Colored Non-Crossing Euclidean Steiner Forest

Sergey Bereg[1], Krzysztof Fleszar[2], Philipp Kindermann[2], Sergey Pupyrev[3,4], Joachim Spoerhase[2], and Alexander Wolff[2]

[1] University of Texas, Dallas, USA
[2] Lehrstuhl für Informatik I, Universität Würzburg, Germany
[3] Department of Computer Science, University of Arizona, Tucson, Arizona, USA
[4] Institute of Mathematics and Computer Science, Ural Federal University, Russia.

Abstract. Given a set of k-colored points in the plane, we consider the problem of finding k trees such that each tree connects all points of one color class, no two trees cross, and the total edge length of the trees is minimized. For $k = 1$, this is the well-known Euclidean Steiner tree problem. For general k, a $k\rho$-approximation algorithm is known, where $\rho \le 1.21$ is the Steiner ratio.
We present a PTAS for $k = 2$, a $(5/3 + \varepsilon)$-approximation for $k = 3$, and two approximation algorithms for general k, with ratios $O(\sqrt{n} \log k)$ and $k + \varepsilon$.

1 Introduction

Steiner tree is a fundamental problem in combinatorial optimization. Given an edge-weighted graph and a set of vertices called *terminals*, the task is to find a minimum-weight subgraph that connects the terminals. For *Steiner forest*, the terminals are colored, and the desired subgraph must connect, for each color, the terminals of that color.

In this paper, we consider a geometric variant of Steiner forest where we add the constraint of planarity and require that terminals with distinct colors lie in distinct connected components. More precisely, we consider the problem of computing, for a k-colored set of points in the plane (which we also call *terminals*), k pairwise non-crossing planar Euclidean Steiner trees, one for each color. Such trees exist for every given set of points. We call the problem of minimizing the total length of these trees k-*Colored Non-Crossing Euclidean Steiner Forest* (k-CESF). Figure 1 shows some instances.

The problem is motivated by a method that Efrat et al. [7] suggested recently for visualizing embedded and clustered graphs. They visualize clusters by regions in the plane that enclose related graph vertices. Their method attempts to reduce visual clutter

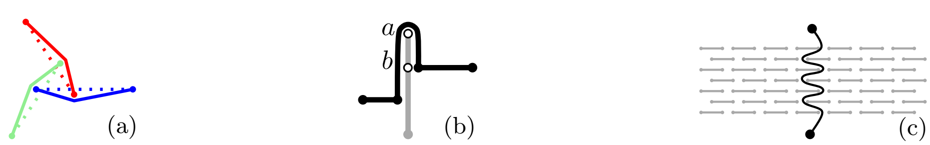

Fig. 1: Difficult examples for k-CESF. (a) The optimum contains no straight-line edge. (b) Segment ab is used twice by the black curve. (c) The black curve can be made arbitrarily longer than the corresponding straight-line segment. (Gray segments represent different colors.)

k	k-CESF	planar graph
1	EST: NP-hard [10], $1+\varepsilon$ [1,15]	ST: NP-hard [10], $1+\varepsilon$ [4]
2	$1+\varepsilon$ (Thm. 4)	
3	$5/3+\varepsilon$ (Thm. 5)	
general k	$k+\varepsilon$ (Thm. 1), $O(\sqrt{n}\log k)$ (Thm. 3)	k const.-size nets on 2 faces, exact [12]
$n/2$	NP-hard [2], $O(\sqrt{n}\log n)$ [5]	k size-2 nets on h faces, exact [9]

Table 1: Known and new results for k-CESF (hardness and approximation ratios)

and optimize "convexity" of the resulting regions by reducing the amount of "ink" necessary to connect all elements of a cluster. Efrat et al. [7] proposed the problem k-CESF and provided a simple $k\rho$-approximation algorithm, where ρ is the *Steiner ratio*, that is, the supremum, over all finite point sets in the plane, of the ratio of the total edge length of a minimum spanning tree over the total edge length of a Euclidean Steiner tree (EST). Chung and Graham [6] showed that $\rho \le 1.21$.

Our contribution. The middle column of Table 1 shows our results. For k-CESF, we present a deterministic $(k+\varepsilon)$- and a randomized $O(\sqrt{n}\log k)$-approximation algorithm; see Sect. 2. The main result of our paper is that 2-CESF admits a polynomial-time approximation scheme (PTAS); see Sect. 3. By a non-trivial modification of the PTAS, we prove that 3-CESF admits a $(5/3+\varepsilon)$-approximation algorithm; see Sect. 4.

Our PTAS for 2-CESF uses some ideas of Arora's algorithm [1] for EST, which is equivalent to 1-CESF. Since, in a solution to 2-CESF, the two trees are not allowed to cross, our approach differs from Arora's algorithm in several respects. We use a different notion of *r-lightness*, and by a *portal-crossing reduction* we achieve that each portal is crossed at most three times. More care is also needed in the perturbation step and in the base case of the dynamic program.

Related Work. Apart from the result of Efrat et al. [7], so far the only two variants of k-CESF that have been studied are those with extreme values of k. As mentioned above, 1-CESF is the same as EST, which is NP-hard [10]. Arora [1] and Mitchell [15] showed independently that EST admits a PTAS. The other extreme value of k, for which k-CESF has been considered, is $k = n/2$. This is the problem of joining specified pairs of points via non-crossing curves of minimum total length. Liebling et al. [13] gave some heuristics for this problem. Bastert and Fekete [2] claimed that $(n/2)$-CESF is NP-hard, but their proof has not been formally published. Recently, Chan et al. [5] considered $(n/2)$-CESF in the context of embedding planar graphs at fixed vertex locations. They gave an $O(\sqrt{n}\log n)$-approximation algorithm based on an idea of Liebling et al. [13] for computing a short non-crossing tour.

There is substantial work on the case where there are obstacles in the plane. Note that, in contrast to k-CESF, a valid solution may not exist in that setting. For a single color (that is, 1-CESF with obstacles), Müller-Hannemann and Tazari [17] give a PTAS. Papadopoulou [18] gave an algorithm for finding minimum-length non-crossing paths joining pairs of points (that is, $n/2$-CESF) on the boundary of a single polygon. A practical aspect of the problem—computing non-crossing paths of specified thickness—was studied by Polishchuk and Mitchell [19]. Their algorithm computes a representation of the thick paths inside a simple polygon; they also show how to find shortest thick disjoint paths joining endpoints on the boundaries of polygonal obstacles (with exponential dependence on the number of obstacles). The main difficulty with multiple obstacles is deciding which homotopy class of the paths gives the minimum length. If the homotopy class of the paths is specified, then the problem is significantly easier [8, 20]. Hurtado et al. [11] studied a set visualization problem where points can be blue or red or both, and the points of either color must be connected. Their aim was to minimize the total length of the network. The blue and red subgraphs may intersect.

The graph version of the problem has been studied in the context of VLSI design. Given an edge-weighted plane graph G and a family of k vertex sets (called nets), the goal is to find a set of k non-crossing Steiner trees interconnecting the nets such that the total weight is minimized. The problem is clearly NP-hard, as the special case $k = 1$ is the graph Steiner tree problem (ST), which is known to be NP-hard [10]. ST admits a PTAS [4]. On planar graphs, k-CESF can be solved in $O(2^{O(h^2)} n \log k)$ time [9] for k terminal pairs (that is, size-2 nets) if all terminals lie on h faces of the given n-vertex graph and in $O(n \log n)$ time for $h = 2$ and k constant-size nets [12]. We list these results in Table 1; many entries are still open.

In the *group Steiner tree* problem, one is given a k-colored point set and the task is to find a minimum-length tree that connects at least one point of each color. The problem is discussed in a survey by Mitchell [16]. Another related problem is that of constructing a minimum-length non-crossing path through a given sequence of points in the plane. Its complexity status remains open [14].

2 Algorithms for k-CESF

Despite its simple formulation, the k-CESF problem seems to be rather difficult. There are instances where the optimum contains no straight-line edges or contains paths with repeated line segments; see Figs. 1(a) and 1(b). This shows that obvious greedy algorithms fail to find an optimal solution, as Liebling et al. [13] observed. They also provided an instance of the problem in a unit square for $k = n/2$ in which the length of an optimal solution is in $\Omega(n\sqrt{n})$, whereas the trivial lower bound (the sum of lengths of straight-line segments connecting the pairs of terminals) is only $O(n)$. The example is based on the existence of expander graphs with a quadratic number of edge crossings. In Fig. 1(c), we provide an example in which one of the curves in the optimal solution can be arbitrarily longer than the trivial lower bound for the corresponding color.

Efrat et al. [7] suggested an approximation algorithm for k-CESF. The key ingredient of their algorithm is the following observation, which shows how to make a pair of given trees non-crossing: reroute one of the trees using a "shell" around the other tree. For any geometric graph G, we denote its total edge length by $|G|$.

Lemma 1 (Efrat et al. [7]). *Let R and B be two trees in the plane spanning red and blue terminals, respectively. Then, there exists a tree R' spanning the red terminals such that (i) R' and B are non-crossing and (ii) $|R'| \leq |R| + 2|B|$.*

Efrat et al. [7] start with k (possibly intersecting) minimum spanning trees, one for each color. Then, they iteratively go through these trees in order of increasing length.

In every step, they reroute the next tree by laying a shell around the current solution as in Lemma 1. Their algorithm has approximation factor $k\rho$. In the full version of the paper [3], we show that the algorithm even yields approximation factor $k+\varepsilon$ if we use a PTAS for EST for the initial solution to each color.

Theorem 1. *For every $\varepsilon > 0$, there is a $(k+\varepsilon)$-approximation algorithm for k-CESF.*

For even k, we can slightly improve on this by using our PTAS for 2-CESF (Thm. 4).

Theorem 2. *For every $\varepsilon > 0$, there is a $(k-1+\varepsilon)$-approximation algorithm for k-CESF if k is even.*

Next, we present an approximation algorithm for k-CESF whose ratio depends only logarithmically on k, but also depends on $\sqrt{n}$. The algorithm employs a space-filling curve through a set of given points. The curve was utilized in a heuristic for $(n/2)$-CESF by Liebling et al. [13]. Recently, Chan et al. [5] showed that the approach yields an $O(\sqrt{n}\log n)$-approximation for $(n/2)$-CESF. We show that similar arguments yield approximation ratio $O(\sqrt{n}\log k)$ for general k.

Theorem 3. *k-CESF admits a (randomized) $O(\sqrt{n}\log k)$-approximation algorithm.*

Proof. Chan et al. [5] gave a randomized algorithm to construct a curve C through the given set P of n points. Their curve has small *stretch*, that is, the ratio between the Euclidean distance $d(p,q)$ of two points $p,q \in P$ and their distance $d_C(p,q)$ along the curve is small. Assuming that the points are scaled to lie in a unit square, Chan et al. showed, for a fixed pair of points $p,q \in P$, $\mathbb{E}[d_C(p,q)] \leq O(\sqrt{n}\log(\frac{1}{d(p,q)}))\cdot d(p,q)$. Using C, we construct a solution to k-CESF so that, for every color, the terminals are visited in the order given by the curve; and thus, the solution to every color is a path. All paths can be wrapped around the curve without intersecting each other; see Fig. 2.

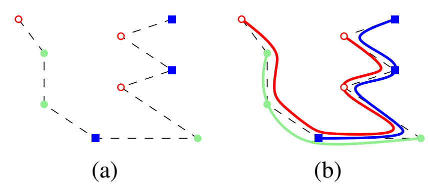

Fig. 2: (a) A low-stretch curve C through the terminals; (b) a 3-CESF solution to the instance created by wrapping paths around C.

If the order of the points along the curve for a specific color i is $p_1^i,\dots,p_{n_i}^i$, then the length of the corresponding path is $\sum_{j=1}^{n_i-1} d_C(p_j^i,p_{j+1}^i) = d_C(p_1^i,p_{n_i}^i)$. Let $\bar{d} = \sum_{i=1}^{k} d(p_1^i,p_{n_i}^i)/k$. The total (expected) length of the solution is

$$\mathrm{ALG} = \sum_{i=1}^{k} \mathbb{E}[d_C(p_1^i,p_{n_i}^i)] \leq \sum_{i=1}^{k} O(\sqrt{n}\log(1/d(p_1^i,p_{n_i}^i)))\cdot d(p_1^i,p_{n_i}^i)\ .$$

Given that $\log$ is concave, this expression is bounded by $\sum_{i=1}^{k} O(\sqrt{n}\log(1/\bar{d}))\cdot\bar{d}$; see Chan et al. [5]. Since the optimal solution to P connects all pairs of terminals of the same color (possibly using non-straight-line curves), $\mathrm{OPT} \geq \sum_{i=1}^{k} d(p_1^i,p_{n_i}^i) = k\bar{d}$. Hence,

$$\mathrm{ALG} \leq \sum_{i=1}^{k} O(\sqrt{n}\log(k/\mathrm{OPT}))\cdot\mathrm{OPT}/k \leq O(\sqrt{n}\log k)\,\mathrm{OPT}\ . \qquad \square$$

3 PTAS for 2-CESF

In this section, we show that 2-CESF admits a PTAS. We follow Arora's approach for computing EST [1], which consists of the following steps. First, Arora performs a recursive geometric partitioning of the plane using a quadtree and snaps the input points to the corners of the tree. Next, he defines an *r-light* solution, which is allowed to cross an edge of a square in the quadtree at most r times and only at so-called *portals*. Then he builds an optimal *portal-respecting* solution using dynamic programming (DP), and finally trims the edges of the solution to get the result. To get an algorithm for 2-CESF, we modify these steps as follows:

(i) The perturbation step, which snaps the terminal to a grid, is modified to avoid crossings between trees. Similarly, the reverse step transforming a perturbed instance solution into one to the original instance is different; see Lemmas 2 and 3.
(ii) We use a different notion of an r-light solution in which every portal is crossed at most r times. We devise a *portal-crossing reduction* that reduces the number of crossings to $r = 3$; see Lemma 5.
(iii) The base case of the DP needs a special modification; it computes a set of crossing-free Steiner trees of minimum total length (see Lemma 6).

We assume that the bounding rectangles of the two sets of input terminals overlap; otherwise, we can use a PTAS for the Steiner tree of each input set individually. We first snap the instance to an $(L \times L)$-grid with $L = O(n)$. We proceed as follows. Let L_0 be the diameter of the smallest bounding box of the given 2-CESF instance. We place an $(L \times L)$-grid of granularity (grid cell size) $g = L_0/L$ inside the bounding box. By scaling the instance appropriately, we can assume that the granularity is $g = 1$. We move each terminal of one color to the nearest grid point in an even row and column, and each terminal of the other color to the nearest grid point in an odd row and column. Thus, the grid point for each terminal is uniquely defined, and no terminals of different color end up at the same location. If there are more terminals of the same color on a grid point, we remove all but one of them. We call the resulting instance a *perturbed* instance.

Lemma 2. *Let OPT_I be the length of an optimal solution to a 2-CESF instance I of n terminals and let $\varepsilon > 0$. There is an $(L \times L)$-grid with $L = O(n/\varepsilon)$ such that $\mathrm{OPT}_{I^*} \leq (1+\varepsilon)\,\mathrm{OPT}_I$, where OPT_{I^*} is the length of an optimal solution to the perturbed instance I^*.*

Proof. Choose L to be a power of 2 within the interval $[3\sqrt{2}n/\varepsilon, 6\sqrt{2}n/\varepsilon]$ and perturb the instance as described above. Consider an optimal solution to I. Iteratively, we connect every terminal in I^* to the optimum solution as follows: Connect the terminal to the closest point of the tree in the optimum solution that has the same color. If this line segment crosses the tree of the other color, then reroute this tree around the line segment by using two copies of the line segment. Two copies suffice even if the other tree is crossed more than once since all crossing edges can be connected to the two new line segments. The distance between the terminal and the tree is at most the distance between the terminal and the corresponding terminal in I, which is bounded by $\sqrt{2}$ as we are assuming the unit grid. Hence, we pay at most $3\sqrt{2}$ for connecting the terminal.

Since the bounding rectangles of the input terminals overlap, $\mathrm{OPT}_I \geq L$. Thus, the additional length of an optimal solution to I^* is

$$\mathrm{OPT}_{I^*} - \mathrm{OPT}_I \leq 3\sqrt{2}n \leq \varepsilon \cdot \mathrm{OPT}_I \ . \qquad \square$$

The next lemma, proven analogously to Lemma 2, shows that we can transform a solution to the perturbed instance into one to the original instance.

Lemma 3. *Given a solution $\mathcal{T}$ to the perturbed instance as defined in Lemma 2, we can transform $\mathcal{T}$ into a solution to the original instance, increasing its length by at most $\varepsilon \cdot \mathrm{OPT}_I$.*

In the following, we assume that the instance is perturbed. We place a quadtree in dependence of two integers $a, b \in [0, \dots, L-1]$ that we choose independently uniformly at random. We place the origin of the coordinate system on the bottom left corner of the bounding box of our instance. Then we take a box B whose width and height is twice the width and height of the bounding box. We place it such that its bottom left corner has coordinates $(-a, -b)$. Note that the bounding box is inside B. We extend the $(L \times L)$-grid to cover B. Thus, we have an $(L' \times L')$-grid with $L' = 2L$.

Then we partition B with a quadtree along the $(L' \times L')$-grid. The partition is stopped when the current quadtree box coincides with a grid cell. We define the *level of a quadtree square* to be its depth in the quadtree. Thus, B has level 0, whereas the level of a leaf is bounded by $\log L' = \log(2L) = O(\log n)$. Then, for each grid line ℓ, we define its *level* as the highest (that is, of minimum value) level of all the quadtree squares that touch ℓ (but which are not crossed by it).

Let $m = \lceil 4 \log L'/\varepsilon \rceil$. On each grid line ℓ of level i, we place $2^i \cdot m$ equally spaced points. We call these points *portals*. Thus, each square contains at most m portals on each of its edges. A solution that crosses the grid lines only at portals is called *portal-respecting*. We show that there is a close-to-optimal portal-respecting solution. Note that, in contrast to Arora, we first make the solution portal-respecting before reducing the number of crossings on each grid line. The proof of the following lemma is similar to the Arora's prove and is provided in the full version of the paper [3].

Lemma 4. *Let OPT_I be the length of an optimal solution to a 2-CESF instance I, and let $\varepsilon > 0$ be as in the definition of m. Then, there exists a position of the quadtree and a portal-respecting solution to I of length at most $(1+\varepsilon)\,\mathrm{OPT}_I$.*

The last ingredient of our DP is to reduce the number of crossings in every portal. We call a solution *r-light* if each portal is crossed at most r times.

In the following, we explain an operation which we call a *portal-crossing reduction*. We are given a portal-respecting solution consisting of two Steiner trees R and B (red and blue) and we want to reduce (that is, modify without increasing its length) it such that R and B pass through each portal at most three times in total.

Lemma 5. *Every portal-respecting solution of 2-CESF can be transformed into a 3-light portal-respecting solution without increasing its length.*

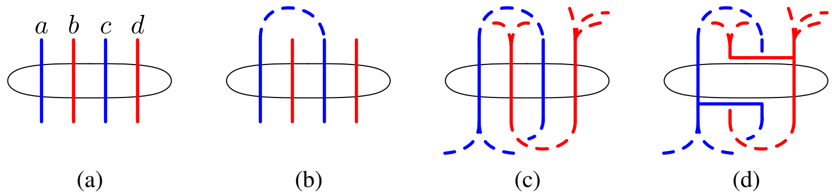

Fig. 3: A portal modification for four passes.

Proof. Consider a sequence of passes through a portal. We assume that there are no terminals in the portals. If two adjacent passes belong to the same tree, then we can eliminate one of them by snapping it to the other one. Note that this may create cycles, but they can be broken by removing the longest part of each cycle. Therefore, we can assume that the passes form an alternating sequence. It suffices to show that any alternating sequence of four passes can be reduced to two passes by shortening the trees. Let a, b, c, and d be such a sequence as shown in Fig. 3a, where a and c belong to B and b and d to R. We cut the passes b and c. This results in two connected components in each tree. W.l.o.g., a and the upper part of c belong to the same connected component; see Fig. 3b. Otherwise, we can change the colors because (i) a and the lower part of c are connected, and (ii) the upper part of b and d are connected.

Since R and B are disjoint, d and the lower part of b are in the same connected component; see Fig. 3c. Then, we connect the component as shown in Fig. 3d and shorten the trees (e.g., the lower part of b can be reduced to a terminal of R). Note that the passes a and d remain in the solution, while the passes b and c are eliminated. We repeat the procedure for the remaining passes, until there are at most three passes left. The length of the solution does not increase because the portal has zero width. □

With the next Lemma 6, we show how to find a close-to-optimal 3-light portal-respecting solution to the perturbed instance. We assume that an appropriate quadtree (as defined in Lemma 4) is given.

Lemma 6. *Let $\varepsilon > 0$ be as above. Given a perturbed instance I^* of an n-terminal 2-CESF instance, we can compute, in time $O(n^{O(1/\varepsilon)})$ and $O(n^{O(1/\varepsilon)})$ space, a solution of length at most $(1+\varepsilon)\,\mathrm{OPT}_{I^*}$, where OPT_{I^*} is the length of an optimal 3-light portal-respecting solution to I^*.*

Proof. We use DP with a subproblem consisting of (a) a square of the quadtree, (b) a sequence of up to three red and blue points on each portal on the border of the square, and (c) a *non-crossing* partition of these points into sets of the same color. A partition of these points is non-crossing if for no four points a, b, c, d, occurring in that order on the boundary of the square, it holds that a and c belong to one set of the partition, and b and d to another one. The goal is to find an optimal collection of crossing-free red and blue Steiner trees, such that each set of the partition and each terminal inside the square is contained in a tree of the same color.

The base case of our DP is a unit square, which is either empty or contains terminals only at corners of the square. If the square is empty, we consider each set of the partition as an instance of 1-CESF and solve it by the PTAS for EST [1]. For each point set, we force its Steiner tree to lie inside its convex hull, by projecting any part outside the convex hull to its border. Since the partition is non-crossing, the convex hulls of its point sets are pairwise disjoint. Therefore, the Steiner trees and their union is also a close-to-optimal solution to the base case. If the square contains (up to four) terminals at the corners, these terminals are treated in a similar way as portals.

For composite squares in the quadtree, we proceed as follows. For the four squares that subdivide the composite square, we consider all combinations of all possible (b) and (c) that match together and match the subproblem. In the DP, we already have computed a close-to-optimal solution to every choice of (b) and (c) of each of the four squares; taking the best combination gives a close-to-optimal solution.

The size of the DP table is proportional to the number of subproblems, that is, (a)×(b)×(c). There are $O(n^2)$ squares in the quadtree in total. Each square contains at most $m = O(\log n/\varepsilon)$ portals. For each portal, there is a constant number of possible sequences of up to three colored points. Thus, there are $2^{O(\log n/\varepsilon)} = n^{O(1/\varepsilon)}$ possibilities for (b). Since the number of non-crossing partitions of a set of k elements is the k'th Catalan number C_k, we have $C_{O(\log n/\varepsilon)} < 2^{O(\log n/\varepsilon)} = n^{O(1/\varepsilon)}$ possibilities for (c). In total, we consider $n^{O(1/\varepsilon)}$ subproblems in the DP.

The running time to solve an instance of the base case is polynomial in $m = O(\log n/\epsilon)$. The running time to handle a composite square is polynomial in $(n^{O(1/\epsilon)})^4$, which is $n^{O(1/\epsilon)}$. Thus, the total running time is bounded by $n^{O(1/\epsilon)}$. □

Now we prove the main result of this section.

Theorem 4. *2-CESF admits a PTAS.*

Proof. Consider a 2-CESF instance I. Let OPT be the length of an optimum solution. For any $\varepsilon > 0$, by Lemmas 2, 4 and 5, the length, OPT', of an optimal 3-light portal-respecting solution to the perturbed version of I is a most $(1+\varepsilon)\,\mathrm{OPT}$. Using Lemma 6, we find a 3-light portal-respecting solution to the perturbed instance of length at most $(1+\varepsilon)\,\mathrm{OPT}' \le (1+\varepsilon)(1+\varepsilon)\,\mathrm{OPT}$. By Lemma 3, we transform the solution into a solution to I by increasing its length by at most $\varepsilon \cdot \mathrm{OPT}$. Therefore, for every $\varepsilon' > 0$, we can construct a solution to I of length $(1+\varepsilon)(1+\varepsilon)\,\mathrm{OPT} + \varepsilon \cdot \mathrm{OPT} \le (1+\varepsilon')\,\mathrm{OPT}$ by choosing $\varepsilon > 0$ appropriately. □

4 Algorithm for 3-CESF

The above approach for 2-CESF cannot be directly applied to 3-CESF since optimal trees may need to pass portals many times. For example, the three paths crossing the portal in Fig. 4 are difficult because we cannot locally reroute to make them $O(1)$-light as in Lemma 5.

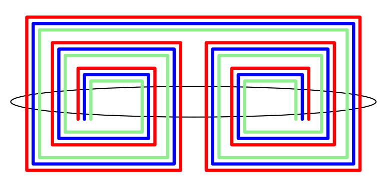

Fig. 4: A difficult portal crossing of a 3-CESF instance.

Instead, we now improve the approximation ratio of $3 + \varepsilon$ (from Theorem 1) to $5/3 + \varepsilon$. We re-use some ideas of the approach for 2-CESF.

To this end, take an optimal solution T for 3-CESF. The terminals are red, green, and blue; we call the corresponding trees R, G, and B. We assume that B is the cheapest among the three trees. In the beginning, we construct a quadtree partitioning the plane and choose the portals, for a given ε, as described in Sect. 3. We then make the solution portal-respecting, which results in a solution T^* consisting of trees R^*, G^*, and B^*. In expectation, this increases the length of each of the trees (and hence, of T) by a factor of at most $1 + \varepsilon$.

First, we show that we have few portal passes if the blue and the green tree do not *meet* at any portal, that is, no blue and green passes are adjacent.

Lemma 7. *Consider a portal-respecting solution T^* to 3-CESF consisting of trees R^*, G^*, B^*. If B^* and G^* do not meet at any portal, then T^* can be transformed into a 7-light portal-respecting solution.*

Proof. Apply the portal-crossing reduction from Lemma 5 and consider a portal. Recall that, after this operation, there are no $rbrb$ and $rgrg$ subsequences in the passes of the portal. Here, r, b, and g correspond to the passes of the trees R^*, B^*, and G^*, respectively. If the portal has only one blue or one green pass, then the solution is already 7-light at the portal (with the longest possible sequences $rgrbrgr$ and $rbrgrbr$, respectively). Otherwise, it contains at least two blue and at least two green passes. Notice that the sequence of passes must be r-alternate, that is, of the form $\ldots r \circ r \circ r \ldots$ since blue and green do not meet. Thus, a sequence of more than 7 passes must contain a subsequence $grbrgrb$ (or a symmetric one, $brgrbrg$). These subsequences are reducible. See Fig. 5 for one of the possible cases, the other cases are analogous. □

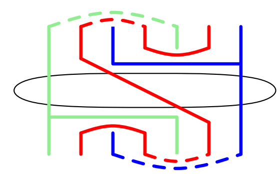

Fig. 5: Constructing a 7-light solution to an instance without adjacent blue-green passes (one of several possible cases).

Now, we show that T^* can be transformed into a 10-light portal-respecting solution T' of length at most $|R^*| + |G^*| + 3|B^*|$.

Lemma 8. *A portal-respecting solution T^* to 3-CESF, consisting of trees B^*, R^*, and G^*, can be transformed into a portal-respecting solution T' such that*
(i) T' passes at most 10 times through each portal, and
(ii) $|T'| \leq |R^*| + |G^*| + 3|B^*|$.

Proof sketch. We define a BG-*solution*; informally, this is a solution in which we are allowed to connect green branches to the blue tree (if they never meet, we can apply Lemma 7). We prove the lemma in two steps. First, we show that T can be transformed to a portal-respecting BG-solution T^{BG} with at most 6 passes per portal having the same (or smaller) length. In this solution, B^* remains connected and passes each portal at most twice. Then, we further modify T^{BG} to get a portal-respecting solution T^* with at most 10 passes per portal and the desired length by laying a shell around B^* to reroute G^*. The full proof is given in the full version of the paper [3]. □

Before we describe our approximation algorithm, we first need to discuss the perturbation step. The perturbation itself is the same as in Sect. 3: we move each terminal to a uniquely defined closest grid point (we assign the grid points of even row and odd column to the third color) and merge terminals of the same color to one terminal. However, we need a different technique to transform a solution to the original instance into a solution to the perturbed instance and vice versa.

Lemma 9. *Let I be a 3-CESF instance with n terminals, let OPT_I be the length of an optimal solution to I, and let $\varepsilon > 0$. Then, we can place an $(L \times L)$-grid with $L = O(n/\varepsilon)$ such that, for the perturbed instance I^* of I, $\mathrm{OPT}_{I^*} \leq (1+\varepsilon)\,\mathrm{OPT}_I$.*

Proof sketch. We proceed similar to the proof of Lemma 2 by connecting each terminal of I^* to the nearest point of its corresponding tree. Since this connection can cross segments of two colors, we have to be more careful with the rerouting. The full proof is given in the full version of the paper [3]. □

Analogously to the proof of Lemma 9, we transform a solution to a perturbed instance back into one to the original instance not increasing the length by much. Then, we combine the lemmas to prove the main result of this section.

Lemma 10. *We can transform a solution $\mathcal{T}$ to the perturbed instance I^* into a solution to the original instance I, increasing the length by at most $\varepsilon\,\mathrm{OPT}_I$.*

Proof. Iteratively connect each terminal of the original instance to the solution $\mathcal{T}$ analogously to the proof of Lemma 9. □

Theorem 5. *For every $\varepsilon > 0$, 3-CESF admits a $(5/3 + \varepsilon)$-approximation algorithm.*

Proof. Let $\varepsilon' = \sqrt[3]{1 + 3\varepsilon/5} - 1$. Let T be an optimal solution to a 3-CESF instance I with trees R, G and B. W.l.o.g., assume that $|B| \leq |R|, |G|$. Denote by $\mathrm{OPT}_I = |R| + |G| + |B|$ the length of T. We first construct a portal-respecting solution T^* of length $|T^*| = |R^*| + |G^*| + |B^*| \leq (1+\varepsilon')(|R| + |G| + |B|)$. Then, Lemma 8 yields an optimal 10-light portal-respecting solution T' of length

$$\begin{aligned}|T'| \leq |R^*| + |G^*| + 3|B^*| \leq 5/3 \cdot |T^*| &\leq 5/3 \cdot (1+\varepsilon') \cdot (|R| + |G| + |B|) \\ &= 5/3 \cdot (1+\varepsilon') \cdot \mathrm{OPT}_I \ .\end{aligned}$$

Using a DP similar to the one described in Sect. 3 and using Lemma 9, we find a 10-light portal-respecting solution of length $(1+\varepsilon')|T'|$ to the perturbed instance I^* of I. By Lemma 10, we can transform our solution to I^* into a solution to I whose total length is bounded by

$$(1+\varepsilon')^2|T'| \leq 5/3(1+\varepsilon')^3\,\mathrm{OPT}_I < (5/3+\varepsilon)\,\mathrm{OPT}_I \ . \qquad \square$$

5 Conclusion

We have presented approximation algorithms for k-CESF. We leave the following questions open. Is k-CESF APX-hard for some $k \geq 3$? Can we improve the running time of the PTAS for 2-CESF from $O(n^{O(1/\varepsilon)})$ to $O(n(\log n)^{O(1/\varepsilon)})$ as Arora [1] did for EST?

Currently, we are studying an "anchored" version of k-CESF where the only allowed Steiner points are input points of a different color. Any α-approximation for k-CESF yields an $\alpha(1+\sqrt{3})/2$- approximation for the anchored version.

Acknowledgments. We are grateful to Alon Efrat, Jackson Toeniskoetter, and Thomas van Dijk for the initial discussion of the problem.

References

1. Arora, S.: Polynomial time approximation schemes for Euclidean traveling salesman and other geometric problems. J. ACM 45(5), 753–782 (1998)
2. Bastert, O., Fekete, S.P.: Geometric wire routing. Tech. Rep. 96.247, Universität zu Köln (1998), available at http://e-archive.informatik.uni-koeln.de/247
3. Bereg, S., Fleszar, K., Kindermann, P., Pupyrev, S., Spoerhase, J., Wolff, A.: Colored non-crossing Euclidean steiner forest. CoRR abs/1509.05681 (2015), http://arxiv.org/abs/1509.05681
4. Borradaile, G., Klein, P., Mathieu, C.: An $O(n \log n)$ approximation scheme for Steiner tree in planar graphs. ACM Trans. Algorithms 5(3), 31 (2009)
5. Chan, T.M., Hoffmann, H.F., Kiazyk, S., Lubiw, A.: Minimum length embedding of planar graphs at fixed vertex locations. In: Wismath, S., Wolff, A. (eds.) Graph Drawing (GD'13). LNCS, vol. 8242, pp. 376–387. Springer, Heidelberg (2013)
6. Chung, F.R.K., Graham, R.L.: A new bound for Euclidean Steiner minimal trees. Annals New York Acad. Sci. 440(1), 328–346 (1985)
7. Efrat, A., Hu, Y., Kobourov, S.G., Pupyrev, S.: Mapsets: Visualizing embedded and clustered graphs. In: Duncan, C., Symvonis, A. (eds.) Graph Drawing (GD'14). LNCS, vol. 8871, pp. 452–463. Springer, Heidelberg (2014)
8. Efrat, A., Kobourov, S.G., Lubiw, A.: Computing homotopic shortest paths efficiently. Comput. Geom. Theory Appl. 35(3), 162–172 (2006)
9. Erickson, J., Nayyeri, A.: Shortest non-crossing walks in the plane. In: Proc. ACM-SIAM Symp. Discrete Algorithms (SODA'11). pp. 297–308 (2011)
10. Garey, M.R., Johnson, D.S.: Computers and Intractability: A Guide to the Theory of NP-Completeness. W. H. Freeman & Co., New York, NY, USA (1979)
11. Hurtado, F., Korman, M., van Kreveld, M., Löffler, M., Sacristán, V., Silveira, R.I., Speckmann, B.: Colored spanning graphs for set visualization. In: Wismath, S., Wolff, A. (eds.) Graph Drawing (GD'13), LNCS, vol. 8242, pp. 280–291. Springer, Heidelberg (2013)
12. Kusakari, Y., Masubuchi, D., Nishizeki, T.: Finding a noncrossing Steiner forest in plane graphs under a 2-face condition. J. Combin. Optim. 5, 249–266 (2001)
13. Liebling, T.M., Margot, F., Müller, D., Prodon, A., Stauffer, L.: Disjoint paths in the plane. ORSA J. Comput. 7(1), 84–88 (1995)
14. Löffler, M.: Existence and computation of tours through imprecise points. Int. J. Comput. Geom. Appl. 21(1), 1–24 (2011)
15. Mitchell, J.S.: Guillotine subdivisions approximate polygonal subdivisions: A simple polynomial-time approximation scheme for geometric TSP, k-MST, and related problems. SIAM J. Comput. 28(4), 1298–1309 (1999)
16. Mitchell, J.S.: Geometric shortest paths and network optimization. In: Urrutia, J., Sack, J.R. (eds.) Handbook of Computational Geometry, chap. 15, pp. 633–701. North-Holland (2000)
17. Müller-Hannemann, M., Tazari, S.: A near linear time approximation scheme for Steiner tree among obstacles in the plane. Comput. Geom. Theory Appl. 43(4), 395–409 (2010)
18. Papadopoulou, E.: k-pairs non-crossing shortest paths in a simple polygon. Int. J. Comput. Geom. Appl. 9(6), 533–552 (1999)
19. Polishchuk, V., Mitchell, J.S.B.: Thick non-crossing paths and minimum-cost flows in polygonal domains. In: Proc. ACM Symp. Comput. Geom. (SoCG'07). pp. 56–65 (2007)
20. Verbeek, K.: Homotopic $\mathcal{C}$-oriented routing. In: Didimo, W., Patrignani, M. (eds.) Graph Drawing (GD'12). LNCS, vol. 7704, pp. 272–278. Springer, Heidelberg (2013)