# Network Optimization on Partitioned Pairs of Points*†

**Esther M. Arkin[‡1], Aritra Banik[2], Paz Carmi[3], Gui Citovsky[4], Su Jia[5], Matthew J. Katz[§6], Tyler Mayer[¶7], and Joseph S. B. Mitchell[‖8]**

1 **Dept. of Applied Mathematics and Statistics, Stony Brook University, Stony Brook, USA**
`esther.arkin@stonybrook.edu`
2 **Dept. of Computer Science, Ben-Gurion University, Beersheba, Israel**
`aritrabanik@gmail.com`
3 **Dept. of Computer Science, Ben-Gurion University, Beersheba, Israel**
`carmip@gmail.com`
4 **Google, Manhattan, USA**
`gcitovsky@gmail.com`
5 **Dept. of Applied Mathematics and Statistics, Stony Brook University, Stony Brook, USA**
`su.jia@stonybrook.edu`
6 **Dept. of Computer Science, Ben-Gurion University, Beersheba, Israel.**
`matya@cs.bgu.ac.il`
7 **Dept. of Applied Mathematics and Statistics, Stony Brook University, Stony Brook, USA**
`tyler.mayer@stonybrook.edu`
8 **Dept. of Applied Mathematics and Statistics, Stony Brook University, Stony Brook, USA.**
`joseph.mitchell@stonybrook.edu`

## Abstract

Given $n$ pairs of points, $\mathcal{S} = \{\{p_1, q_1\}, \{p_2, q_2\}, \ldots, \{p_n, q_n\}\}$, in some metric space, we study the problem of two-coloring the points within each pair, red and blue, to optimize the cost of a pair of node-disjoint networks, one over the red points and one over the blue points. In this paper we consider our network structures to be spanning trees, traveling salesman tours or matchings. We consider several different weight functions computed over the network structures induced, as well as several different objective functions. We show that some of these problems are NP-hard, and provide constant factor approximation algorithms in all cases.

**1998 ACM Subject Classification** F.2.2 Nonnumerical Algorithms and Problems

**Keywords and phrases** Spanning Tree, Traveling Salesman Tour, Matching, Pairs, NP-hard

**Digital Object Identifier** 10.4230/LIPIcs.ISAAC.2017.6

* We acknowledge support from the US-Israel Binational Science Foundation (Projects 2010074, 2016116).
† A full version of the paper is available at `https://arxiv.org/abs/1710.00876`.
‡ E. Arkin acknowledges support from the National Science Foundation (CCF-1526406).
§ M. Katz was partially supported by grant 1884/16 from the Israel Science Foundation.
¶ T. Mayer's research is partially supported by the NSF (IIS-1247726, CNS-1408695, CNS-1755615, CCF-1439084, CCF-1617618, CCF-BSF-1716252, CCF 1725543, CCF-1526406, CCF-1737939) and by Sandia National Laboratories and NetApp.
‖ J. Mitchell acknowledges support from the National Science Foundation (CCF-1526406).


28th International Symposium on Algorithms and Computation (ISAAC 2017).
Editors: Yoshio Okamoto and Takeshi Tokuyama; Article No. 6; pp. 6:1–6:12
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

**Table 1** Table of results: $\alpha$ is the Steiner ratio and $\beta$ the best approximation factor of the TSP in the underlying metric space. Unless specified otherwise, all other results in this table apply to general metric spaces.

| | min $\|f(R)\| + \|f(B)\|$ | min-max$\{\|\lambda(R)\|, \|\lambda(B)\|\}$ | min-max$\{\|f(R)\|, \|f(B)\|\}$ |
|---|---|---|---|
| Spanning tree | $3\alpha$ | 9<br>3 for $\mathbb{R}$ | $4\alpha$ |
| Matching | 2 | 3 | 3 |
| TSP tour | $3\beta$ | 18 | $6\beta$ |

## 1 Introduction

We study a class of network optimization problems on *pairs* of sites in a metric space. Our goal is to determine how to split each pair, into a "red" site and a "blue" site, in order to optimize *both* a network on the red sites and a network on the blue sites. In more detail, given $n$ pairs of points, $\mathcal{S} = \{\{p_1, q_1\}, \{p_2, q_2\}, \ldots, \{p_n, q_n\}\}$, in the Euclidean plane or in a general metric space, we define a *feasible coloring* of the points in $S = \bigcup_{i=1}^{n} \{p_i, q_i\}$ to be a coloring, $S = R \cup B$, such that $p_i \in R$ if and only if $q_i \in B$. Among all feasible colorings of $\mathcal{S}$, we seek one which optimizes the cost function over a *pair* of network structures, spanning trees, traveling salesman tours (TSP tours) or matchings, one on the red set and one on the blue set. Let $f(X)$ be a certain structure computed on point set $X$ and let $\lambda(X)$ be the longest edge of a bottleneck structure, $f(X)$, computed on point set $X$. For each of the aforementioned structures we consider the objective of (over all feasible colorings $S = R \cup B$) minimizing $|f(R)| + |f(B)|$, minimizing $\max\{|f(R)|, |f(B)|\}$ and minimizing $\max\{|\lambda(R)|, |\lambda(B)|\}$. Here, $|\cdot|$ denotes the cost (e.g., sum of edge lengths) of the structure.

The problems we study are natural variants of well-studied network optimization problems. Our motivation comes also from a model of secure connectivity in networks involving facilities with replicated data. Consider a set of facilities each having two (or more) replications of their data; the facilities are associated with pairs of points (or $k$-tuples of points in the case of higher levels of replication). Our goal may be to compute two networks (a "red" network and a "blue" network) to interconnect the facilities, each network visiting exactly one data site from each facility; for communication connectivity, we would require each network to be a tree, while for servicing facilities with a mobile agent, we would require each network to be a Hamiltonian path/cycle. By keeping the red and blue networks distinct, a malicious attack within one network is isolated from the other.

**Our results.** We show that several of these problems are NP-hard and give O(1)-approximation algorithms for each of them. Table 1 summarizes our O(1)-approximation results.

**Related work.** Several optimization problems have been studied of the following sort: Given sets of tuples of points (in a Euclidean space or a general metric space), select exactly one point or at least one point from each tuple in order to optimize a specified objective function on the selected set. Gabow et al. [12] explored the problem in which one is given a directed acyclic graph with a source node $s$ and a terminal node $t$ and a set of $k$ pairs of nodes, where the objective was to determine if there exists a path from $s$ to $t$ that uses at most one node from each pair. Myung et al. [17] introduced the Generalized Minimum Spanning Tree Problem: Given an undirected graph with the nodes partitioned into subsets, compute a minimum spanning tree that uses exactly one point from each subset. They show that this problem is NP-hard and that no constant-factor approximation algorithm exists for this problem unless $P = NP$. Related work addresses the generalized traveling salesperson problem [6, 18, 19, 20], in which a tour must visit one point from each of the given subsets. Arkin et al. [4] studied the problem in which one is given a set $V$ and a set of subsets of $V$, and one wants to select at least one element from each subset in order to minimize the diameter of the chosen set. They also considered maximizing the minimum distance between any two elements of the chosen set. In another recent paper, Consuegra et al. [8] consider several problems of this kind. Abellanas et al. [1], Das et al. [10] and Khantemouri et al. [15] considered the following problem. Given colored points in the Euclidean plane, find the smallest region of a certain type (e.g., strip, axis-parallel square, etc.) that encloses at least one point from each color. Barba et al. [5] studied the problem in which one is given a set of colored points (of $t$ different colors) in the Euclidean plane and a vector $c = (c_1, c_2, \ldots, c_t)$, and the goal is to find a region (axis-aligned rectangle, square, disk) that encloses exactly $c_i$ points of color $i$ for each $i$. Efficient algorithms are given for deciding whether or not such a region exists for a given $c$.

While optimization problems of the "one of a set" flavor have been studied extensively, the problems we study here are fundamentally different: we care not just about a single structure (e.g., network) that makes the best "one of a set" choices on, say, pairs of points; we must consider also the cost of a second network on the "leftover" points (one from each pair) *not* chosen. As far as we know, the problem of partitioning points from pairs into two sets in order to optimize objective functions on *both* sets has not been extensively studied. One recent work of Arkin et al. [3] does address optimizing objectives on both sets: Given a set of pairs of points in the Euclidean plane, color the points red and blue so that if one point of a pair is colored red (resp. blue), the other must be colored blue (resp. red). The objective is to optimize the radii of the minimum enclosing disk of the red points and the minimum enclosing disk of the blue points. They studied the objectives of minimizing the sum of the two radii and minimizing the maximum radius.

## 2 Spanning Trees

Let $MST(X)$ be a minimum spanning tree over the point set $X$, and $|MST(X)|$ be the cost of the tree, i.e. sum of edge lengths. Let $\lambda(X)$ be the longest edge in a bottleneck spanning tree on point set $X$ and $|\lambda(X)|$ be the cost of that edge. Given $n$ pairs of points in a metric space, find a feasible coloring which minimizes the cost of a pair of spanning trees, one built over each color class.

### 2.1 Minimum Sum

In this section we consider minimizing $|MST(R)| + |MST(B)|$.

▶ **Theorem 1.** *The Min-Sum 2-MST problem is NP-hard in general metric spaces.* [See full paper [2] for proof.]

#### An $O(1)$-approximation algorithm for Min-Sum 2-MST problem.

Compute $MST(S)$, a minimum spanning tree on all $2n$ points. Imagine removing the heaviest edge, $h$, from $MST(S)$. This leaves us with two trees; $T_1$ and $T_2$. Perform a preorder traversal on $T_1$, coloring nodes red as long as there is no conflict. If there is a conflict ($q_i$ is reached in the traversal and $p_i$ was already colored to red) then color the node blue. Repeat this for $T_2$. We then return the coloring $S = R \cup B$ as our approximate coloring.

**Case 1: All nodes in $T_1$ are of the same color and all nodes in $T_2$ are of the same color.** This partition is optimal. To see this, note that the weight of $MST(S) \setminus \{h\}$ is a lower bound on the cost of the optimal solution as it is the cheapest way to create two trees, the union of which span all of the input nodes. Since each tree is single colored, we know that each tree must have $n$ points, exactly one from each pair, and thus is also feasible to our problem.

**Case 2: One tree is multicolored and the other is not.** Let $OPT$ be the optimal solution. Suppose without loss of generality that $T_1$ contains only red nodes and $T_2$ contains both blue and red nodes. Then, there must be a pair with both nodes in $T_2$. Imagine also constructing an MST on each color class of an optimal coloring. By definition, in the MSTs built over each color class, at least one point in $T_2$ must be connected to a point in $T_1$. This implies that the weight of the optimal solution is at least as large as $|h|$, as $h$ is the cheapest edge which spans the cut $(T_1, T_2)$. Therefore, $|h| \leq |OPT|$.

Consider $MST(R)$. By the Steiner property, we have that an MST over a subset $U \subseteq S$ has weight at most $\alpha|MST(S)|$ where $\alpha$ is the Steiner ratio of the metric space. Recall that $|MST(S) \setminus \{h\}| \leq |OPT|$. In this case, since $|h| \leq |OPT|$, we have that $|MST(R)| \leq \alpha|MST(S)| \leq 2\alpha|OPT|$.

Next, consider building $MST(B)$. Since no blue node exists in $T_1$, there does not exist an edge that crosses the cut $(T_1, T_2)$ in $MST(B)$, and thus we have that $|MST(B)| \leq \alpha|MST(S) \setminus \{h\}| \leq \alpha|OPT|$. Therefore, $|MST(R) \cup MST(B)| \leq 3\alpha|OPT|$.

**Case 3: Both trees are multicolored.** In this case, there are two pairs one with both nodes contained in $T_1$ and one with both nodes contained in $T_2$. Imagine, again, constructing an MST on each color class in this optimal coloring. In this case, there must be at least two edges crossing the cut $(T_1, T_2)$, one edge belonging to each tree. Note that each of these edges has weight at least $|h|$ as $h$ is the cheapest edge spanning the cut $(T_1, T_2)$, implying that $|h| \leq |OPT|/2$. Thus, $|MST(S)| \leq 1.5|OPT|$ as $|MST(S) \setminus \{h\}| \leq |OPT|$ and $|h| \leq |OPT|/2$.

Using our approximate coloring, one can compute $MST(B)$ and $MST(R)$, each with weight at most $\alpha|MST(S)|$. Therefore $|MST(R) \cup MST(B)| \leq 2\alpha|MST(S)| \leq 3\alpha|OPT|$, where $\alpha$ is again the Steiner ratio of the metric space.

Using the above case analysis, we have the following theorem.

▶ **Theorem 2.** *There exists a $3\alpha$-approximation for the Min-Sum 2-MST problem.*

▶ Remark. The Steiner ratio is the supremum of the ratio of length of an minimum spanning tree and a minimum Steiner tree over a point set. In a general metric space $\alpha = 2$ and in the Euclidean plane $\alpha \leq 1.3546$ [14].

## 2.2 Min-max

In this section the objective is to $\min \max\{|MST(R)|, |MST(B)|\}$.

▶ **Theorem 3.** *The Min-Max 2-MST problem is strongly NP-hard in general metric spaces.*

**Proof.** The reduction is from a problem which we will call *connected partition* [11]. In *connected partition* one is given a graph $G = (V, E)$, where $|V| = n$, and asked if it is possible to remove a set of edges from $G$ which breaks it into two connected components each of size $n/2$.

Given an instance of *connected partition*, $G = (V, E)$, we will create an instance of min-max 2-MST as follows. For each vertex $v_i \in V$ create an input pair $\{p_i, q_i\}$. For each edge $e = (v_i, v_j) \in E$ set the distance between the corresponding points, $p_i$ and $p_j$ to be one. Set the distances $d(q_i, q_j)$ to be zero for all $i, j$, and the distances $d(p_i, q_j)$ to be two for all $i, j$. In order to complete the construction, set all remaining distances to be the shortest path length among the distances defined above.

*Claim:* $G$ can be partitioned into two connected components of size $n/2$ if and only if there is a solution to the corresponding instance of min-max 2-MST with value $n/2 + 1$.

To show the first direction, suppose that the graph $G$ can be split into two connected components, $C_1, C_2$, of size exactly $n/2$. Without loss of generality suppose $\{v_i : 1 \le i \le n/2\} \in C_1$ and $\{v_j : n/2 < j \le n\} \in C_2$. Then, it is easy to verify based on the pairwise distances in the metric space described above that the coloring $\{p_i : 1 \le i \le n/2\} \cup \{q_j : n/2 < j \le n\} \in R$, $\{p_i : n/2 < i \le n\} \cup \{q_j : 1 \le j \le n/2\} \in B$, achieves a cost of $n/2 + 1$.

To show the opposite direction, suppose that there is a solution to the instance of min-max 2-MST of cost $n/2 + 1$. Notice that the minimum distance from point $p_i$ to any other point is at least one; therefore, there can be at most $n/2 + 2$ points from the set $P = \{p_i : 1 \le i \le n\}$ colored either red or blue in the solution which achieves this cost. Thus there are at least $n/2 - 2$ points from the set $Q = \{q_j : 1 \le j \le n\}$ colored either red or blue in this solution in order for it to be a feasible coloring. This implies that there will be at least one edge crossing the cut $(P, Q)$ in both the red and blue MST which realize the cost of this solution, and this edge has cost two. Then, of the remaining budget of $n/2 - 1$ units in order to complete the trees which realize the cost of this solution, it must be the case that we can utilize $n/2 - 1$ edges of length one which interconnect exactly $n/2$ nodes from the set $P$ in each color class.

The edges of length one in our metric space correspond directly to original edges of the graph $G$ in *connected partition* thus showing that there exists two spanning trees each of which spans exactly $n/2$ nodes of $G$ and thus $G$ can be partitioned into two connected components of size exactly $n/2$. ◀

▶ **Theorem 4.** *There exists a $4\alpha$-approximation for the Min-Max 2-MST problem.*

**Proof.** We use the same algorithm as we did for the Min-Sum 2-MST problem. The approximation factor is dominated by case 2 in the Min-Sum 2-MST analysis. For the Min-Max objective function, we have that $\max\{|MST(B)|, |MST(R)|\} \le \alpha|MST(S)|$ and that $|MST(S)| \le 4|OPT|$. Thus, $\max\{|MST(B)|, |MST(R)|\} \le 4\alpha|OPT|$. ◀

## 2.3 Bottleneck

In this section the objective is to $\min \max\{|\lambda(R)|, |\lambda(B)|\}$.

▶ **Lemma 5.** *Given $n$ pairs of points on a line in $\mathbb{R}^2$ where consecutive points on the line are unit separated, there exists a feasible coloring of the points, such that* $\max\{|\lambda(R)|, |\lambda(B)|\} \le 3$.

**Proof.** The proof will be constructive, using Algorithm 1. We partition the points into $n$ disjoint buckets, where a bucket consists of two consecutive points on the line.

Observe that at the end of Algorithm 1, each bucket has exactly one red point and one blue point by construction. Thus, the maximum distance between any two points of the same color is 3. ◀

**Algorithm 1:** Coloring points on a line.

```
Color the leftmost point, p, red
Let p′ be the point that is in p's bucket
Let R be a set of red points and B be a set of blue points
R ← {p}; B ← ∅
while There exists an uncolored point do
    while p′ is uncolored do
        if p is red then
            Color p's pair, q, blue
            B ← B ∪ {q}
            p ← q
        else
            Let p″ be the point in p's bucket
            Color p″ red
            R ← R ∪ {p″}
            p ← p″
        end
    end
    Find the leftmost uncolored point x and color it red. Let x′ be the point in x's
      bucket
    p ← x; p′ ← x′
end
return {R, B}
```

▶ **Theorem 6.** *There exists a 3-approximation algorithm for the Bottleneck 2-MST problem on a line.*

**Proof.** Note that if the leftmost $n$ points do not contain two points from the same pair, then it is optimal to let $R$ be the leftmost $n$ points and $B$ be the rightmost $n$ points. Suppose now that the leftmost $n$ points contain two points from the same pair. We run Algorithm 1 on the input. Imagine building two bottleneck spanning trees over the approximate coloring as well as over an optimal coloring. Let $\lambda$ be the longest edge (between two points of the same color) in our solution and $\lambda^*$ be the longest edge in the optimal solution.

Consider any two consecutive input points $s_i$ and $s_{i+1}$ on the line. We first show that $|\lambda^*| \geq |s_i s_{i+1}|$ by arguing that the optimal solution must have an edge that covers the interval $[s_i, s_{i+1}]$. Suppose to the contrary that no such edge exists. This means that $s_i$ is connected to $n-1$ points only to its left and $s_{i+1}$ is connected to $n-1$ points only to its right. This contradicts the assumption that the leftmost $n$ points contain two points from the same pair.

Let the longest edge in our solution be defined by two points, $p_i$ and $p_j$. Consider the number of input points in interval $[p_i, p_j]$. Input points in this interval other than $p_i$ and $p_j$ will have a different color than $p_i$ and $p_j$. It is easy to see that if $[p_i, p_j]$ consists of two input points, that $|\lambda^*| = |\lambda|$, and if $[p_i, p_j]$ consists of three input points, that $|\lambda^*| \geq |\lambda/2|$. We know by lemma 5 that $[p_i, p_j]$ can consist of no more than four input points. In this last case, $|\lambda^*|$ must be at least the length of the longest edge of the three edges in $[p_i, p_j]$. Thus, we see that $|\lambda^*| \geq |\lambda|/3$. ◀

▶ **Theorem 7.** *There exists a 9-approximation algorithm for the Bottleneck 2-MST problem in a metric space.*

**Proof.** First, we compute $MST(S)$ and consider the heaviest edge, $h$. The removal of this edge separates the nodes into two connected components, $H_1$ and $H_2$. If $\nexists i : p_i, q_i \in H_j$ for $1 \leq i \leq n$ and $1 \leq j \leq 2$, then we let $R = H_1$ and $B = H_2$ and return $R$ and $B$. Let $\lambda^*$ be the heaviest edge in the bottleneck spanning trees built on an optimal coloring. Note that $MST(S)$ lexicographically minimizes the weight of the $k$th heaviest edge, $1 \leq k \leq 2n-1$, among all spanning trees over $S$, and thus the weight of the heaviest edge in $MST(S) \setminus \{h\}$ is a lower bound on $|\lambda^*|$. Thus, in this case, our solution is clearly feasible and is also optimal as $MST(R)$ and $MST(B)$ are subsets of $MST(S) \setminus \{h\}$.

Now suppose $\exists\, j \in \{1,2\} : p_i, q_i \in H_j, 1 \leq i \leq n$. This means that $|\lambda^*| \geq |h|$. In this case, we compute a bottleneck TSP tour on the entire point set. It is known that that a bottleneck TSP tour with bottleneck edge $\lambda$ can be computed from $MST(S)$ so that $|\lambda| \leq 3|h| \leq 3|\lambda^*|$ [9].

Next we run Algorithm 1 on the TSP tour and return two paths, each having the property that the largest edge has weight no larger than $9|\lambda^*|$. ◀

▶ Remark. Consider the problem of computing a feasible partition which minimizes the bottleneck edge across two bottleneck TSP tours. Let the heaviest edge in the bottleneck TSP tours built on the optimal partition be $\lambda^{**}$. The above algorithm gives a 9-approximation to this problem as well because the algorithm returns two Hamilton paths and we know that (using the notation in the above proof) $|\lambda^*| \leq |\lambda^{**}|$. Thus, $|\lambda| \leq 9|\lambda^*| \leq 9|\lambda^{**}|$.

The following is a generalization of Lemma 5. Let $\mathcal{S} = \{S_1, S_2, \ldots, S_n\}$ be a set of $n$ $k$-tuples of points on a line. Each set $S_i$, $1 \leq i \leq n$, must be colored with $k$ colors. That is, no two points in set $S_i$ can be of the same color.

Consider two consecutive points of the same color, $p$ and $q$. We show that there exists a polynomial time algorithm that colors the points in $S$ so that the number of input points in interval $(p,q)$ is $O(k)$.

▶ **Lemma 8.** *There exists a polynomial time algorithm to color $S$ so that for any two consecutive input points of the same color, $p$ and $q$, the interval $(p,q)$ contains at most $2k-2$ input points.*

**Proof.** The algorithm consists of $k$ steps, where in the $j$th step, we color $n$ of the yet uncolored points with color $j$. We describe the first step.

Divide the $kn$ points into $n$ disjoint buckets, each of size $k$, where the first bucket $B_1$ consists of the $k$ leftmost points, the second bucket $B_2$ consists of the points in places $k+1, k+2, \ldots, 2k$, etc. Let $G = (V, E)$ be the bipartite graph, with node set $V = \{\mathcal{S} \cup B = \{B_1, \ldots, B_n\}\}$, in which there is an edge between $B_i$ and $S_j$ if and only if at least one of $S_j$'s points lies in bucket $B_i$. According to Hall's theorem [13], there exists a perfect matching in $G$. Let $M$ be such a matching and for each edge $e = (B_i, S_j)$ in $M$, color one of the points in $B_i \cap S_j$ with color 1. Now, remove from each tuple the point that was colored 1, and remove from each bucket the point that was colored 1. In the second step we color a single point in each bucket with the color 2, by again computing a perfect matching between the buckets (now of size $k-1$) and the $(k-1)$-tuples. It is now easy to see that for any two consecutive points of the same color, $p$ and $q$, at most $2k-2$ points exist in interval $(p,q)$. ◀

## 3 Matchings

Let $M(X)$ be the minimum weight matching on point set $X$ and $|M(X)|$ be the cost of the matching. Let $\lambda(X)$ be the longest edge in a bottleneck matching on point set $X$ and $|\lambda(X)|$ be the cost of that edge edge. Given $n$ pairs of points in a metric space, find a feasible coloring which minimizes the cost of a pair of matchings, one built over each color class.

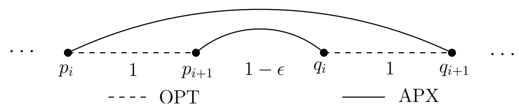


**Figure 1** $\frac{|APX|}{|OPT|} \approx 2$.

## 3.1 Minimum Sum

In this section the objective is to minimize $|M(R)| + |M(B)|$.

▶ **Theorem 9.** *There exists a 2-approximation for the Min-Sum 2-Matching problem in general metric spaces.*

**Proof.** First, note that the weight of the minimum weight perfect matching on $S$, $M^*$, which forbids edges $(p_i, q_i)$ for all $i$ is a lower bound on $|OPT|$. Next, we define the minimum weight one of a pair matching, $\hat{M}$, to be a minimum weight perfect matching which uses exactly one point from each input pair $\{p_i, q_i\}$ . Observe that $|\hat{M}|$, is a lower bound on the weight of the smaller of the matchings of OPT and therefore has weight at most $|OPT|/2$.

Our algorithm is to compute $\hat{M}$, and color the points involved in this matching red, and the remainder blue. We return the coloring $R \cup B$ as our approximate solution.

We have that $|M(R)| = |\hat{M}| \leq |OPT|/2$. To bound $|M(B)|$, consider the multigraph $G = (V = S, E = M^* \cup \hat{M})$. All $v \in B$ have degree 1 (from $M^*$), and all $u \in R$ have degree 2 (from $M^*$ and $\hat{M}$). For each $v_i \in B$, either $v_i$ is matched to $v_j \in B$ by $M^*$, or $v_i$ is matched to $u_i \in R$ by $M^*$. In the former case we can consider $v_i$ and $v_j$ matched in $B$ and charge the weight of this edge to $|M^*|$. In the latter case, note that each $u \in R$ is part of a unique cycle, or a unique path. If $u \in R$ is part of a cycle then no vertex in that cycle belongs to $B$ due to the degree constraint. Thus, if $v_i \in B$ is matched to $u_i \in R$, $u_i$ is part of a unique path whose other terminal vertex $x$ belongs to $B$, due to the degree constraint. We can consider $v_i$, and $x$ matched and charge the weight of this edge to the unique path connecting $v_i$ and $x$ in $G$. Thus, $|M(B)|$ can be charged to $|M^* \cup \hat{M}|$ and has weight at most $1.5|OPT|$.

Therefore, our partition guarantees $|M(R)| + |M(B)| \leq 2|OPT|$. Figure 1 shows the approximation factor using our algorithm is tight. ◀

## 3.2 Min-max

In this section the objective is to $\min \max\{|M(R)|, |M(B)|\}$.

▶ **Theorem 10.** *The Min-Max 2-Matching problem is weakly NP-hard in the Euclidean plane.* [See full paper [2] for proof.]

▶ Remark. The reduction used can be easily modified to also show that the *Min-Max 2-MST problem* is weakly NP-hard in the Euclidean plane.

▶ **Theorem 11.** *The approximation algorithm for the Min-Sum 2-Matching problem serves as a 3-approximation for the Min-Max 2-Matching problem in general metric spaces.* [See full paper [2] for proof.]

## 3.3 Bottleneck

In this section the objective is to $\min \max\{|\lambda(R)|, |\lambda(B)|\}$.

**Algorithm 2:** Algorithm $A(\mu, \beta)$. $0 < \mu < 1$ and $\beta > 1$.

Let $\overline{TSP_\beta}(X)$ denote a $\beta$-factor approximate TSP tour on set $X$.

1. Compute $\overline{TSP_\beta}(S)$.
2. Let $2k$ be the largest even number not exceeding $(2+\frac{1}{\mu})\beta$. Enumerate all ways of decomposing $\overline{TSP_\beta}(S)$ into $2k$ connected components: for each decomposition, color the nodes from consecutive components red and blue alternately (i.e. color all nodes in component one red, all nodes in component two blue, etc.). If this coloring is infeasible, then skip to the next decomposition; otherwise compute $\overline{TSP_\beta}(R)$ and $\overline{TSP_\beta}(B)$.
3. Compute a random feasible coloring, $S = R \cup B$, and compute $\overline{TSP_\beta}(R)$ and $\overline{TSP_\beta}(B)$.
4. Among all pairs of tours produced in steps 2 and 3, choose the pair of minimum sum.

▶ **Theorem 12.** *There exists a 3-approximation to the Bottleneck 2-Matching problem in general metric spaces.* [See full paper [2] for proof]

## 4 TSP Tours

Let $TSP(X)$ be a TSP tour on point set $X$ and $|TSP(X)|$ be the cost of the tour. Let $\lambda(X)$ be the longest edge in a bottleneck TSP tour on point set $X$ and $|\lambda(X)|$ be the cost of that longest edge. Given $n$ pairs of points in a metric space, find a feasible coloring which minimizes the cost of a pair of TSP tours, one built over each color class.

It is interesting to note the complexity difference emerging here. In prior sections, the structures to be computed on each color class of a feasible coloring were computable exactly in polynomial time. Thus, the decision versions of these problems, which ask if there exists a feasible coloring such that some cost function over the pair of structures is at most $k$, are easily seen to be in NP. However, when the cost function is over a set of TSP tours or bottleneck TSP tours, this is no longer the case. That is, suppose that a non-deterministic Touring machine could in polynomial time, for a point set $S$ and $k \in \mathbb{R}$, return a coloring for which it claimed the cost of the TSP tours generated over both color classes is at most $k$. Unless $P = NP$, the verifier cannot in polynomial time confirm that this is a valid solution, and therefore the problem is not in NP. Thus, the problems considered in this section are all NP-hard.

### 4.1 Minimum Sum

In this section the objective is to minimize $|TSP(R)| + |TSP(B)|$.

We will show for $\beta > 1$ and for the proper choice of $\mu$, that Algorithm 2 gives a $3\beta$-approximation for the Min-Sum 2-TSP problem. Fix a constant $\mu < 1$. Let $OPT$ be the optimal (feasible) coloring $S = R^* \cup B^*$. Let $d(R, B)$ be the minimum point-wise distance between sets $R$ and $B$. We call an instance of the problem $\mu$-separable if there exists a feasible coloring $S = R \cup B : d(R, B) \geq \mu(|TSP(R)| + |TSP(B)|)$.

Let $APX$ be the coloring returned by our algorithm. We will show that if $S$ is not $\mu$-separable, then $|APX| \leq \frac{2}{1-4\mu}\beta|OPT|$ (see Lemma 13) and that if $S$ is $\mu$-separable, then $|APX| \leq \frac{1}{4\mu}\beta|OPT|$ (see Lemma 14). Supposing both of these are true, then the approximation factor of our algorithm is $\max\{\frac{1}{4\mu}, \frac{2}{1-4\mu}\}\beta$. One can easily verify that $\mu =$

1/12 is the minimizer which gives the desired $3\beta$ factor. The following lemma states that if $S$ is not $\mu$-separable, then any feasible coloring yields a "good" approximation.

▶ **Lemma 13.** *If $S$ is not $\mu$-separable, then $|APX| \le \frac{2}{1-4\mu}\beta|OPT|$.*

**Proof.** If $S$ is not $\mu$-separable, then for any feasible coloring $S = R \cup B$ we have $d(R, B) \le \mu(|TSP(R)| + |TSP(B)|)$. In particular, for the coloring induced by the optimal solution, $S = R^* \cup B^*$, $d(R^*, B^*) \le \mu(|TSP(R^*)| + |TSP(B^*)|)$. Then,

$$\begin{aligned}|TSP(S)| &\le |OPT| + 2d(R^*, B^*)\\ &\le |OPT| + 2\mu(|TSP(R^*)| + |TSP(B^*)|)\\ &\le |OPT| + 4\mu|TSP(S)|.\end{aligned}$$

Hence, when $\mu < \frac{1}{4}$, $|TSP(S)| \le \frac{1}{1-4\mu}|OPT|$. Let $S = \hat{R} \cup \hat{B}$ be the random feasible coloring computed by $\mathcal{A}(\mu, \beta)$. Then, as we are returning the best coloring between $\hat{R} \cup \hat{B}$ and all $O(n^{2k})$ colorings of $\overline{TSP_\beta}(S)$, we have $|APX| \le \beta(|TSP(\hat{R})| + |TSP(\hat{B})|) \le 2\beta|TSP(\mathcal{S})| \le \frac{2\beta}{1-4\mu}|OPT|$. ◀

The following lemma states that if $S$ is $\mu$-separable, then any witness coloring to the $\mu$-separability of $S$ gives a "good" approximation.

▶ **Lemma 14.** *If $S$ is $\mu$-separable, then $|APX| \le \frac{1}{4\mu}\beta|OPT|$.*

**Proof.** Suppose we successfully guessed a coloring $X_0 = R^0 \cup B^0$ that is a "witness" to the $\mu$-separability of $S$ (we will show how to guess $X_0$ later).

- Case 1: $OPT = X_0$. Then $|APX| \le \beta(|TSP(R^0)| + |TSP(B^0)|) = \beta(|TSP(R^*)| + |TSP(B^*)|) = \beta|OPT|$.
- Case 2: $OPT \ne X_0$. Then $R^* \ne R^0, B^* \ne B^0$ which means each tour in OPT must contain at least 2 edges crossing the cut $(R^0, B^0)$, hence the optimal solution must contain at least 4 edges crossing the cut $(R^0, B^0)$. So $|OPT| \ge 4d(R^0, B^0) \ge 4\mu(|TSP(R^0)| + |TSP(B^0)|) \ge \frac{4\mu}{\beta}|APX|$. Equivalently, $|APX| \le \frac{\beta}{4\mu}|OPT|$. ◀

The next two lemmas show how to guess a witness coloring $X_0$ in polynomial time. First, we show that if $S$ is $\mu$-separable with a witness coloring $X_0$, then $\overline{TSP_\beta}(S)$ cannot cross the red/blue cut defined by this coloring "too many" times.

▶ **Lemma 15.** *Let $\overline{TSP_\beta}(S)$ be an $\beta$-factor approximation for $TSP(S)$. Also, suppose $S$ is $\mu$-separable with witness $X_0$. Then $\overline{TSP_\beta}(S)$ crosses the cut $(R^0, B^0)$ at most $(2 + \frac{1}{\mu})\beta$ times.*

**Proof.** One can construct a TSP tour for $S$ by adding two bridges to $TSP(R^0)$ and $TSP(B^0)$, thus we have $|TSP(S)| \le |TSP(R^0)| + |TSP(B^0)| + 2d(R^0, B^0) \le (2 + \frac{1}{\mu})d(R^0, B^0)$. Also, suppose $\overline{TSP_\beta}(S)$ crosses the cut $(R^0, B^0)$ $2k$ times. Then, $2kd(R^0, B^0) \le |\overline{TSP_\beta}(S)| \le \beta|TSP(S)|$. Combining the above two inequalities, we obtain $2k \le (2 + \frac{1}{\mu})\beta$. ◀

The next lemma completes our proof.

▶ **Lemma 16.** *Suppose $S$ is $\mu$-separable. Let $X_0$ be any coloring which serves as a "witness". Then, in step 2 of $\mathcal{A}(\mu, \beta)$, we will encounter $X_0$ at some stage.*

**Proof.** Given a nonnegative integer $k$ and a TSP tour $P$, define $\Pi(P,k)$ =$\{X\colon X$ is a feasible coloring and $P$ crosses $X$ at most $k$ times$\}$. By Lemma 15, we know $X_0 \in \Pi(\overline{TSP_\beta}(S), (2+\frac{1}{\mu})\beta)$. Since step 2 of $\mathcal{A}(\mu,\beta)$ is actually enumerating all colorings in $\Pi(\overline{TSP_\beta}(S), (2+\frac{1}{\mu})\beta)$, this completes the proof. ◀

Note that step 2 considers $O(n^{2k}) = O(n^{14\beta})$ decompositions and for each coloring that is feasible, we compute two approximate TSP tours. Suppose the running time to compute a $\beta$-factor TSP tour on $n$ points is $h_\beta(n)$. Then the worst case running time of Algorithm 2 is $O(h_\beta(2n)n^{14\beta})$. Thus, we have the following Theorem.

▶ **Theorem 17.** *For any $\beta > 1$, the algorithm $\mathcal{A}(\frac{1}{12},\beta)$ is a $3\beta$-approximation for the Min-Sum 2-TSP problem with running time $O(h_\beta(2n)n^{14\beta})$.*

▶ Remark. If $S$ is in the Euclidean plane then $\beta = 1+\epsilon$ for some $\epsilon > 0$ [16] yielding a $(3+\epsilon)$-approximation and if $S$ is in a general metric space then $\beta = 3/2$ [7] yielding a 4.5-approximation. In both cases $h_\beta(2n)$ is polynomial.

## 4.2 Min-Max

In this section the objective is to $\min\max\{|TSP(R)|, |TSP(B)|\}$.

▶ **Theorem 18.** *There exists a $6\beta$-approximation to the Min-Max 2-TSP problem, where $\beta$ is the approximation factor for TSP in a certain metric space.* [See full paper [2] for proof.]

## 4.3 Bottleneck

In this section the objective is to $\min\max\{|\lambda(R)|, |\lambda(B)|\}$.

▶ **Theorem 19.** *There exists an 18-approximation algorithm for the Bottleneck 2-TSP problem.* [See full paper [2] for proof.]

## References


1 Manuel Abellanas, Ferran Hurtado, Christian Icking, Rolf Klein, Elmar Langetepe, Lihong Ma, Belén Palop, and Vera Sacristán. Smallest color-spanning objects. In FriedhelmMeyer auf der Heide, editor, *Algorithms — ESA 2001*, volume 2161 of *Lecture Notes in Computer Science*, pages 278–289. Springer Berlin Heidelberg, 2001. doi:10.1007/3-540-44676-1_23.

2 Esther M Arkin, Aritra Banik, Paz Carmi, Gui Citovsky, Su Jia, Matthew J Katz, Tyler Mayer, and Joseph SB Mitchell. Network optimization on partitioned pairs of points. *arXiv, submission 2025570*, 2017. URL: https://arxiv.org/abs/1710.00876.

3 Esther M. Arkin, José M. Díaz-Báñez, Ferran Hurtado, Piyush Kumar, Joseph S. B. Mitchell, Belén Palop, Pablo Pérez-Lantero, Maria Saumell, and Rodrigo I. Silveira. Bichromatic 2-center of pairs of points. *Comput. Geom.*, 48(2):94–107, 2015. doi:10.1016/j.comgeo.2014.08.004.

4 Esther M. Arkin and Refael Hassin. Minimum-diameter covering problems. *Networks*, 36(3):147–155, 2000.

5 Luis Barba, Stephane Durocher, Robert Fraser, Ferran Hurtado, Saeed Mehrabi, Debajyoti Mondal, Jason Morrison, Matthew Skala, and Mohammad Abdul Wahid. On k-enclosing objects in a coloured point set. In *Proceedings of the 25th Canadian Conference on Computational Geometry, CCCG 2013, Waterloo, Ontario, Canada, August 8-10, 2013*. Carleton

University, Ottawa, Canada, 2013. URL: http://cccg.ca/proceedings/2013/papers/paper_35.pdf.

6 Binay K. Bhattacharya, Ante Custic, Akbar Rafiey, Arash Rafiey, and Vladyslav Sokol. Approximation algorithms for generalized MST and TSP in grid clusters. In Zaixin Lu, Donghyun Kim, Weili Wu, Wei Li, and Ding-Zhu Du, editors, *Combinatorial Optimization and Applications - 9th International Conference, COCOA 2015, Houston, TX, USA, December 18-20, 2015, Proceedings*, volume 9486 of *Lecture Notes in Computer Science*, pages 110–125. Springer, 2015. doi:10.1007/978-3-319-26626-8_9.

7 Nicos Christofides. Worst-case analysis of a new heuristic for the travelling salesman problem. Technical report, DTIC Document, 1976.

8 Mario E. Consuegra and Giri Narasimhan. Geometric avatar problems. In Anil Seth and Nisheeth K. Vishnoi, editors, *IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS 2013, December 12-14, 2013, Guwahati, India*, volume 24 of *LIPIcs*, pages 389–400. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2013. doi:10.4230/LIPIcs.FSTTCS.2013.389.

9 Thomas H Cormen, Charles E Leiserson, Ronald L Rivest, and Clifford Stein. Introduction to algorithms second edition, 2001.

10 Sandip Das, Partha P. Goswami, and Subhas C. Nandy. Smallest color-spanning object revisited. *International Journal of Computational Geometry and Applications*, 19(05):457–478, 2009. doi:10.1142/S0218195909003076.

11 Martin E Dyer and Alan M Frieze. On the complexity of partitioning graphs into connected subgraphs. *Discrete Applied Mathematics*, 10(2):139–153, 1985.

12 Harold N. Gabow, Shachindra N. Maheshwari, and Leon J. Osterweil. On two problems in the generation of program test paths. *IEEE Transactions on Software Engineering*, 2(3):227–231, 1976.

13 Philip Hall. On representatives of subsets. *J. London Math. Soc*, 10(1):26–30, 1935.

14 Dan Ismailescu and Joseph Park. Improved upper bounds for the steiner ratio. *Discrete Optimization*, 11:22–30, 2014. doi:10.1016/j.disopt.2013.10.004.

15 Payam Khanteimouri, Ali Mohades, MohammadAli Abam, and MohammadReza Kazemi. Computing the smallest color-spanning axis-parallel square. In Leizhen Cai, Siu-Wing Cheng, and Tak-Wah Lam, editors, *Algorithms and Computation*, volume 8283 of *Lecture Notes in Computer Science*, pages 634–643. Springer Berlin Heidelberg, 2013. doi:10.1007/978-3-642-45030-3_59.

16 Joseph SB Mitchell. Guillotine subdivisions approximate polygonal subdivisions: A simple polynomial-time approximation scheme for geometric tsp, k-mst, and related problems. *SIAM Journal on Computing*, 28(4):1298–1309, 1999.

17 Young-Soo Myung, Chang-ho Lee, and Dong-wan Tcha. On the generalized minimum spanning tree problem. *Networks*, 26(4):231–241, 1995. doi:10.1002/net.3230260407.

18 Petrica C. Pop. New models of the generalized minimum spanning tree problem. *J. Math. Model. Algorithms*, 3(2):153–166, 2004. doi:10.1023/B:JMMA.0000036579.83218.8d.

19 Petrica C. Pop, Walter Kern, Georg Still, and Ulrich Faigle. Relaxation methods for the generalized minimum spanning tree problem. *Electronic Notes in Discrete Mathematics*, 8:76–79, 2001. doi:10.1016/S1571-0653(05)80085-X.

20 Petr Slavik. *Approximation algorithms for set cover and related problems.* PhD thesis, State University of New York at Buffalo, Buffalo, NY, USA, 1998.