Mathematical Programming 80 (1998) 171–194

A maximum b-matching problem arising from median location models with applications to the roommates problem

Arie Tamir [a,*], Joseph S.B. Mitchell [b]

[a] *Department of Statistics and Operations Research, School of Mathematical Sciences, Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 69978, Israel*

[b] *Department of Applied Mathematics and Statistics, State University of New York at Stony Brook, Stony Brook, NY 11777-3600, USA*

Received 1 May 1994; revised manuscript received 23 December 1996

Abstract

We consider maximum b-matching problems where the nodes of the graph represent points in a metric space, and the weight of an edge is the distance between the respective pair of points. We show that if the space is either the rectilinear plane, or the metric space induced by a tree network, then the b-matching problem is the dual of the (single) median location problem with respect to the given set of points. This result does not hold for the Euclidean plane. However, we show that in this case the b-matching problem is the dual of a median location problem with respect to the given set of points, in some extended metric space. We then extend this latter result to any geodesic metric in the plane. The above results imply that the respective fractional b-matching problems have integer optimal solutions. We use these duality results to prove the nonemptiness of the core of a cooperative game defined on the roommate problem corresponding to the above matching model. © 1998 The Mathematical Programming Society, Inc. Published by Elsevier Science B.V.

Keywords: The b-matching problem; Median location problems; The roommates problem

1. Introduction

Let $G = (V, E)$ be an undirected graph with a node set V and an edge set E. A *matching* in G is a subset of E such that each node of G is met by at most one edge

* Corresponding author.

0025-5610/98/$19.00
Published by Elsevier Science B.V.
PII S0025-5610(97)00032-4

in the subset. In 1965 Edmonds discovered a strongly polynomial algorithm for the weighted matching problem, as well as a characterization of a linear system that defines the convex hull of the incidence vectors of all the matchings of a graph. The concept of a matching was generalized by Edmonds to b-matchings, defined as follows: For each node $v \in V$ let $\delta(v)$ denote the set of edges in E which meet v. For each vector $x = (x_e : e \in E)$ and a subset E' of E, let $x(E') = \sum\{x_e : e \in E'\}$.

Let $b = (b_v : v \in V)$ be a positive integer vector. A *b-matching* of G is a nonnegative integer vector $x = (x_e : e \in E)$, satisfying the following degree constraints,

$$x(\delta(v)) \leqslant b_v \quad \forall v \in V. \tag{1}$$

(A b-matching of G with $b_v = 1$ for all $v \in V$ is a *matching*.) A b-matching is *perfect* if all the degree constraints are satisfied as equalities. Given a vector of weights $c = (c_e : e \in E)$, the weighted b-matching problem is

$$\begin{aligned} \text{Maximize} \quad & \sum_{e \in E} c_e x_e \\ \text{s.t.} \quad & x(\delta(v)) \leqslant b_v \quad \forall v \in V, \\ & x \geqslant 0 \quad \text{and integer.} \end{aligned} \tag{2}$$

The linear relaxation of (2), called the *fractional b-matching problem*, is

$$\begin{aligned} \text{Maximize} \quad & \sum_{e \in E} c_e x_e \\ \text{s.t.} \quad & x(\delta(v)) \leqslant b_v \quad \forall v \in V, \\ & x \geqslant 0. \end{aligned} \tag{3}$$

A characterization of a linear system of the convex hull of the b-matchings of a graph was given by Edmonds and Pulleyblank and appeared in [23]. Pulleyblank has also provided a pseudopolynomial algorithm to find a maximum weighted b-matching. Cunningham and Marsh [5], found the first polynomial time algorithm for finding an optimal b-matching, and several years later Anstee [2], presented the first strongly polynomial algorithm for this problem. The reader is referred to [20] and [4] for additional results on general matchings, b-matchings and capacitated b-matchings.

In this paper we study a special case of (2), where the nodes of the graph represent points in a metric space, and the weight of an edge is the distance between the respective pair of points. We first consider, in Section 2.1, the case in which the metric space is induced by a tree network. Specifically, let $T = (V, E')$ be an undirected tree with positive edge lengths. For each pair of nodes $v, u \in V$ let $d(v, u)$ denote the length of the unique simple path on T connecting v and u. We consider the special instance of the weighted b-matching problem on the complete graph $G = (V, E)$, where for each edge $e = (v, u) \in E$, $c_e = d(v, u)$. (Note that G and T have the same node set.) We prove that for this case the fractional b-matching problem (3) has an integer optimal solution

where the number of positive components is at most $|V|$. The proof is based on a duality argument, showing that the dual of (3) is the problem of locating a weighted median on the given tree T. In Section 3 we extend the duality result to a restricted weighted median problem.

We demonstrate that the above integrality property might not hold for the capacitated b-matching version of (2), where positive upper bounds are imposed on the x variables, even if the tree T is a simple path.

In Section 2.2 we consider the case where the nodes of the graph represent points in the rectilinear plane. Assuming that $\sum\{b_v : v \in V\}$ is even, we prove that the above integrality result holds for this case as well. Again, the proof is based on showing that the dual of the linear relaxation of (2) is the problem of finding a weighted median of the given set of points in the rectilinear plane.

In Section 2.3 we consider the Euclidean plane. The linear relaxation of (2) is not the dual of the problem of finding the weighted median of the given points in the Euclidean plane. However, using a different argument, we prove that if $\sum\{b_v : v \in V\}$ is even the linear relaxation of (2) has the integrality property for this case as well. This integrality property implies that the b-matching problem for the Euclidean plane is the dual of a weighted median problem in some metric space extension of the Euclidean space. We then extend these results to general geodesic metrics in the plane (Section 2.4). In Section 2.5 we give counterexamples to the duality and integrality properties in three dimensions.

In Section 3, we consider an extended b-matching problem, for a metric induced by a tree. Finally, in Section 4, we apply our results to the roommates model introduced by Gale and Shapley [7]. In the roommates problem there is a group of an even number of boys who wish to divide up into pairs of roommates. Each boy ranks all other boys in accordance with his preference for a roommate. A matching of the group of boys to pairs is said to be *stable* if under it there does not exist a pair of boys who would prefer to room with each other rather than with their assigned roommates. Using a simple example, Gale and Shapley demonstrated that, in general, the roommates problem may have no stable matching. The reader is referred to [11] for recent structural and algorithmic results on this problem. Bartholdi and Trick [3] studied the special case where preferences are single peaked and narcissistic, and proved the existence of a unique stable matching. (Roughly speaking this corresponds to the case where there is a one-dimensional frame of reference, e.g., temperature scale or highway distance map, that is common to all boys, and alternatives are preferred less as they are "farther" from each boy's ideal.) In contrast, we consider the case where alternatives are preferred more as they are farther from each boy's location. We discuss a cooperative game with side payments based on this roommates problem, and show that the core is nonempty for the metric spaces discussed in Section 2. This is to be compared with the result of Granot [10], who showed that, in general, the core of the roommates problem cooperative game with side payments may be empty.

2. Maximum b-matchings and median location

2.1. The tree metric space problem

Given is an undirected tree $T = (V, E')$ with positive edge lengths. For each pair of nodes $v, u \in V$ we let $d(v,u)$ be the length of the unique simple path connecting v and u on T. $d(v,u)$ is called the *distance* between v and u. Let $G = (V, E)$ denote the complete undirected graph having V as its node set. Consider the maximum b-matching problem (2), and assume that for each edge $e = (v,u) \in E$, $c_e = d(v,u)$. The linear programming dual of (3) is

$$\begin{aligned} \text{Minimize} \quad & \sum_{v \in V} b_v z_v \\ \text{s.t.} \quad & z_v + z_u \geqslant d(v,u) \quad \text{for all pairs } (v,u), \\ & z \geqslant 0. \end{aligned} \tag{4}$$

We now interpret the dual problem as a median location problem on the given tree $T = (V, E')$. Suppose that the edges of the tree are rectifiable, and the tree is embedded in the plane. Each edge is a line segment of the appropriate length. We refer to interior points on an edge by their distance (on the line segment) from the two nodes of the edge. The continuum set of all points of these segments is a metric space that we denote by $A(T)$. Suppose that each node v of the tree is associated with a nonnegative weight b_v, and consider the problem of finding a point in $A(T)$ that will minimize the sum of weighted distances from the point to all nodes. It is well known that if b_v is a positive real for all $v \in V$, then the set of optimal points (known as *weighted medians* of T), is either a single node or the set of points on a single edge of T. In particular, there is always one weighted median which is a node. The next theorem follows from the general results in [6] on monotonic tree location problems. It states that (4) is a reformulation of the above weighted median problem.

Theorem 1. *Suppose that b_v is nonnegative for each $v \in V$. Let $z^* = (z_v^* : v \in V)$ be an optimal solution to problem (4). Then there exists a weighted median of T, $v^* \in V$, with respect to the weights $\{b_v\}$, such that $z_v^* = d(v, v^*)$ for all $v \in V$.*

We will need the following characterization of weighted medians due to Sabidussi [24], Zelinka [26], Kang and Ault [13] and Kariv and Hakimi [14].

Theorem 2. *Let v^* be a node of T. Then v^* is a weighted median with respect to the nonnegative weights $\{b_v\}$ if and only if the sum of the weights of the nodes in each connected component of T, obtained by the removal of v^*, does not exceed $(1/2)\sum\{b_v : v \in V\}$.*

We are now ready to prove the main result of this section.

Theorem 3. *Suppose that b_v is a positive integer for each $v \in V$. The linear program (3) has an integer optimal solution, x^*, where the number of its positive components is at most $|V|$.*

Proof. It is sufficient to show that there is an integer feasible solution to (3) with at most $|V|$ positive components, whose objective value is equal to the optimal value of the linear program (4). Let v^* be a weighted median of T. Consider the connected components of the tree obtained by the deletion of v^*. From Theorem 2, the total node weights in each component does not exceed $B/2$, where $B = \sum\{b_v : v \in V\}$. By aggregating components we can obtain a partition of the node subset $V - \{v^*\}$ into three subsets $\{V^1, V^2, V^3\}$ satisfying the following properties:

(i) $B^j = \sum\{b_v : v \in V^j\} \leqslant B/2$, for $j = 1, 2, 3$.

(ii) Every simple path connecting a pair of nodes v, u, where v is in some V^j and u is in $V - V^j$ must contain the median v^*.

We will now construct a b-matching x^*, of the complete graph $G = (V, E)$, with the following two properties:

(iii) No pair of nodes belonging to the same subset V^j are matched together.

(iv) For each node $v \neq v^*$, $\sum\{x_e^* : e \in \delta(v)\} = b_v$.

Therefore, if the edge e connects the pair of nodes v and $u \in G$, then the contribution of the variable x_e^* to the objective is $x_e^* d(v,u) = x_e^* \big(d(v,v^*) + d(v^*,u)\big)$. Moreover, from (iv) it follows that the objective value of any b-matching satisfying (iii) and (iv) is $\sum\{b_v d(v,v^*) : v \in V\}$, which is exactly the objective value of the weighted median problem (4). To construct the above b-matching suppose without loss of generality that $B^1 \geqslant B^2 \geqslant B^3$. For convenience denote $V^4 = \{v^*\}$ and $B^4 = b_{v^*}$. Next we replicate each node v, and replace it by b_v copies (*quasi-nodes*). Thus, we assume that V^j contains B^j quasi-nodes, $j = 1, 2, 3, 4$. By saying that we match p pairs from $[V^i, V^j]$, we will mean that a subset of p quasi-nodes from V^i are paired with a subset of p quasi-nodes from V^j.

Consider the following cases:

Case I: $B^1 > B^2$ and $B^1 \geqslant B^2 + B^3$.

Match	B^2	pairs from	$[V^1, V^2]$,
	B^3	pairs from	$[V^1, V^3]$ and
	$B^1 - B^2 - B^3$	pairs from	$[V^1, V^4]$.

Case II: $B^1 > B^2, B^1 < B^2 + B^3$ and $B^2 + B^3 - B^1$ is even.

Match	$(B^1 + B^2 - B^3)/2$	pairs from	$[V^1, V^2]$,
	$(B^1 - B^2 + B^3)/2$	pairs from	$[V^1, V^3]$ and
	$(-B^1 + B^2 + B^3)/2$	pairs from	$[V^2, V^3]$.

Case III: $B^1 > B^2, B^1 < B^2 + B^3$ and $B^2 + B^3 - B^1$ is odd.

Match	$(B^1 + B^2 - B^3 - 1)/2$	pairs from	$[V^1, V^2]$,
	$(B^1 - B^2 + B^3 - 1)/2$	pairs from	$[V^1, V^3]$,
	$(-B^1 + B^2 + B^3 + 1)/2$	pairs from	$[V^2, V^3]$ and
	1	pair from	$[V^1, V^4]$.

Case IV: $B^1 = B^2$ and B^3 is even.

Match	$B^1 - B^3/2$	pairs from	$[V^1, V^2]$,
	$B^3/2$	pairs from	$[V^1, V^3]$ and
	$B^3/2$	pairs from	$[V^2, V^3]$.

Case V: $B^1 = B^2$ and B^3 is odd.

Match	$B^1 - B^3/2 - 1/2$	pairs from	$[V^1, V^2]$,
	$B^3/2 - 1/2$	pairs from	$[V^1, V^3]$,
	$B^3/2 + 1/2$	pairs from	$[V^2, V^3]$ and
	1	pair from	$[V^1, V^4]$.

It is easy to check that each one of the five cases constructs a b-matching satisfying properties (iii) and (iv) above. For example consider Case I. Each matched pair involves a node from V^1. The only questionable matching is from $[V^1, V^4]$. However, this matching is feasible if $B^1 - B^2 - B^3$ is between 0 and B^4. Since we focus on Case I, the nonnegativity of $B^1 - B^2 - B^3$ is ensured. The second constraint is equivalent to $2B^1 \leqslant B^1 + B^2 + B^3 + B^4$, which follows from property (i) above. (Note that $B^1 + B^2 + B^3 + B^4 = B$.)

It remains to prove that in each one of the five cases the b-matching can be implemented using at most $|V|$ edges. Consider for example Case I. (Similar arguments can be applied to the other cases.) Let V_1 be a minimal subset of V^1 (with respect to containment), such that $\sum\{b_v : v \in V_1\} \geqslant B^2 + B^3$. Match $B^2 + B^3$ pairs from $[V_1, V^2 \cup V^3]$. Using a basic solution argument, (e.g., the classical North–West corner rule applied to the transportation problem), we conclude that this matching can be implemented using at most $|V_1| + |V^2| + |V^3| - 1$ edges. From the minimality of V_1, all nodes in V_1, but possibly one, say v'', are perfectly matched (i.e. they satisfy the equality in property (iv) above). Next we match $B^1 - B^2 - B^3$ from $[(V^1 - V_1) \cup \{v''\}, V^4]$. Again this matching can be implemented using at most $|V^1| - |V_1| + 1 + |V^4| - 1$ edges. To conclude $|V^1| + |V^2| + |V^3| = |V| - 1$ edges will suffice to implement the b-matching in Case I. As noted above similar arguments can be used for the other four cases. This completes the proof. □

We note that recently Gerstel and Zaks [8] considered the case where all edges are of unit length and $b_v = 1$ for each $v \in V$. They did not discuss the fractional matching

problem (3). However, they did prove that the optimal value of the respective integer program (2) is equal to the optimal value of the (unweighted) median problem. It is clear that if $b_v = 1$ for each $v \in V$, there is an optimal integer solution where the number of positive variables is at most $|V|/2$.

Theorem 3 and the duality of problems (3) and (4) imply that in our case an optimal integer b-matching can be found in $O(|V|)$ time. The complexity bound is dominated by the effort to find a weighted median of a tree. The latter task can be performed in linear time, (see [9,14]). For comparison purposes recall that the best complexity bound known for the general maximum weighted b-matching problem is super cubic in $|V|$ [2]. The bound is cubic when $b_v = 1$ for all $v \in V$ [19]. If, in addition, the edge weights $\{c_e\}$ equal to the Euclidean distances between points in the plane which represent the nodes of the graph, then some subcubic algorithms are known. Vaidya [25] presented such an algorithm for finding a minimum-weight geometric perfect matching. Marcotte and Suri [21] considered the case where the representing points are on the boundary of a simple polygon in the plane. They gave an $O(|V| \log^2 |V|)$ time algorithm for the minimum-weight perfect matching problem (which has been improved to $O(|V| \log |V|)$, by Hershberger and Suri [12]), and an $O(|V| \log |V|)$ time algorithm for the maximum-weight perfect matching problem. They also presented improved bounds for the case where the polygon is convex.

We point out another interesting property of our model. From Theorem 2 it follows that an optimal weighted median depends only on the topology of the tree and the node weights $\{b_v\}$ but not on the edge lengths. Due to the above duality we conclude that an optimal b-matching is also independent of the tree distances $\{d(v,u)\}$.

Remark 4. Theorem 3 shows that there is an optimal solution x^* to the b-matching problem considered above, where the number of edges e with a positive variable x_e^* is at most $|V|$. We note that the respective bound for general b-matching problems is higher than $|V|$. It is shown in [23] that there is an optimal solution to the general b-matching problem, where the subset of edges with a positive variable contains no even cycles (simple or nonsimple). Therefore, all cycles are odd and node disjoint. In particular, there are at most $|V|/3$ such cycles in the subgraph induced by the above subset of edges. By removing one edge from each cycle, this subgraph reduces to a forest. Thus, the total number of edges in the subgraph is at most $(4/3)|V| - 1$. Simple examples demonstrate that this bound is tight.

Remark 5. From the proof of Theorem 3 we observe that the strict positivity assumption on the $\{b_v\}$ coefficients can be weakened. Generally, it is not possible to replace the inequality with equality in the degree constraints since the problem with equality constraints might be infeasible. Indeed, in the optimal solution constructed in the above proof the degree of the median v^* is equal to zero in Cases II and IV, and it is equal to one in Cases III and V. (In Case I the median degree is guaranteed to be bounded above by b_{v^*}.) Thus, the only cases of the proof where the positivity is required are III and V.

Furthermore, even in these two cases we only need the positivity of b_{v^*}. Therefore, the positivity assumption can be replaced by the assumption that all the coefficients are nonnegative, and the coefficient of a weighted median is positive. Alternatively, it suffices to require that the sum of all the coefficients is even, since this will exclude cases III and V. The integrality result of Theorem 3 may not hold if neither of the two assumptions is satisfied. Consider the example of a 4-node tree, where one of the nodes, the center, is connected by an edge to each one of the other nodes, *leaves*. Suppose that each edge is of length one. Finally, set the b coefficient of each leaf to be equal to 1, and let the respective coefficient of the center be equal to 0. Generally, to find the best integral solution when the sum of the $\{b_v\}$ coefficients is odd and the coefficient of the weighted median is zero, we first subtract 1 from the positive coefficient of a node which is closest to the weighted median amongst all nodes with positive coefficients, and then solve the modified problem as above. We also note that if the given tree is a general path, then Theorem 3 holds even if the b coefficients are allowed to take on a zero value.

If we impose positive integer upper bounds on the x variables in problem (3), the integrality result may not hold even if the tree is a path.

Example 6. Consider the path defined by the following 5 points (nodes) on the real line: $v_1 = 0$, $v_2 = 2$, $v_3 = 3$, $v_4 = 4$ and $v_5 = 6$. Let $b_{v_1} = b_{v_5} = 3$, $b_{v_2} = b_{v_4} = 2$ and $b_{v_3} = 1$. Suppose that the ten x variables corresponding to the edges of the complete graph with the node set $V = \{v_1, v_2, v_3, v_4, v_5\}$ are all bounded above by 1. The best integer solution satisfying the constraints of (3) and the above upper bound yields the objective value 19. However, the following fractional solution, specified by its nonzero components, yields the value 20:

$$x_{(v_1,v_5)} = x_{(v_1,v_4)} = x_{(v_2,v_5)} = 1,$$

and

$$x_{(v_1,v_2)} = x_{(v_2,v_4)} = x_{(v_4,v_5)} = x_{(v_5,v_3)} = x_{(v_3,v_1)} = 1/2.$$

We note that Theorem 3 cannot be extended to general distance functions, where we replace the assumption that the numbers $\{d(v,u)\}$ represent tree distances by the weaker assumption that they only satisfy the triangle inequality. This is obvious when the number of nodes is odd and $b_v = 1$ for each v, since no perfect matching is feasible. (Consider a 3-node triangle with unit length edges.) More interestingly, assuming $b_v = 1$ for each $v \in V$, the result does not hold for the even case, for which a perfect matching is feasible, even when $|V| = 6$, as demonstrated by the next example.

Example 7. Consider the complete graph with $V = \{v_1, v_2, v_3, v_4, v_5, v_6\}$. Suppose that $b_{v_i} = 1$ for $i = 1, \ldots, 6$. Let $d(v_i, v_j)$, $i \neq j$, be equal to 2 if (i,j) is in

$\{(1,2),(1,3),(2,3)\}$ and be equal to 1 otherwise. The optimal objective value of problem (3) is 4.5, while the best integer solution satisfying the constraints of (3) yields the value 4.

2.2. The rectilinear planar b-matching problem

The duality result between the weighted median problem in the tree metric space $A(T)$ and the respective b-matching problem, does not hold for general metric spaces (see the next section). We now prove that it does hold for the rectilinear plane.

Let $V = \{v_1, \ldots, v_n\}$, $v_i = (v_i^1, v_i^2)$, $i = 1, \ldots, n$, be a collection of points in R^2. Let b_{v_i}, $i = 1, \ldots, n$, be a positive integer weight. The *rectilinear weighted median problem* in the plane is to find a point $x^* = (x^{*1}, x^{*2})$ which minimizes the function

$$\sum_{i=1,\ldots,n} b_{v_i}(|x^1 - v_i^1| + |x^2 - v_i^2|).$$

The rectilinear integer b-matching problem on the complete graph $G = (V, E)$ is the version of problem (2), where for each edge $e = (v_i, v_j)$,

$$c_e = d(v_i, v_j) = |v_i^1 - v_j^1| + |v_i^2 - v_j^2|. \tag{4a}$$

Theorem 8. *Consider the rectilinear weighted median problem in the plane and suppose that* $\sum\{b_{v_i} : v_i \in V\}$ *is even. Then the optimal solution value of this problem is equal to the optimal solution value of the respective rectilinear integer b-matching problem. Moreover, the former is equal to the optimal objective value of problem* (4), *while the latter is equal to the objective value of problem* (3), *where* $\{c_e\}$ *and* $\{d(v_i, v_j)\}$ *are defined above in* (4a).

Proof. Using continuity and perturbation arguments it is sufficient to prove the theorem for the case when the n points $\{v_1, \ldots, v_n\}$ are distinct, $b_{v_i} = 1$ for $i = 1, \ldots, n$, and no pair $\{v_i, v_j\}$, $i \neq j$, lie on a vertical or horizontal line in R^2. The supposition in the theorem implies that n is even. An optimal solution to the rectilinear problem is given by the point $x = (x^1, x^2)$, where x^1 is a median of the set $\{v_i^1\}$ and x^2 is a median of the set $\{v_i^2\}$. Due to the above assumptions the median point $x = (x^1, x^2)$ can be selected such that no point v_i lies on the vertical and horizontal lines passing through x. Consider the four quadrants defined by this pair of lines and let n_k, $k = 1, 2, 3, 4$, denote the number of points in the set $\{v_i\}$ that lie in the kth quadrant. We have $n_1 + n_2 = n_3 + n_4 = n/2$ and $n_2 + n_3 = n_1 + n_4 = n/2$. Thus, $n_1 = n_3$ and $n_2 = n_4$. Consider a pair of points v_i and v_j, where v_i is in the 1st quadrant (2nd quadrant) and v_j is in the 3rd quadrant (4th quadrant). Then $d(v_i, v_j) = d(v_i, x) + d(v_j, x)$. Therefore, if we now match the n_1 (n_2) points in the 1st (2nd) quadrant with the $n_3 = n_1$ ($n_4 = n_2$) points in the 3rd (4th) quadrant, we obtain a perfect matching with an objective value of $\sum d(v_i, x)$. To complete the proof we now show that the matching constructed is optimal. Consider

the pair of linear programs (3) and (4), where $d(v_i, v_j)$ is the rectilinear distance defined above. Problem (3) is certainly a relaxation of the integer matching problem. Problem (4) can be viewed as the following relaxation of the rectilinear median problem: Consider the complete undirected graph with node set $\{v_1, \ldots, v_n\} \cup \{v\}$. Each edge (v_i, v_j) is associated with a length which is equal to the rectilinear distance $d(v_i, v_j)$. The problem is to assign lengths $\{z_{v_i}\}$ to the edges $\{(v_i, v)\}$, such that the triangle inequality is satisfied for all edges of the graph, and $\sum\{z_{v_i} : v_i \in V\}$ is minimized. The rectilinear median problem in the plane corresponds to the special case where the lengths $\{z_{v_i}\}$ are the rectilinear distances from some point in the plane to the points $\{v_i\}$.

The perfect matching constructed above and the rectilinear median x correspond respectively to feasible solutions to problems (3) and (4). Moreover, as noted above, they yield the same objective value to both problems. Therefore, from duality theory, both are optimal solutions to the respective problems. This concludes the proof of the theorem. □

2.3. *The Euclidean planar b-matching problem*

In the previous sections we proved that the problem of finding a weighted median of a set of points $\{v_1, \ldots, v_n\}$ in a tree metric space (or in the rectilinear plane), is the dual of the respective b-matching problem when $\sum\{b_{v_i} : v_i \in V\}$ is even. We used this result to conclude that the respective fractional b-matching problem (3), has an integer solution. Example 9 demonstrates that this duality result does not hold even in the Euclidean plane.

Example 9. Consider the 3 vertices $\{v_1, v_2, v_3\}$ of an equilateral triangle with unit edge lengths, in the Euclidean plane. Let $b_{v_i} = 2$ for $i = 1, 2, 3$. The optimal solution value to the b-matching problem is 3. The weighted median problem in this case is to find a point in the plane that minimizes the sum of weighted Euclidean distances to the 3 vertices. The optimal value is $2\sqrt{3}$, which is greater than 3.

Note, however, that for any metric space the optimal value of the weighted median problem is greater than or equal to the optimal value of the respective b-matching problem.

In spite of the lack of duality in the Euclidean planar case, we prove in this section that the respective fractional b-matching problem (3) has an integer solution. We will then demonstrate in Section 2.5 that the integrality property of the rectilinear and Euclidean planar cases does not hold for higher dimensions.

We will need the following lemma dealing with the general b-matching problem.

Lemma 10. *Let $G = (V, E)$ denote the complete undirected graph having V as its node set. Each edge $e = (v, u) \in E$ is associated with a nonnegative real c_e, and each node*

$v \in V$ is associated with a positive integer b_v. Consider the b-matching problem (2), *and its linear relaxation, the fractional b-matching problem* (3).

(i) *There exists an optimal solution x^* to* (2) ((3)), *such that all nodes, but possibly one, are saturated, i.e., there exists a node v_0, and $x^*(\delta(v)) = b_v$, for each $v \in V - \{v_0\}$.*

(ii) *Suppose that the edge weights $\{c_e\}$ satisfy the triangle inequality. If* (3) *has no integer optimal solution, then* (3) *has an optimal solution x^*, such that all nodes are saturated, and each component of x^* is an integer multiple of* 1/2.

Proof. The first property follows directly from the nonnegativity of the edge weights $\{c_e\}$. To prove the second property consider a basic optimal solution of (3), x^*. It is well known (see [23]), that each component of x^* is an integer multiple of 1/2. If all nodes are not saturated we assume that node v_0 is the only unsaturated node in V. We also assume that x^* has the smallest saturation gap amongst all optimal solutions, that are integer multiples of 1/2. If $x_e^* = 0$, for all edges $e = (v, u)$, $v \neq v_0$, $u \neq v_0$, then x^* is integer, since $x_e^* = b_v$, for each edge $e = (v, v_0)$. Thus, suppose that there exists a pair of nodes v, u, $(v \neq v_0, u \neq v_0)$, and x_e^*, $e = (v, u)$, is positive. The saturation gap at v_0 is a positive integer, since it is equal to $\sum\{b_v : v \in V\} - 2\sum\{x_e^* : e \in E\}$.

Finally consider the solution of (3), obtained from x^* by subtracting 1/2 from x_e^*, $e = (v, u)$, and adding 1/2 to both x_e^*, $e = (v, v_0)$ and x_e^*, $e = (u, v_0)$. (Since the saturation gap at v_0 is at least one, this solution is feasible.) Since the edge weights satisfy the triangle inequality the new solution is certainly optimal. Moreover, all nodes, but possibly v_0, are saturated, and the saturation gap at v_0 has decreased. This contradicts the fact that x^* has the smallest possible gap. □

While the duality result does not, in general, hold for the Euclidean plane, we now show that an integrality result does hold.

We will need a few definitions. Let ab denote the directed closed line segment joining two distinct points, a and b, in the plane. A *closed polygonal walk*, $P = (v_1, \ldots, v_n)$, is the union of the line segments (*edges*) $v_1v_2, v_2v_3, \ldots, v_{n-1}v_n, v_nv_1$, joining n distinct points v_i.

If l_{ab} is the line containing ab, then the two rays that comprise the set $l_{ab} \setminus ab$ are called the *extension rays* (or, simply, *extensions*) of ab. We say that two line segments, ab and cd, are in *convex position* if the lines containing these two segments intersect at a point that is not contained in the interior of either of the two segments (i.e., the closure of an extension of ab intersects the closure of an extension of cd).

We will need the following lemma, which may have independent interest.

Lemma 11. *Let P and Q be closed polygonal walks in the plane, each having an odd number of vertices. Then there exists an edge of P and an edge of Q such that the two edges are in convex position.*

Proof. For simplicity, we assume that no three vertices of $P \cup Q$ are colinear (in particular, then, the vertices of $P \cup Q$ are distinct), and that no edge of P is parallel to an edge of Q; the lemma remains true in degenerate cases, but the details are tedious.

Let us begin with some notation. The *lines of* Q are simply the lines through edges of Q. For a polygonal walk $P = (v_1, \ldots, v_n)$, we say that vertex v_i *induces* the closed convex cone C_i whose boundary consists of the two rays (with apex v_i) that are the extensions of (directed) edges $v_{i-1}v_i$ and $v_{i+1}v_i$. If a line intersects a cone in a ray, then we say that the line and the cone are *incident.*

The proof of the lemma is based on two simple claims:

Claim 1. *Let P be a closed polygonal walk having an odd number of vertices. Then, any line l that is not parallel to an edge of P must be incident on an odd number of cones induced by P. (In particular, this implies that any line l must be incident on* some *induced cone of P.)*

Proof. Without loss of generality, assume that l is the y-axis. Since l is not parallel to any edge of P, a vertex of P can be classified as one of three types, according to the local geometry of the two incident edges: left (both incident edges are to the right of the vertex), right (vice versa), or middle (otherwise). There must be an even number of left/right vertices, since between any 2 left (resp., right) vertices there is a right (resp., left) vertex. Thus, since P is odd, there are an odd number of middle vertices. But, for any middle vertex, the induced cone necessarily includes the vertical direction, so the y-axis will intersect such a cone in a ray. □

Claim 2. *Let C_i be a cone induced by P. If C_i is incident on an odd number of lines of Q, then one of the two rays bounding C_i must intersect an extension ray of Q.*

Proof. Assume that there are an odd number of lines of Q incident on C_i. Consider a line l of Q that is incident on C_i. Consider the two vertices of Q that are on l. If neither vertex lies in C_i or if both vertices lie in C_i, then an extension ray of Q (namely, one of the two on l) must cross the boundary of C_i, and we are done. Thus, we can assume that, for any line l of Q that is incident on C_i, there is exactly one vertex, v, of Q on l that lies within C_i. Each vertex of Q has exactly two lines through it; thus, there is a second line, l', of Q through v. If l' is *not* incident on C_i, then the extension ray of Q that has apex v and lies on line l' must intersect the boundary of C_i, and we are done. Thus, we can assume that each vertex of Q that lies within C_i has both of its lines incident on C_i. But since each line of Q that is incident on C_i must have exactly one of its two vertices within C_i, this implies that there are an even number of lines of Q incident on C_i – a contradiction. □

Now, by Claim 1, each line of Q is incident on an odd number of cones induced by P. Since there are an odd number of lines of Q, this implies that the total number of

line-cone incidences is odd.

This implies that some cone, C_i, of P is incident on an *odd* number of lines of Q. But then, by Claim 2, a ray bounding C_i must intersect an extension ray of Q, implying that there is a crossing of extensions. This concludes the proof of the lemma. □

Remark 12. Janos Pach has suggested to us an alternate (but logically equivalent) proof of this lemma; this same proof was previously discovered (but not published) by M. Perles. [1] The problem arises in the study of "geometric graphs", where the following Turán-type question arises: What is the smallest number $k(n)$ such that in any geometric graph on n points in the plane that has at least $k(n)$ edges, there must exist a pair of edges in convex position? A simple construction shows that $k(n) \geqslant 2n - 1$. Kupitz [16,17] conjectured that $k(n) = 2n - 1$. Recently, Katchalski and Last [15] have shown that $k(n) \leqslant 2n$.

Theorem 13. *Let V be a finite set of points in R^2. Each point v in V is associated with a positive integer b_v. Suppose that $\sum\{b_v : v \in V\}$ is even. Let $G = (V, E)$ denote the complete undirected graph having V as its node set. For each edge (pair of points) $e = (v, u)$, let c_e denote the Euclidean distance between v and u. Then, the fractional b-matching problem (3) has an integer optimal solution, where the number of its positive components is at most $|V|$.*

Proof. We assume, without loss of generality, that the points in V are distinct. (Otherwise, identical points can be aggregated without affecting the evenness assumption on the sum of the node weights.)

Moreover, for the sake of brevity, in order to avoid complicated and messy perturbation arguments, we assume that the points in V satisfy the following general position assumption: no three points of V are colinear.

If problem (3) has no integer solution, then by Lemma 10 there exists an optimal solution to (3), which is an integer multiple of 1/2, and all nodes are saturated.

We will now prove the claim that if problem (3) has a saturated optimal solution, then every basic optimal solution that is saturated is integer.

Let x^* be a saturated basic optimal solution. Suppose that x^* is not integer. x^* is an integer multiple of 1/2.

Let E^* be the subset of edges associated with integer components of x^*. Since x^* is not integer, E^* is a proper subset of E. For each node $v \in V$, let $\delta^*(v)$ be the subset of edges of E^* that are incident to v. Consider the instance of problem (3), obtained by subtracting $x^*(\delta^*(v))$ from b_v, for each $v \in V$. Refer to this instance as the reduced problem. Let y^* be the vector obtained from x^* by equating to zero each component x_e^* such that $e \in E^*$. Then, y^* is an optimal solution to the reduced problem, and all

[1] We thank Noga Alon for pointing out the existence of Perles' unpublished result and for acquainting us with the relevant literature.

its nonzero components are odd integer multiples of 1/2. To prove the claim it is now sufficient to contradict the optimality of y^*.

Consider the subset E' of edges associated with positive components of y^*. Since x^* is a basic solution, it follows from Pulleyblank [23] that the subgraph $G' = (V, E')$ has no even cycles. In particular, all cycles are odd, and they are node disjoint. If G' has no cycles, it is a forest. Since all nodes are saturated, y^* is integer. If G' has exactly one cycle, say C, then all edges which are not on C must be associated with integer components of y^*. But y^* has no positive integer components. Therefore, all edges of G' are on C. Since C is an odd cycle, $2\sum\{y_e^* : e \in C\}$ is odd. This is not possible, since $2\sum\{x_e^* : e \in E - C\}$ is even, and

$$\sum\{b_v : v \in V\} = 2\sum\{y_e^* : e \in C\} + 2\sum\{x_e^* : e \in E - C\}.$$

Thus, G' contains two odd cycles C_1 and C_2, which are node disjoint. The two cycles correspond to two closed polygonal walks P_1 and P_2 in the plane, each having an odd number of vertices. From Lemma 11 there exist an edge of P_1 and an edge of P_2 such that the two edges are in convex position.

To simplify the notation, suppose that $V = \{v_1, v_2, \ldots, v_n\}$, and that $P_1 = (v_1, v_2, \ldots, v_k)$ and $P_2 = (v_{k+1}, \ldots, v_p)$. Also, assume that the edges (v_1, v_2) and (v_{k+1}, v_{k+2}) are in convex position. This implies that the four points $\{v_1, v_2, v_{k+1}, v_{k+2}\}$ are on the boundary of their convex hull polygon, where v_2 is adjacent to v_1, and v_{k+2} is adjacent to v_{k+1}. Without loss of generality, suppose that their order along the boundary is given by $v_1, v_2, v_{k+1}, v_{k+2}$. Consider the solution, z^*, obtained from y^* by subtracting 1/2 from $y^*_{(v_1,v_2)}$ and from $y^*_{(v_{k+1},v_{k+2})}$ and adding 1/2 to the components $y^*_{(v_1,v_{k+1})}$ and $y^*_{(v_2,v_{k+2})}$.

Then, z^* is a saturated feasible solution that is an integer multiple of 1/2. From the assumption that the points in V are in general position, it follows that the convex hull of the above four points is not a line segment. Therefore, it is easy to see that the objective value at z^* is larger than that of y^*- contradicting the optimality of y^*. This completes the proof of the claim. From the above claim it follows that problem (3) has an integer optimal solution. We can also use the above claim to show that there is an integer optimal solution, where the number of its positive components is at most $|V|$.

If there is a saturated optimal solution, it follows directly from the above claim that every saturated basic optimal solution is integer. If there is no saturated integer optimal solution, then there is an integer optimal solution x^* and a node v_0, such that all nodes but v_0 are saturated. Moreover, the saturation gap at v_0 is even. Consider the problem obtained from (3) by reducing the (even) saturation gap from b_{v_0}. This modified problem is a saturated optimal solution. Therefore, by the above claim each one of its saturated basic optimal solutions is integer. This completes the proof of the theorem. □

Remark 14. The proof above applies not only to the Euclidean metric, but to any metric that obeys the simple property that a straight path between p and q is an instance

of a shortest path in the metric. In particular, this geometric proof gives an alternative method of proof for the integrality result of Theorem 8.

Remark 15. We have already noted above that the planar Euclidean median problem is not the dual of the respective matching problem. However, we can use the integrality result of Theorem 13. to define an appropriate dual. Problem (4) can be viewed as the following relaxation of the weighted planar Euclidean median problem: Consider a metric space extension of the Euclidean plane, which preserves the Euclidean distances between the points in V. The extended weighted median problem is to find a point w in such a space so that the sum of the weighted distances (with respect to the new metric), of w from all points in V is minimized. For each $v \in V$, z_v will denote the distance from w to v. The duality of (3) and (4), and the integrality result of Theorem 13, imply that the extended weighted median problem is the dual of the b-matching problem for the Euclidean planar case.

2.4. *Geodesic metrics in the plane*

We will next show that the integrality result of Theorem 13 holds if we replace the Euclidean metric with the *geodesic* metric induced by a set of polygonal "obstacles" in the plane. More formally, let F denote a bounded, closed, (multiply) connected region in R^2, whose boundary consists of a union of a finite number of (bounded) line segments. (Such a set F is commonly called a *polygonal domain.*) We refer to the set F as the *free space*; the complement, $R^2 \setminus F$, is a finite set S of (connected) polygonal *obstacles.* Since F is bounded and connected, the obstacle set S consists of one unbounded (open) polygon, and h bounded (open) simple polygons, where h is the number of holes in F. Note that we consider the free space F to be *closed* (it includes its boundary). Let W denote the (finite) vertex set for F (and, hence, for S).

The *geodesic distance with respect to* F, $d_F(p,q)$, between point $p \in F$ and point $q \in F$ is defined to be the (Euclidean) length of a shortest obstacle-free path from p to q (i.e., the length of a shortest path joining p to q within the free space F). A shortest path within F is called a *geodesic*(F) *path.* It is well known (see [1]) that a *geodesic*(F) path between any two points of the free space is a simple (non-self-intersecting) polygonal path, whose vertices (bend points) are among the set W of obstacle vertices. In general, there may be many shortest paths between two points; in the special case that F is a simple polygon (without holes), there is a unique shortest path.

Theorem 16. *Let F be a polygonal domain (free space) in the plane, having vertex set W. Let $U = \{u_0, \ldots, u_{k-1}\}$ and $V = \{v_0, \ldots, v_{l-1}\}$ be finite sets of points in the free space, having odd cardinalities k and l. Let P (resp., Q) be a closed polygonal walk linking up the points U (resp., V) using geodesic(F) paths $\pi(u_i, u_{i+1})$ (resp., $\pi(v_j, v_{j+1})$). Then there exists a closed polygonal walk R linking up the points $U \cup V$*

using geodesic(F) paths, such that the length of R is greater than or equal to the sum of the lengths of P and of Q.

Proof. We make the following four nondegeneracy assumptions about the input: (1) F equals the closure of the interior of F. (2) No three points of $\mathcal{V} = U \cup V \cup W$ are colinear. (3) There is a *unique geodesic*(F) path between any pair of points $\mathcal{V} = U \cup V \cup W$. (4) No point $w \in U \cup V$ lies on some geodesic path between a pair of points $w_1, w_2 \in U \cup V$ ($w \neq w_1, w_2$). (This is analogous to the assumption in the Euclidean plane (without obstacles) that no three points are colinear.) By a perturbation argument, the proof can be extended to handle the cases when these assumptions do not hold, but the details are tedious and omitted here.

Let *VG* denote the *visibility graph* whose node set is $\mathcal{V} = U \cup V \cup W$ and whose edges (called *VG-edges*) connect two nodes if and only if the straight line segment joining the two nodes is contained in the free space F. By our nondegeneracy assumption (2), a VG-edge intersects no points of $\mathcal{V}$ other than at its endpoints. Since F is closed, the VG-edges include each edge (line segment) of the boundary of F. Any geodesic path between two nodes in $\mathcal{V}$ must lie on the visibility graph (i.e., consist of a union of segments that are VG-edges); see [1]. In particular, the geodesic paths $\pi(u_i, u_{i+1})$ and $\pi(v_j, v_{j+1})$ that constitute the "edges" (call them *g-edges*) of the walks P and Q must lie on the visibility graph.

For a set of three distinct points $z_1, z_2, z_3 \in U \cup V$, let $\gamma(z_1, z_2, z_3)$ denote the closed (oriented) Jordan cycle going from z_1 to z_2 along $\pi(z_1, z_2)$, from z_2 to z_3 along $\pi(z_2, z_3)$, and then from z_3 back to z_1 along $\pi(z_3, z_1)$. The fact that the three paths, $\pi(z_1, z_2)$, $\pi(z_2, z_3)$, and $\pi(z_3, z_1)$ form a closed Jordan curve follows from our nondegeneracy assumption, which ensures that any two *geodesic*(F) paths that intersect do so in a connected set (i.e., a single subpath).

From our assumptions, we know that $\gamma(z_1, z_2, z_3)$ separates R^2 into two regions – a bounded region, which we call a *geodesic triangle*, and an unbounded region. We say that $z_3 \in U \cup V$ *lies to the left of* (z_1, z_2) if the region to the left of $\gamma(z_1, z_2, z_3)$ is bounded (a *geodesic triangle*). (The region will necessarily have nonempty interior, from our nondegeneracy assumptions (1) and (4).) We let $L(z_1, z_2)$ denote the set of all such points $z_3 \in U \cup V$. Similarly, we say that $z_3 \in U \cup V$ *lies to the right of* (z_1, z_2) if the region to the right of $\gamma(z_1, z_2, z_3)$ is bounded, and we let $R(z_1, z_2)$ denote the set of such points $z_3 \in U \cup V$.

The following definitions apply to distinct points $z_1, z_2, z_3, z_4 \in U \cup V$. We say that $\pi(z_1, z_2)$ *points at* $\pi(z_3, z_4)$ if $z_3 \in L(z_1, z_2)$ and $z_4 \in R(z_1, z_2)$, or $z_3 \in R(z_1, z_2)$ and $z_4 \in L(z_1, z_2)$.

We say that $\pi(z_1, z_2)$ and $\pi(z_3, z_4)$ *cross* each other if they intersect (in a connected subpath) *and* the ordering of the points about the union $\pi(z_1, z_2) \cup \pi(z_3, z_4)$ is z_1, z_3, z_2, z_4 (in either clockwise or counterclockwise traversal of the union). Note that $\pi(z_1, z_2)$ and $\pi(z_3, z_4)$ cross if and only if $\pi(z_1, z_2)$ points at $\pi(z_3, z_4)$ and $\pi(z_3, z_4)$ points at $\pi(z_1, z_2)$.

We say that g-edge $\pi(z_1, z_2)$ and g-edge $\pi(z_3, z_4)$ are in *convex position* if $\pi(z_1, z_2)$ does not point at $\pi(z_3, z_4)$ *and* $\pi(z_3, z_4)$ does not point at $\pi(z_1, z_2)$. (To motivate

this definition, note that it is equivalent to saying that the g-edges do not cross and the "geodesic convex hull" of the endpoints $A = \{z_1, z_2, z_3, z_4\}$, $CH(A)$, has all four endpoints on it. $\mathrm{CH}(A)$ is the simply-connected closed region having minimum-length perimeter, such that $A \subseteq \mathrm{CH}(A) \subseteq F$.)

Consider a g-edge $e = \pi(u_i, u_{i+1})$ of P and a g-edge $f = \pi(v_j, v_{j+1})$ of Q. There are four cases for any pair (e, f):

(1) e points at f *and* f points at e, implying that e and f cross; or
(2) e points at f *and* f does *not* point at e;
(3) f points at e *and* e does *not* point at f;
(4) e does *not* point at f *and* f does *not* point at e, implying that e and f are in convex position.

Now, if P and Q are both odd cycles, then there are an odd number of pairs (e, f). Our goal is to prove that the number of pairs (e, f) corresponding to case 4 above is *odd*, which, in particular, will show that there is at least one such pair in convex position. Thus, it suffices to show that each of the cases 1–3 arise an *even* number of times. This is done in the following two claims.

Claim 1. *There are an even number of pairs of crossing g-edges, $e = \pi(u_i, u_{i+1})$ of P and $f = \pi(v_j, v_{j+1})$ of Q.*

Proof. Since any pair of g-edges intersect in a connected set (or not at all), we can perturb the g-edges of P and of Q so that in the arrangement of perturbed g-edges of P and Q, any pair of crossing g-edges share exactly one point, any pair of adjacent g-edges along P (Q) share exactly one point, while all other pairs are disjoint. (In particular, two non-adjacent g-edges that coincide along a subpath, but do not cross, can be perturbed so that they are disjoint.) Since P is a closed walk, the above perturbation implies that the planar arrangement of P is a planar graph, with all of its nodes having an even degree. Thus, the planar dual of this arrangement (with nodes corresponding to the 2-faces) is bipartite, and hence can be 2-colored. Now, if we start at the vertex v_0 of Q and walk along the closed curve Q until we return to v_0, we will cross from face to face in the arrangement of P. If we cross from a face of one color to a face of the *same* color, then the total number of g-edges of P that are crossed is even. We call such a crossing an *even crossing*. (This can happen only if the two faces share a common vertex of the arrangement of P, but do not share a common 1-face.) If we cross from a face of one color to a face of the other color, then the number of g-edges of P that are crossed is odd. Such a crossing is called an *odd crossing*. The total number of odd crossings must be even, since we will end up back in the cell where we started (namely, the 2-face containing v_0). Thus, the total number of crossings of g-edges of P by the curve Q is even. Now, the claim follows directly from the fact that a g-edge of Q can intersect a g-edge of P in at most one point. □

Claim 2. *A g-edge $f = \pi(v_j, v_{j+1})$ of Q points at an even number of g-edges $\pi(u_i, u_{i+1})$ of P. Similarly, a g-edge $e = \pi(u_i, u_{i+1})$ of P points at an even number of g-edges $\pi(v_j, v_{j+1})$ of Q.*

Proof. A vertex u_i of P is either left of $f = \pi(v_j, v_{j+1})$ or right of $f = \pi(v_j, v_{j+1})$. (Here, we are invoking the nondegeneracy assumptions (3) and (4).) Starting with vertex u_0 of P, consider walking along P until we return to u_0. Each time that we go from a vertex that is left (resp., right) of $\pi(v_j, v_{j+1})$ to a vertex that is right (resp., left) of $\pi(v_j, v_{j+1})$, the corresponding g-edge of P is pointed at by $\pi(v_j, v_{j+1})$. Since P is a closed walk, it is easy to see that the number of such transitions is even, implying the claim. □

This allows us to conclude that:

Lemma 17. *There exist g-edges $e = \pi(u_i, u_{i+1})$ of P and $f = \pi(v_j, v_{j+1})$ of Q that are in convex position.*

Finally, we claim that

Claim 3. *If $e = \pi(u_i, u_{i+1})$ and $f = \pi(v_j, v_{j+1})$ are in convex position, then either*
(a) *$\pi(u_i, v_{j+1})$ intersects $\pi(v_j, u_{i+1})$, or*
(b) *$\pi(u_i, v_j)$ intersects $\pi(u_{i+1}, v_{j+1})$.*

Proof. Since e and f are in convex position, we know that e does not point at f, and f does not point at e. Assume, without loss of generality, that $v_j, v_{j+1} \in L(u_i, u_{i+1})$. If (a) and (b) both fail to hold, then we must have a planar embedding of the two geodesic triangles, $\gamma(u_i, u_{i+1}, v_{j+1})$ and $\gamma(u_i, u_{i+1}, v_j)$. Since they share the g-edge $e = \pi(u_i, u_{i+1})$, we must have either v_j contained in the region bounded by $\gamma(u_i, u_{i+1}, v_{j+1})$ or v_{j+1} contained in the region bounded by $\gamma(u_i, u_{i+1}, v_j)$. But, in either case, this implies that u_i and u_{i+1} do not lie on the same side of $f = \pi(v_j, v_{j+1})$ (both in $L(v_j, v_{j+1})$, or both in $R(v_j, v_{j+1})$), contradicting the fact that f does not point at e. □

From the above claim, and the lemma, we know that there exist g-edges $e \in P$ and $f \in Q$ such that either (a) or (b) holds; in particular, we may assume that $e = \pi(u_i, u_{i+1})$ and $f = \pi(v_j, v_{j+1})$ exist such that $\pi(u_i, v_{j+1})$ intersects $\pi(v_j, u_{i+1})$ at some point $c \in P \cap Q$. By performing an "exchange" at point c, we can construct a new closed walk, R, as follows: Starting at u_0, follow P to u_i, then follow $\pi(u_i, v_{j+1})$ to v_{j+1}, then follow Q to v_j, then follow $\pi(v_j, u_{i+1})$ to u_{i+1}, then follow P back to u_0. The fact that R is longer than or equal to the sum of the lengths of P and of Q follows from the triangle inequality. In particular, the exchange has saved length

$$d(u_i, u_{i+1}) + d(v_j, v_{j+1}),$$

while adding length

$$\begin{aligned} d(u_i, v_{j+1}) + d(v_j, u_{i+1}) &= d(u_i, c) + d(c, v_{j+1}) + d(v_j, c) + d(c, u_{i+1}) \\ &\geqslant d(u_i, u_{i+1}) + d(v_j, v_{j+1}). \end{aligned}$$

□

Theorem 16 can be used to generalize Theorem 13 to the case where for each edge (pair of points) $e = (v, u)$, c_e is the Euclidean length of a geodesic path connecting v

and u. Following the proof of Theorem 13, we observe that in the case of the geodesic metric Theorem 16 enables us to (successively) replace a pair of odd cycles by an even cycle whose length is equal to or greater than the sum of the lengths of the two odd cycles. A solution with no odd cycles can be represented as a convex combination of integer solutions. For the sake of brevity we skip the formal technical details. For completeness we state the generalized theorem.

Theorem 18. *Let F be a polygonal domain (free space) in R^2, and let V be a finite set of points in F. Each point v in V is associated with a positive integer b_v. Suppose that $\sum\{b_v : v \in V\}$ is even. Let $G = (V, E)$ denote the complete undirected graph having V as its node set. For each edge (pair of points) $e = (v, u)$, let c_e denote the Euclidean length of a geodesic(F) path between v and u. Then, the fractional b-matching problem (3) has an integer optimal solution, where the number of its positive components is at most $|V|$.*

2.5. Higher dimensions

We now consider the higher dimensional versions of the problems discussed above. First, we note that the duality result of Theorem 8 does not extend to the case where the points $\{v_i\}$ are in R^3 and the metric is rectilinear. This is demonstrated by the next example.

Example 19. Consider the four points $v_1 = (-1, 1, -1)$, $v_2 = (1, 1, 1)$, $v_3 = (1, -1, -1)$ and $v_4 = (-1, -1, 1)$. Let $b_{v_i} = 1$, $i = 1, 2, 3, 4$. The solution value to the b-matching problem is 8, while the optimal solution value to the respective median problem is 12.

Furthermore, we also show that the integrality property does not hold in R^3, even if there are an even number of nodes and the weights $\{b_v\}$ are all 1 (i.e., even for the case of maximum weight perfect matching on an even set of nodes). Example 20 considers the Euclidean case, while Example 21 considers the rectilinear case.

Example 20. Consider the following six points in R^3: $v_1 = (0, 1, 0)$, $v_2 = (1.4, 0, 0)$, $v_3 = (0, -1, 0)$, $v_4 = (0, 0, 0)$, $v_5 = (1.4, 0, -1)$, $v_6 = (1.4, 0, 1)$. Then, each of the two odd cycles, $P = (v_1, v_2, v_3)$ and $Q = (v_4, v_5, v_6)$, has Euclidean length $2(1 + \sqrt{2.96})$, implying that the fractional solution having $x_e = 1/2$ for edges e in P or Q, and $x_e = 0$ otherwise, has objective value $2(1 + \sqrt{2.96})$. On the other hand, it is easy to check that every perfect matching on these six points has length less than $2(1 + \sqrt{2.96})$.

Example 21. Consider the following six points in R^3: $v_1 = (0, 0, 2)$, $v_2 = (1, 1, 1)$, $v_3 = (2, 0, 0)$, $v_4 = (1, 0, 1)$, $v_5 = (2, 1, 2)$, $v_6 = (0, 1, 0)$. Then, each of the two odd cycles, $P = (v_1, v_2, v_3)$ and $Q = (v_4, v_5, v_6)$, has rectilinear length 10, implying that the fractional solution having $x_e = 1/2$ for edges e in P or Q, and $x_e = 0$ otherwise, has objective value 10. On the other hand, it is easy to check that every perfect matching on these six points has length at most 9.

3. An extended b-matching problem

In this section we generalize the results of Section 2.1, and consider the following extended model.

Referring to problem (3) above, where c_e is defined as the respective tree distance $d(v,u)$, suppose that the degree coefficient b_v of each node v can be increased linearly provided that a certain nonnegative penalty p_v is paid. The objective is now to maximize the net reward obtained by subtracting the total penalty from the weighted matching. Specifically, consider the following problem:

$$\begin{aligned} \text{Maximize} \quad & \sum_{e\in E} c_e x_e - \sum_{v\in V} p_v y_v \\ \text{s.t.} \quad & x(\delta(v)) \leqslant b_v + y_v \quad \forall v \in V, \\ & x \geqslant 0, \ y \geqslant 0. \end{aligned} \tag{6}$$

We will extend Theorem 3 and show that problem (6) has an integer optimal solution. The dual linear program of (6) is given by

$$\begin{aligned} \text{Minimize} \quad & \sum_{v\in V} b_v z_v \\ \text{s.t.} \quad & z_v + z_u \geqslant d(v,u) \quad \text{for all pairs } v,u, \\ & z_v \leqslant p_v \quad \text{for all nodes } v, \\ & z \geqslant 0. \end{aligned} \tag{7}$$

Again, using the general results in [6], we note that problem (7) is a reformulation of the following restricted weighted median problem. Find a point in $A(T)$ whose distance from each node v is at most p_v, and such that the sum of weighted distances of all nodes to the point is minimized. The optimal solution is called a *restricted weighted median*. If an unrestricted median satisfies the distance constraints it is optimal. Otherwise, suppose that the set of all points on the tree which satisfy the distance constraints is nonempty. It is easily observed that this feasible set is a connected set (closed subtree) of $A(T)$. The optimal restricted weighted median is the closest point to the unrestricted weighted median in the feasible set.

Theorem 22. *Suppose that b_v is a positive integer for each $v \in V$. If problem (6) has a finite optimal value, then there is an optimal integer solution x^*, y^* to problem (6), where the number of positive variables $\{x_e^*\}$ is at most $|V|$, and the number of positive variables $\{y_v^*\}$ is at most 1.*

Proof. Let t be a point on the tree where the optimal restricted weighted median is located. In particular, using the above duality the optimal value of (6) and (7) is $\sum\{b_v d(v,t) : v \in V\}$. Suppose without loss of generality that t is a point on some edge (v,u) of T and t is not the unrestricted weighted median v^*. Let v' be such that $d(v',t) = p_{v'}$ and the point t is on the unique path connecting v^* and v'. Let V' denote

the set of nodes in V having the point t on the unique simple path connecting them to v^*. Define

$$B' = \sum\{b_v : v \in V'\}.$$

By definition we have $B' \leqslant B/2$. Next define $y_{v'} = B - 2B'$. Add $y_{v'}$ to the weight $b_{v'}$. It then follows from Theorem 2 that each point on the edge (v, u) which contains t, is an optimal unrestricted median with respect to the modified set of weights $\{b_v\}$, $v \neq v'$, $\{b_{v'} + y_{v'}\}$. Using Theorem 3 for the modified weights we conclude that there exists an integer matching x^*, with at most $|V|$ positive components such that

$$x^*(\delta(v)) \leqslant b_v \quad \forall v \neq v',$$
$$x^*(\delta(v')) \leqslant b_{v'} + B - 2B',$$

and

$$\sum_{e \in E} c_e x_e^* = \sum_{v \in V} b_v d(v, t) + (B - 2B') d(v', t). \tag{8}$$

Define y^* by

$$y_v^* = \begin{cases} B - 2B' & \text{if } v = v', \\ 0 & \text{otherwise.} \end{cases}$$

Then (x^*, y^*) is a feasible solution to problem (6) with objective value of

$$\sum_{e \in E} c_e x_e^* - (B - 2B') d(v', t) = \sum_{e \in E} c_e x_e^* - (B - 2B') p_{v'}.$$

From (8) this objective value is equal to the optimal value of the restricted weighted median problem, problem (7). Using the duality of problems (6) and (7), we conclude that (x^*, y^*) satisfies the statement of the theorem. □

4. The roommates problem

Consider the roommates problem defined in the Introduction. Let V indicates the group of n boys. The boys' preferences are defined according to the following scheme. Each boy is represented by some point v of a given metric space X. For each pair of boys v, u let $d(v, u)$ denote the distance between them. For simplicity suppose that the $n(n-1)/2$ distances between all pairs of boys v, u, $v \neq u$, are distinct. We consider the case where alternatives are preferred more as they are farther from each boy's location. Each boy ranks all other boys by their distances from him, where the highest preference is given to the farthest one. The roommate problem with these preferences is easily observed to have a unique stable matching: First match the unique pair of boys whose distance is the diameter of the set of n representing points, i.e., the pair v, u corresponding to the largest distance. Remove this pair of points from the set, and repeat the process with the remaining set. Having established the existence of a stable matching, we follow Granot

[10] and analyze a cooperative game associated with the roommates problem. (We consider the multiple version of this roommates problem.) This game with side payments, $(V; f)$, is defined by its characteristic function $f : 2^V \to R$. Suppose that $G = (V, E)$ is the complete undirected graph with node set V. Each node $v \in V$ represents a boy.

For each nonempty subset S of V (coalition), let

$$f(S) = \text{Max}\Big\{\sum_{e\in E} c_e x_e : \text{s.t. } x(\delta(v)) \leqslant b_v \text{ for each } v \in S,$$
$$x(\delta(v)) \leqslant 0 \text{ for each } v \in V - S,$$
$$x \geqslant 0 \text{ and integer}\Big\}.$$

The core of the cooperative roommates problem game, $C[(V; f)]$, is defined by

$$C[(V; f)] = \Big\{y = (y_v : v \in V) : \sum\{y_v : v \in S\} \geqslant f(S) \quad \forall S \subset V,$$
$$\text{and } \sum\{y_v : v \in V\} = f(V)\Big\}.$$

Granot [10] considered the cooperative game corresponding to the classical matching model, where each player (boy) v is to be paired with a single roommate, i.e. $b_v = 1$ for each $v \in V$, and demonstrated that the core might be empty when the $\{c_e\}$ coefficients are arbitrary nonnegative reals. The games corresponding to Example 7 and Example 21 show that the core might be empty even when these coefficients satisfy the triangle inequality and $|V| = 6$. (The core of any game with $|V| = 4$ is nonempty if the triangle inequality holds.) We now use the above integrality results to prove that the core is nonempty if the $\{c_e\}$ coefficients are defined by any of the distance functions discussed in Section 2.

Theorem 23. *Let $G = (V, E)$ be the complete undirected graph with node set V. Each edge $e \in E$ is associated with a nonnegative weight c_e, and each node $v \in V$ is associated with a positive integer b_v. Suppose that the fractional b-matching problem (3) has an integer optimal solution. Let $z^* = (z_v^* : v \in V)$ be an optimal solution to problem (4). Then the vector $y = (y_v = b_v z_v^* : v \in V)$, is in the core $C[(V; f)]$ of the respective roommates problem game.*

Proof. Given the game $(V; f)$ defined above we consider the fractional roommates problem game, $(V; f^*)$, where for each $S \subset V$,

$$f^*(S) = \text{Max}\Big\{\sum_{e\in E} c_e x_e : \text{ s.t. } x(\delta(v)) \leqslant b_v \text{ for each } v \in S,$$
$$x(\delta(v)) \leqslant 0 \text{ for each } v \in V - S,$$
$$x \geqslant 0\Big\}.$$

Since we have dropped the requirement that x should be integer, we have $f^*(S) \geqslant f(S)$ for all $S \subset V$. Moreover, since (3) has an integer optimal solution $f^*(V) = f(V)$.

Therefore, $C[(V; f^*)] \subset C[(V; f)]$. Since the game $(V; f^*)$ can be viewed as a linear production game, it follows from [22] that if z^* is an optimal solution to problem (4), the vector $y = (y_v = b_v z_v^* : v \in V)$ is in the core of the game $(V; f^*)$. This completes the proof of the theorem. □

Remark 24. If b_v is not constant for all $v \in V$, the core $C[(V; f)]$ may contain vectors that cannot be generated from the solutions to problem (4) by the above process. As an example consider the case of a 3-node tree, $V = \{v_1, v_2, v_3\}$, where $b_{v_1} = b_{v_2} = 1$, $b_{v_3} = 2$, and $d(v_1, v_2) = d(v_2, v_3) = 1$. The vector (0,1,2) is in the core, but there is no optimal solution z to problem (4) with $z_{v_1} = 0$. The example given in Remark 5 shows that the core $C[(V; f)]$ may be empty if the sum of all $\{b_v\}$ coefficients is odd and the b coefficient of the weighted median is zero.

Theorem 25. *Let $G = (V, E)$ be the complete undirected graph with node set V. Each edge $e \in E$ is associated with a nonnegative weight c_e, and each node $v \in V$ is associated with a positive integer b_v. Assume that $b_v = b'$, for all $v \in V$, for some positive integer b'. Suppose that the fractional b-matching problem (3) has an integer solution. If y is a vector in $C[(V; f)]$, then the vector y/b' is an optimal solution of (4).*

Proof. We prove that if $b_v = b'$ for each $v \in V$, then for each vector $y \in C[(V; f)]$, the vector $z = y/b'$ is an optimal solution to problem (4). If y is in the core, then it follows from the duality of (3)-(4) that $\sum\{y_v : v \in V\}$ is equal to the optimal objective value of problem (4). But $\sum\{b' z_v : v \in V\} = \sum\{y_v : v \in V\}$. Thus, it is sufficient to prove that z satisfies the constraints of problem (4). The monotonicity of f implies that y is nonnegative. Therefore, by definition, z is nonnegative. We have to prove that $z_v + z_u \geqslant d(v, u)$ for each pair of nodes v, u. Indeed, consider the coalition $S = \{v, u\}$. Since $b'z$ is in the core we have $b' z_v + b' z_u \geqslant f(S)$. But $f(S)$ is easily observed to be equal to $b' d(v, u)$. □

It may happen that the set of optimal matchings corresponding to the game discussed in Theorem 23 has no stable matching with respect to the preferences defined above. This is demonstrated by the next example.

Example 26. Consider a 6-node path, with $d(v_1, v_2) = 5$, $d(v_i, v_{i+1}) = 1$, $i = 2, 3, 4, 5$. Then add a new node, v_7, and connect it to v_2 with an edge of length 5. The node v_3 is the unique median of this tree. The nodes v_1 and v_7 should be matched in any stable matching. However this is not the case in any optimal matching to problem (3).

Acknowledgements

We thank Noga Alon, Esther Arkin, Micha Sharir, and Emo Welzl for useful discussions on portions of this research. We also thank the referees for several suggestions that improved the presentation of this paper. Joe Mitchell thanks Noga Alon and Micha Sharir for hosting his visit to the Special Semester in Computational Geometry at Tel Aviv University, which provided the opportunity for this research collaboration.

References

[1] H. Alt and E. Welzl, Visibility graphs and obstacle-avoiding shortest paths, *Zeitschrift für Operations Research* 32 (1988) 145–164.
[2] R.P. Anstee, A polynomial algorithm for b-matchings: An alternative approach, *Information Processing Letters* 24 (1987) 153–157.
[3] J. Bartholdi and M.A. Trick, Stable matching with preferences derived from psychological model, *Operations Research Letters* 5 (1986) 165–169.
[4] W. Cook and W.R. Pulleyblank, Linear systems for constrained matching problems, *Mathematics of Operations Research* 12 (1987) 97–120.
[5] W.H. Cunningham and A.B. Marsh III, A primal algorithm for optimum matching, *Mathematical Programming Study* 8 (1978) 50–72.
[6] E. Erkut, R.L. Francis, T.J. Lowe and A. Tamir, Equivalent mathematical programming formulations of tree network location problems, *Operations Research* 37 (1989) 447–461.
[7] D. Gale and L.S. Shapley, College admissions and the stability of marriage, *American Mathematical Monthly* 69 (1962) 9–14.
[8] O. Gerstel and S. Zaks, A new characterization of tree medians with applications to distributed sorting, *Networks* 24 (1994) 23–39.
[9] A.J. Goldman, Optimal center location in simple networks, *Transportation Science* 5 (1971) 212–221.
[10] D. Granot, A note on the roommates problem and a related revenue allocation problem, *Management Science* 30 (1984) 633–643.
[11] D. Gusfield and R.W. Irving, *The Stable Marriage Problem* (MIT Press, Cambridge, MA, 1989).
[12] J. Hershberger and S. Suri, Matrix searching with the shortest path metric, *Proceedings of the 25th Annal ACM Symposium on Theory of Computing* (1993) 485–494.
[13] A.N.C. Kang and D.A. Ault, Some properties of a centroid of a free tree, *Information Processing Letters* 4 (1975) 18–20.
[14] O. Kariv and L.S. Hakimi, An algorithmic approach to network location problems. II: The p-medians, *SIAM Journal on Applied Mathematics* 37 (1979) 539–560.
[15] M. Katchalski and H. Last, Disjoint edges in convex position in a geometric graph, in: *4th Israeli Workshop on Computational and Combinatorial Geometry*, Eilat, Israel, 1995.
[16] Y.S. Kupitz, Extremal problems in combinatorial geometry, M.Sc. Thesis, Aarhus University Lecture Notes Series, Vol. 53, 1979.
[17] Y.S. Kupitz, On pairs of disjoint segments in convex position in the plane, *Annals of Discrete Mathematics* 20 (1984) 203–208.
[18] H. Last, Turan Type problems for geometric graphs, M.Sc. Thesis, Technion, Israel Institute of Technology, 1994 (in Hebrew).
[19] E. Lawler, *Combinatorial Optimization: Networks and Matroids* (Holt, Rinehart and Winston, New York, 1976).
[20] L. Lovasz and M.D. Plummer, *Matching Theory*, Annals of Discrete Mathematics, Vol. 29 (North-Holland, Amsterdam, 1986).
[21] O. Marcotte and S. Suri, Fast matching algorithms for points on a polygon, *SIAM Journal on Computing* 20 (1991) 405–422.
[22] G. Owen, On the core of linear production games, *Mathematical Programming* 9 (1975) 358–370.
[23] W.R. Pulleyblank, Faces of matching polyhedra, Ph.D. Thesis, University of Waterloo, Department of Combinatorics and Optimization, 1973.
[24] G. Sabidussi, The centrality index of a graph, *Psychometrika* 31 (1966) 58–603.
[25] P. Vaidya, Geometry helps in matching, *SIAM Journal on Computing* 18 (1989) 1201–1225.
[26] B. Zelinka, Medians and peripherians of trees, *Archivum Mathematicum (Brno)* 4 (1968) 87–95.