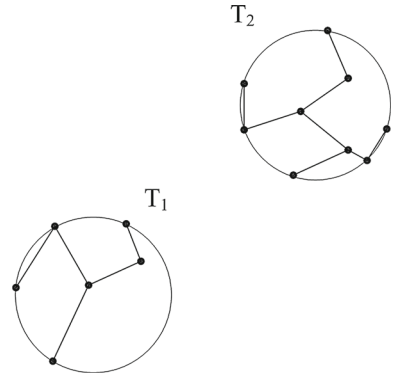
Assertion 3 *For the given graph $\bar{G}$, suppose that the lower bound*

$$\rho(T_1, T_2) > 5R \tag{5}$$

is valid. Then, an arbitrary minimum-weight 2-size cycle cover of the graph consists of the minimum-weight Hamiltonian cycles for the subgraphs $G(T_1)$ and $G(T_2)$ induced by the trees T_1 and T_2. Otherwise, the internode distances of the given instance can be upper-bounded by

$$\frac{7\sqrt{3}}{3}MSF \leqslant \frac{7\sqrt{3}}{3}OPT. \tag{6}$$

Proof Let inequality (5) be valid for the given instance. Suppose $\mathcal{C} = \{C_1, C_2\}$ is a minimum-weight 2-size cycle cover of the graph G. By contradiction, assume that some cycle contains nodes from the both trees T_1 and T_2. W.l.o.g., we suppose that the equations

$$C_1 \cap T_1 \neq \varnothing, \quad C_1 \cap T_2 \neq \varnothing, \quad C_2 \cap T_2 \neq \varnothing, \quad \text{and } C_2 \cap T_1 = \varnothing$$

are valid[2] (Fig. 4).

By assumption, the cycle C_1 contains at least two edges e_1 and e_2 spanning T_1 and T_2. By (5), the weights $w(e_1)$ and $w(e_2)$ are greater than $3R$, simultaneously. Let us remove these edges and close the cycles inside the circumscribed circles (for T_1 and T_2) (Fig. 4) To make such a transformation, we should add three new edges (and remove one). Summarizing, the total weight of the removed edges is greater than $6R$ and the total weight of the added edges is at most $6R$. Therefore, we construct a lighter 2-size cycle cover than $\mathcal{C}$, which contradicts the optimality of $\mathcal{C}$.

Suppose the given instance violates (5). Then, the distance between any two nodes in the graph G is at most $7R$ due to the triangle inequality. Applying the right-hand side of (4) and taking into account the straightforward bound $D \leqslant MSF \leqslant OPT$, we have

$$R \leqslant \frac{\sqrt{3}}{3}D \leqslant \frac{\sqrt{3}}{3}OPT.$$

Therefore, the maximum internode distance of the graph G is upper-bounded by (6). □

As it follows from Assertion 3, for any instance of the Euclidean Min-2-SCCP satisfying (5), a PTAS can be composed of PTASes for two smaller instances of the Euclidean TSP defined by the subgraphs $G(T_1)$ and $G(T_2)$. Hereafter, we consider the special case of the Euclidean Min-2-SCCP, which violates (5).

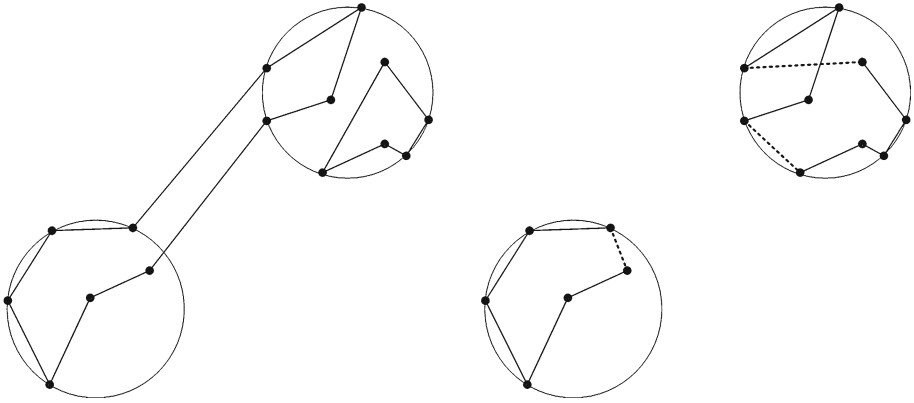[2] Other cases can be considered similarly.

Fig. 4 Cycle C_1 has a nonempty intersection with both T_1 and T_2 and transformation of cycles

6.2 Main idea of the algorithm

The main idea of the proposed approximation scheme, similarly to Arora's PTAS [1], is based on randomized recursive partitioning of the bounding box of the given instance into smaller squares and successive searching for the pair of closed tours satisfying the following conditions:

   (i) any node of the given graph is visited by some route and only once;
  (ii) segments of the routes are continuous piece-wise linear curves and can cross the borders of all squares only in predefined points called *portals*;
 (iii) locations of the portals and the maximum count of crossings for each border-line of the squares depend on parameter c;
 (iv) the required pair of routes has the minimum weight among other pairs satisfying conditions (i)–(iii).

Our proof of the polynomial-time approximation scheme for the Euclidean Min-2-SCCP problem in the plane consists of the following stages.

1. First, in Sect. 6.3, we prove that, for any instance of the Euclidean Min-2-SCCP in the plane and for any value of $c > 1$, an instance of a special kind (we call such an instance *well-rounded*) can be assigned in such a way that any $(1 + 1/c)$-approximate solution of the well-rounded instance induces an $(1 + c_1/c)$-approximate solution of the initial instance (for some independent constant $c_1 > 1$); all these transformations can be carried out in polynomial time.
2. Then, we describe (in Sect. 6.4) a randomized recursive partitioning procedure for the axis-aligned square that contains the well-rounded instance of the Min-2-SCCP (we call this square a *bounding box*).
3. Further, we prove (in Sect. 6.5) that, for the well-rounded instance, there exists an $(1 + 1/c)$-approximate 2-size cycle cover satisfying the above conditions (i)–(iv) with a probability at least 1/2.
4. Finally, using the dynamic programming approach and standard derandomization procedure, we construct (in Sect. 6.6) the approximate solution mentioned above.

6.3 Well-rounded problem

To approximate well the Euclidean Min-2-SCCP in the plane, it is sufficient to design an efficient approximation algorithm for the special case of this problem, which is called a well-rounded Euclidean Min-2-SCCP. In particular, any PTAS for the well-rounded Euclidean Min-2-SCCP induces the corresponding PTAS for the general case with the same running-time bound.

Definition 5 An instance of the Euclidean Min-2-SCCP in the plane is called *well-rounded*, if the following conditions are valid:

 (i) any node i of the input graph $\bar{G}$ has integral coordinates $x_i, y_i \in \mathbb{N}^0_{O(n)}$;
(ii) for each edge $e \in E$, $w(e) \geqslant 4$.

Lemma 1 *Any PTAS for the well-rounded Euclidean Min-2-SCCP induces the appropriate PTAS for the Euclidean Min-2-SCCP.*

Proof Suppose some value $c > 1$ is fixed. To any instance of the Euclidean Min-2-SCCP in the plane, we assign the appropriate well-rounded instance.

Let us take an arbitrary instance of the Min-2-SCCP [violating condition (5)]. We denote the weights of a minimum-weight 2-size cycle cover and a 2-minimum spanning forest by OPT and MSF, respectively. According to Assertion 3, there is an axis-aligned bounding box for this instance with side-length

$$L = \frac{7\sqrt{3}}{3}MSF \leqslant \frac{7\sqrt{3}}{3}OPT.$$

We place an axis-aligned grid of granularity $L/(2nc)$ and move any node to the nearest grid-point. It should be noted that it is possible to map several nodes into the same grid-point, and, therefore, to have a well-rounded instance with a smaller number of nodes. After this transformation, every internode distance is changed at most by $L/(nc)$. Also, the weight of any 2-size cycle cover can be changed at most by L/c.

To obtain the grid of granularity 4, we scale the instance multiplying all the coordinates by $8nc/L$. We shift the origin to the left-bottom corner of the scaled bounding box. As a result, the side-length of the bounding box becomes equal to $O(nc) = O(n)$ for any fixed c.

By construction, the instance obtained is well-rounded. Further, consider an arbitrary 2-size cycle cover $\mathcal{C}$ in the initial instance and the corresponding cover $\mathcal{C}'$ in the well-rounded one. We denote the weights of $\mathcal{C}$ and $\mathcal{C}'$ by W and W', respectively. Then, the bound

$$\frac{8nc}{L}\left(W - \frac{L}{c}\right) \leqslant W' \leqslant \frac{8nc}{L}\left(W + \frac{L}{c}\right)$$

is valid.

Hereafter, *OPT* and *OPT'* stand for the weights of optimal solutions (minimum-weight 2-size cycle covers) of the appropriate instances. Let $\mathcal{C}'$ be the approximate solution obtained by PTAS for the well-rounded instance, i.e.

$$OPT' \leqslant W' \leqslant \left(1 + \frac{1}{c}\right)OPT'.$$

Taking into account that

$$OPT' \leqslant \frac{8nc}{L}\left(OPT + \frac{L}{c}\right),$$

we have

$$\frac{8nc}{L}\left(W - \frac{L}{c}\right) \leqslant W' \leqslant \left(1 + \frac{1}{c}\right) OPT' \leqslant \left(1 + \frac{1}{c}\right)\frac{8nc}{L}\left(OPT + \frac{L}{c}\right);$$

therefore,

$$W - \frac{L}{c} \leqslant \left(1 + \frac{1}{c}\right)\left(OPT + \frac{L}{c}\right)$$

and, finally,

$$OPT \leqslant W \leqslant \left(\frac{7\sqrt{3}}{3c^2} + \frac{17\sqrt{3}}{3c} + 1\right) OPT.$$

$\square$

6.4 Quadtree

We construct a geometric partition of the well-rounded Min-2-SCCP following the approach introduced in [1].

Definition 6 For an instance of the Euclidean Min-2-SCCP given by a graph $\bar{G}$, a bounding box $\mathcal{S}$ is the smallest axis-aligned square containing $\bar{G}$. The side-length L of $\mathcal{S}$ is some power of two.

We construct a dissection of $\mathcal{S}$ into smaller squares using vertical and horizontal lines. These lines are crossing the coordinate axes in integer-coordinate points with a step of length 1. By construction, every smallest-size square contains at most one node of the given instance.

Further, we proceed with using of a 4-regular tree of a special kind known as *a quadtree* [12]. In our case, the root of the tree is the bounding box $\mathcal{S}$. Each non-leaf square in the tree is partitioned into four equal child squares. This recursive partitioning stops on a square containing at most one node (Fig. 5).

We define the levels of the squares of the constructed quadtree as follows. The bounding box $\mathcal{S}$ is the unique square of the first level, its four children belong to the second level, and so on. By construction, the quadtree contains $O(n)$ leaves, $O(\log L) = O(\log n)$ levels and thus $O(n \log n)$ squares in all.

We refer to the point $(L/2, L/2)$ as the centre point. This is the point of crossing of the inner edges of the first level squares. We consider a quadtree with a centre point picked at random in $\mathcal{S}$.

Definition 7 Let $a, b \in \mathbb{N}_L$ be constants. The quadtree with the centre point $((L/2 + a)$ mod $L, (L/2 + b)$ mod $L)$, is called a *shifted* quadtree $T(a, b)$.

For any level greater than 1, the corresponding squares of $T(a, b)$ are considered *modulo* L (Fig. 5). These squares are called *wrapped-around*.

In the next section, we show that, if the stochastic variables a and b defining the quadtree $T(a, b)$ are distributed uniformly in $\mathbb{N}_L$, then there exists a $(1 + 1/c)$-approximation of the Euclidean Min-2-SCCP in the plane with a probability at least $1/2$.

6.5 Structure theorem

To some parameter values $m, r \in \mathbb{N}$, and any square S [a node in the quadtree $T(a, b)$], we assign a regular partition of the border ∂S consisting of $4(m + 1)$ points including all the corners of S (Fig. 6). We call this partition *m-regular*; its points are referred to *portals*.
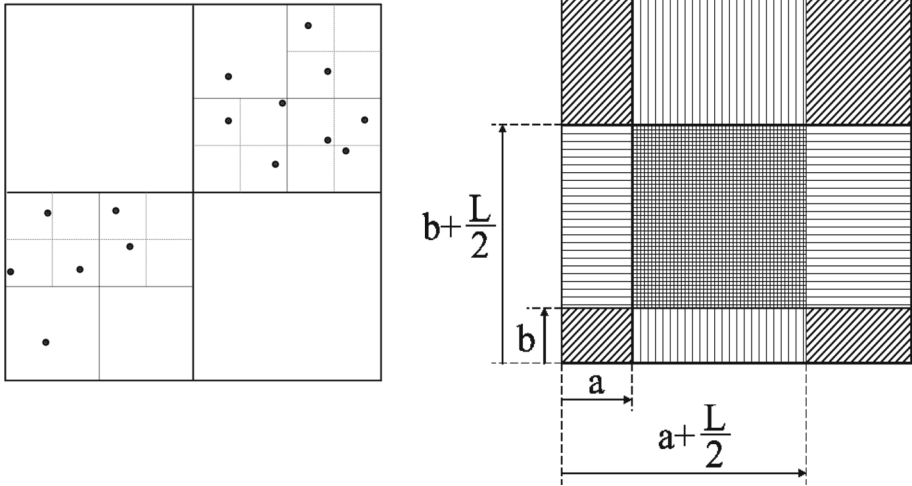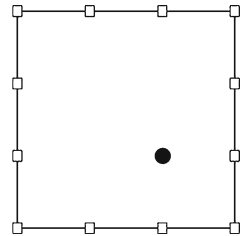
Fig. 5 The quadtree and the shifted quadtree

Fig. 6 Example of the 2-regular partition

Definition 8 The union of m-regular partitions of borders for all nodes of the quadtree $T(a, b)$ (excluding the root) is called the m-regular portal set and is denoted by $P(a, b, m)$.

Definition 9 Let C be an arbitrary simple cycle in the graph $\bar{G}$ in the plane. The closed continuous piecewise linear route $l(C)$ is called an (m, r)-approximation of the cycle C if

(i) $l(C)$ bends only at points of $V(C) \cup P(a, b, m)$;
(ii) the nodes of $V(C)$ are visited by $l(C)$ in the same order as by C;
(iii) for any side of any node of $T(a, b)$, the route $l(C)$ crosses this side at points of $P(a, b, m)$ and at most r times.

We use a result of Arora [1], which for our purposes can be phrased as follows.

Theorem 2 (Structure Theorem) *Let an instance of the well-rounded TSP in the plane be given by the graph G, let L be the side-length of the bounding box S, and let constants $c > 1$ and $\eta \in (0, 1)$ be fixed. If the stochastic variables a and b are distributed uniformly in $\mathbb{N}_L$ and the parameters m and r are defined by the formulas*

$$m = \lceil 2s \log L \rceil, \quad r = s + 4, \quad and \quad s = \lceil 36c/\eta \rceil,$$

then, for an arbitrary simple cycle C of weight $W(C)$, there exists an (m, r)-approximation $l(C)$ of weight $W(l(C)) \leqslant (1 + 1/c)W(C)$ with a probability at least $1 - \eta$.

Following the seminal Arora's idea to approximate Hamiltonian tours by portal-respecting (m, r)-approximations, we define a similar construction for our problem.
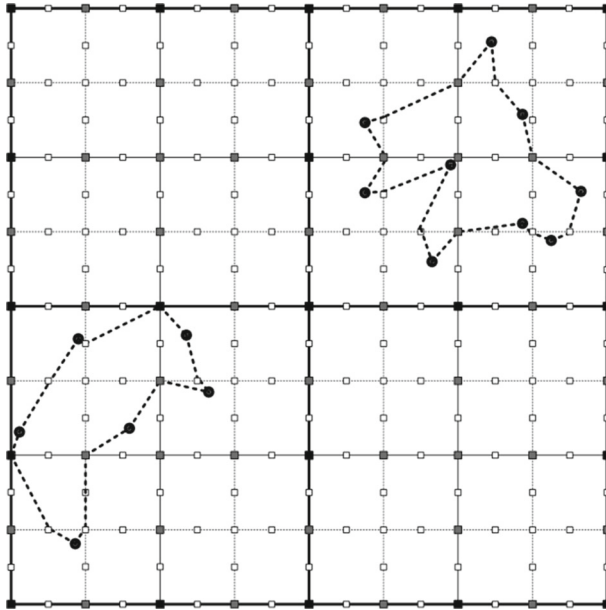
Fig. 7 Example of a cycle $(m, r, 2)$-cover

Definition 10 Let $\mathcal{C} = \{C_1, \ldots, C_k\}$ be an arbitrary k-size cycle cover in the graph $\bar{G}$, and let $l(C_i)$ be some (m, r)-approximation of the cycle C_i. Then, the family $\mathcal{L}(\mathcal{C}) = \{l(C_1), \ldots, l(C_k)\}$ is called a cycle (m, r, k)-cover in the graph $\bar{G}$.

Obviously, an arbitrary cycle $(m, r, 1)$-cover consists of the only bent Hamiltonian cycle. We are interested in $(m, r, 2)$-covers (Fig. 7).

Theorem 3 *Let a constant $c > 1$ be fixed, let L be the side-length of the bounding box $\mathcal{S}$ for the given instance of the well-rounded Min-2-SCCP in the plane. If the discrete stochastic variables a and b are distributed uniformly in $\mathbb{N}_L$ and*

$$m = O(c \log L), \ r = O(c),$$

then, for the graph $\bar{G}$, there exists a cycle $(m, r, 2)$-cover of weight at most $(1 + 1/c)OPT$ with a probability at least $1/2$.

Proof Let $\mathcal{C}^* = \{C_1^*, C_2^*\}$ be an optimal solution of the given instance of the well-rounded Min-2-SCCP, for which $W(C_1^*) + W(C_2^*) = OPT$. By Theorem 2, for $\eta = 1/4$, an arbitrary number $i \in \{1, 2\}$ and parameter values

$$m = \lceil 2s \log L \rceil, \ r = s + 4, \ \text{and} \ s = \lceil 144c \rceil,$$

there exists an (m, r)-approximation $l(C_i^*)$ of weight

$$W(l(C_i^*)) \leqslant (1 + 1/c)W(C_i^*) \tag{7}$$

with a probability at least $1 - \eta = 3/4$.

Then, due to the uniform distribution of variables a and b, (m, r)-approximations $l(C_1^*)$ and $l(C_2^*)$ satisfying inequality (7) exist simultaneously with a probability at least $1/2$. This implies that there exists a cycle $(2, m, 2r)$-cover $\mathcal{L}(\mathcal{C}) = \{l(C_1^*), l(C_2^*))\}$ such that

Fig. 8 Example of an
$(m, r, 2, S)$-segment, $\kappa = 2$.
Lines of different styles represent
the parts of l_1 and l_2
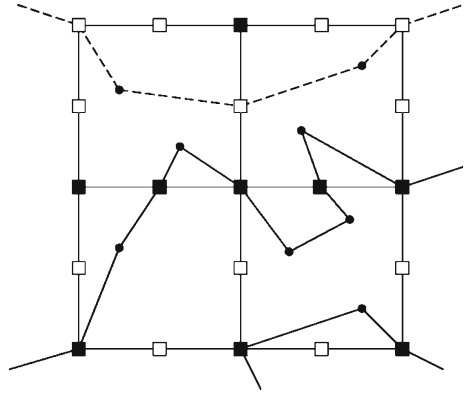
$$W(l(C_1^*)) + W(l(C_2^*)) \leqslant \left(1 + \frac{1}{c}\right)(W(C_1^*) + W(C_2^*)) = \left(1 + \frac{1}{c}\right) OPT$$

with a probability at least $1/2$.   □

6.6 Dynamic programming

For parameter values $m = O(c \log n)$ and $r = O(c)$, we describe the dynamic programming
procedure for searching a cycle $(m, r, 2)$-cover $\mathcal{L} = \{l_1, l_2\}$ of minimum weight, with time
complexity of $O(n(\log n)^{O(c)})$.

Definition 11 A part of a cycle $(m, r, 2)$-cover belonging to some square S (a node of the
quadtree $T(a, b)$) and visiting all the nodes of the given graph $\bar{G}$ located inside the square S
is called an $(m, r, 2, S)$-segment of this cover.

The example of an $(m, r, 2, S)$-segment is represented in Fig. 8.
To describe the Bellman equation, we define the inner task, which is solved recursively
for each entry of the dynamic programming lookup table.

Inner Task (S, R_1, R_2, κ).
Input The square S is a node of the quadtree $T(a, b)$. The cortege $R_i : \mathbb{N}_{q_i} \to (P(a, b, m) \cap \partial S)^2$ defines a sequence of ordered pairs (s_j^i, t_j^i) of portals; i.e., the crossing-points of the
(m, r)-approximation l_i and the border ∂S. The number κ defines the number of (m, r)-
approximations that intersect the square S. If $q_1 q_2 > 0$, then $\kappa = 2$; otherwise, $\kappa \in \{1, 2\}$.
Output A $(m, r, 2, S)$-segment defined by the values of the input parameters of the minimum
weight $W(S, R_1, R_2, \kappa)$.

Depending on the values of the input parameters, there are several cases of this Inner Task
listed below.

  (i) The case of $q_1 q_2 > 0$ seems to be regular. In this case, the resulting segment consists
       of parts of both (m, r)-approximations l_1 and l_2.
 (ii) The case of $q_1 \neq 0, q_2 = 0$ (or $q_1 = 0, q_2 \neq 0$) and $\kappa = 1$ is similar to the previous
       one, except that all subroutes in the square S belong to the only (m, r)-approximation
       (either l_1, or l_2).
(iii) In the case of $q_1 \neq 0, q_2 = 0$ (or $q_1 = 0, q_2 \neq 0$) and $\kappa = 2$, one of the generated
       (m, r)-approximations is supposed to belong entirely to the square S. In this case, the

resulting segment is constructed by augmenting the output of case (ii) by a closed (m, r)-approximation located inside the square S.

(iv) If $q_1 = q_2 = 0$ and $\kappa = 2$, then l_1 and l_2 are supposed to belong to the square S, and the output is a cycle $(m, r, 2)$-cover, which is located in the square S.

(v) Finally, if $q_1 = q_2 = 0$ and $\kappa = 1$, then it is necessary to find a (m, r)-approximation of the minimum weight, which belongs entirely to the square S. This case is the same as the TSP.

The Bellman equation Similarly to the Arora's PTAS for the Euclidean TSP, our dynamic programming procedure starts with leaves of the quadtree $T(a, b)$. Let S be an arbitrary leaf in the tree $T(a, b)$. By construction, S contains at most one node of the graph $\bar{G}$ and, for any input values of R_1, R_2, and the number κ, the Inner Task (S, R_1, R_2, κ) can be solved by brute force in time of $O(r)$.

Any other node (not a leaf) of the quadtree $T(a, b)$ has four child nodes; denote them by $S^I, \ldots, S^{IV}$. According to the recursive assumption, for any entry of the lookup table related to the square S, all possible instances of the Inner Task for squares S^I–S^{IV} should be solved. The appropriate entries should be also filled.

Let us explain the recursive solution of the inner task (S, R_1, R_2, κ) for fixed values of corteges R_1, R_2, and the parameter κ. Let $\mathfrak{P}$ be a family of multisets P consisting of at most $4r$ portals (with their multiplicities) located at the inner sides of the child squares $S^I, \ldots, S^{IV}$.

By the choice of m and r, any such side has $m + 2$ portals, at which the side can be crossed at most r times. Therefore, $|\mathfrak{P}| = O((m + 2)^{4r})$. To any multiset $P \in \mathfrak{P}$, we assign the set Σ_P of maps $\sigma : P \to \mathbb{N}_{q_1+q_2}$. To each inner portal $p \in P$, each map σ assigns the ordered pair

$$(s^1_{\sigma(p)}, t^1_{\sigma(p)}), \text{ if } \sigma(p) \leqslant q_1,$$
$$(s^2_{\sigma(p)-q_1}, t^2_{\sigma(p)-q_1}), \text{ otherwise} .$$

Consequently, the route-segment $l_1(s^1_{\sigma}(p), t^1_{\sigma}(p))$ or $l_2(s^2_{\sigma(p)-q_1}, t^2_{\sigma(p)-q_1})$ crossing one of sides of the child squares at the portal p. In this case, we call such a portal p *matched* to this route-segment. Since $|P| \leqslant 4r$ and $q_1 + q_2 \leqslant 4r$ the bound $|\Sigma_P| = O((2r)^{4r})$, is valid.

To any route-segment $l_i(s_{ij}, t_{ij})$, we assign the set $\Lambda_{ij}(\sigma)$ consisting of permutations α of the preimage multiset $\sigma^{-1}(j + (i - 1)q_1) \subset P$ of the map σ. It is easy to show that $|\Lambda_{ij}(\sigma)| = O((4r)!)$.

Each triple $\tau = (P, \sigma, \alpha)$ induces the instance quadruple of the inner tasks

$$\left((S^I, R_1^I(\tau), R_2^I(\tau), \kappa^I(\tau)), \ldots, (S^{IV}, R_1^{IV}(\tau), R_2^{IV}(\tau), \kappa^{IV}(\tau)) \right)$$

in such a way that

$$W(S, R_1, R_2, \kappa) = \min_{\tau} \sum_{i=I}^{IV} W(S^i, R_1^i(\tau), R_2^i(\tau), \kappa^i(\tau)).$$

The solving time-complexity for the task (S, R_1, R_2, κ) has the upper bound

$$O((m + 2)^{4r} (2r)^{4r} (4r)!).$$

Finally, to construct a minimum-weight cycle $(m, r, 2)$-cover, we should find a solution of the task $(S, R_1^0, R_2^0, 2)$, for empty corteges R_1^0 and R_2^0.

To estimate the total running time of the constructed dynamic programming procedure, we need an upper bound for the number of entries in the lookup table. It can be easily verified that any node S of the quadtree $T(a, b)$ is related to $O(m + 2)^{4r}$ ways to choose a multiset of

portals located at ∂S; any such multiset can be partitioned into pairs at most $O((4r)!)$ times, and any such partition can be assigned to routes l_1 and l_2 at most by $O(2^{2r})$ ways. Taking into account the total number $O(L \log L)$ of nodes (for the quadtree $T(a, b)$), we obtain an upper time complexity bound for the pair (a, b)

$$O\left(L \log L \times (m + 2)^{8r}((4r)!)^2 (2r)^{4r} \times 2^{2r}\right). \tag{8}$$

Derandomization The dynamic programming procedure described above finds an approximate solution of the well-rounded Min-2-SCCP for each pair a, b, which is related to the quadtree $T(a, b)$. We denote the weight of this solution by $APP(a, b)$. It follows from Theorem 3 that

$$P\left(APP(a, b) \leqslant (1 + \frac{1}{c})OPT\right) \geqslant 1/2$$

for the probability measure induced by the distribution of a and b. Therefore, there exists a pair $(a^*, b^*) \in \mathbb{N}_L$, for which the inequality

$$OPT \leqslant APP(a^*, b^*) \leqslant (1 + 1/c)OPT$$

is valid. Such a pair can be found by the exhaustive search in time $O(L^2)$. Taking into account that $m = O(c \log n)$, $r = O(c)$, and $L = O(n)$, we have proved Theorem 4.

Theorem 4 *The well-rounded Min-2-SCCP in the plane has a PTAS with time complexity of*

$$O(n^3 (\log n)^{O(c)}). \tag{9}$$

Using Lemma 1, we obtain our main result.

Theorem 5 *The Euclidean Min-2-SCCP in the plane has a PTAS with a time complexity of* (9).

Remark 1 It follows from (8) that the time complexity of the proposed PTAS for the Euclidean Min-2-SCCP coincides with the complexity of the PTAS from [1] for the Euclidean TSP up to the constant factor $2^{O(c)}$.

7 Conclusion

For any fixed $k \geqslant 1$, we proved the intractability of the Min-k-SCCP and its two special cases, the Metric Min-k-SCCP and the Euclidean Min-k-SCCP.

For the Metric Min-k-SCCP, the 2-approximation polynomial-time algorithm is proposed. For the Euclidean Min-2-SCCP in the plane, the polynomial-time approximation scheme is developed.

While the obtained results imply that the Metric Min-k-SCCP belongs to APX, and the Euclidean Min-2-SCCP belongs to PTAS complexity classes, several issues remain open. In particular, we are interested in constructing approximation algorithms for the Metric Min-k-SCCP with improved approximation ratios. Also, we hope that the presented PTAS can be extended to the Euclidean Min-k-SCCP for any $k > 2$ and any dimension $d > 2$.

References

1. Arora, S.: Polynomial-time approximation schemes for euclidean traveling salesman and other geometric problems. J. ACM **45**, 753–782 (1998)
2. Baburin, A., Della Croce, F., Gimadi, E.K., Glazkov, Y.V., Paschos, V.T.: Approximation algorithms for the 2-peripatetic salesman problem with edge weights 1 and 2. Discrete Appl. Math. **157**(9), 1988–1992 (2009)
3. Bektas, T.: The multiple traveling salesman problem: an overview of formulations and solution procedures. Omega **34**, 209–219 (2006). doi:10.1016/j.omega.2004.10.004
4. Bläser, M., Manthey, B.: Approximating maximum weight cycle covers in directed graphs with weights zero and one. Algorithmica **42**, 121–139 (2005). doi:10.1007/s00453-004-1131-0
5. Bläser, M., Manthey, B., Sgall, J.: An improved approximation algorithm for the asymmetric TSP with strengthened triangle inequality. J. Discrete Algorithms **4**, 623–632 (2006). doi:10.1016/j.jda.2005.07.004
6. Bläser, M., Siebert, B.: Computing cycle covers without short cycles. In: Algorithms ESA 2001, pp. 368–379. Springer, Berlin (2001)
7. Chandran, L.S., Ram, L.S.: On the relationship between ATSP and the cycle cover problem. Theor. Comput. Sci. **370**, 218–228 (2007). doi:10.1016/j.tcs.2006.10.026
8. Christofides, N.: Worst-case analysis of a new heuristic for the traveling salesman problem. In: Symposium on New Directions and Recent Results in Algorithms and Complexity, p. 441 (1975)
9. Cormen, T.H., Stein, C., Rivest, R.L., Leiserson, C.E.: Introduction to Algorithms, 2nd edn. McGraw-Hill Higher Education, New York (2001)
10. Dantzig, G.B., Ramser, J.H.: The truck dispatching problem. Manag. Sci. **6**(1), 80–91 (1959)
11. De Kort, J.: Lower bounds for symmetric k-peripatetic salesman problems. Optimization **20**, 113–122 (1991)
12. Finkel, R.A., Bentley, J.L.: Quad trees: a data structure for retrieval on composite keys. Acta Inf. **4**, 1–9 (1974). doi:10.1007/BF00288933
13. Garey, M.R., Johnson, D.S.: Computers and Intractability: A Guide to the Theory of NP-Completeness. W. H. Freeman & Co., New York (1979)
14. Gimadi, E.: Asymptotically optimal algorithm for finding one and two edge-disjoint traveling salesman routes of maximal weight in euclidean space. Proc. Steklov Inst. Math. **263**, S57–S67 (2008)
15. Golden, B., Raghavan, S., Wasil, E.A.: The Vehicle Routing Problem: Latest Advances and New Challenges. Operations Research/Computer Science Interfaces Series, vol. 43. Springer, Berlin (2008)
16. Jung, H.: Über die kleinste kugel, die eine räumliche figur einschliesst. J. Reine Angew. Math. **123**, 241–257 (1901)
17. Krarup, J.: The peripatetic salesman and some related unsolved problems. In: Roy, B. (ed.) Combinatorial Programming: Methods and Applications, pp. 173–178. Springer, Netherlands (1975)
18. Kumar, S., Panneerselvam, R.: A survey on the vehicle routing problem and its variants. Intell. Inf. Manag. **4**, 66–74 (2012). doi:10.4236/iim.2012.43010
19. Manthey, B.: On approximating restricted cycle covers. SIAM J. Comput. **38**, 181–206 (2008). doi:10.1137/060676003
20. Manthey, B.: Minimum-weight cycle covers and their approximability. Discrete Appl. Math. **157**, 1470–1480 (2009). doi:10.1016/j.dam.2008.10.005
21. Papadimitriou, C.: Euclidean TSP is NP-complete. Theor. Comput. Sci. **4**(3), 237–244 (1977)
22. Sahni, S., Gonzales, T.: P-complete approximation problems. J. ACM **23**, 555–565 (1976)
23. Szwarcfiter, J., Wilson, L.: The cycle cover problem. Tech. Rep. 131, University of Newcastle upon Tyne (1979)
24. Toth, P., Vigo, D. (eds.): The Vehicle Routing Problem. Society for Industrial and Applied Mathematics, Philadelphia (2001)