# Polynomial Time Approximation Schemes for Metric MIN-BISECTION
# (Extended Version)

W. Fernandez de la Vega* Marek Karpinski† Claire Kenyon‡

**Abstract.** We design the first *polynomial time approximation schemes* (PTASs) for the problem of Metric MIN-BISECTION of dividing a given *finite metric* space into two halves so as to minimize the sum of distances across that partition. This settles the status of this problem which was open for some time now. Our approximation schemes depend on a novel hybrid metric placement method and a new application of linearized quadratic programs.

*LRI, CNRS, Université de Paris-Sud, 91405 Orsay. Research partially supported by the IST grant 14036 (RAND-APX), and by IST APPOL2 and by the PROCOPE project. Email: `lalo@lri.lri.fr`

†Dept. of Computer Science, University of Bonn, 53117 Bonn. Research partially supported by DFG grants, DIMACS, PROCOPE grant 31022, IST grant 14036 (RAND-APX), and Max-Planck Research Prize. Research partially done while visiting Dept. of Computer Science, Yale University. Email: `marek@cs.bonn.edu`

‡LRI, Université de Paris-Sud, 91405 Orsay. Research partically supported by the IST APPOL2 project. Email: `kenyon@lri.lri.fr`

# 1 Introduction

The MIN-BISECTION problem of dividing a given graph into two equal halfs so as to *minimize* the number of edges or the sum of their weights across the partition belongs to the most intriguing problems in the area of combinatorial optimization. The reason is that we do not know at the moment how to deal with the minimization global conditions like partitioning the sets of vertices into two halves. The MIN-BISECTION problem arises naturally in several contexts ranging from combinatorial optimization to computational geometry and statistical physics [H97]. Up to now we do not have any approximation hardness result for MIN-BISECTION cf. [BK01], thus we cannot exclude a possibility of existence of a PTAS for that problem. On other hand the best known approximation factor for that problem is $O(\log^2 n)$ [FK00].

Here we consider the metric version of that problem: we consider a finite set $V$ of points which we call vertices together with a metric $d(.,.)$ on $V$ and we ask for a partition of $V$ into two equal parts such that the sum of the distances from the points of one part to the points in the other is minimized. It is easy to see that the metric MIN-BISECTION is NP-hard in exact setting even if restricted to weights 1 and 2. In this paper we prove somewhat surprisingly that the general metric MIN-BISECTION possesses a PTAS. This answers to an open problem of [FK01].

We draw on two lines of research to develop our algorithm: One is a method of so-called exhaustive sampling for additive approximation for various optimization problems such as MAX-CUT or MAX-kSAT [AKK95], [F96], [GGR96], [FK96], [FK97], [AFKK01]. The second line we draw on connects to the previous papers on approximate algorithms for metric problems and weighted dense problems [FK01] and [FVK00].
We describe some of the new essential ideas we used in contrast to other metric papers [FK01] and [I99]. The main problem was the problem of sampling. Just as in [FK01] uniform sampling does not work, and we need to sample by picking each vertex with a probability proportional to the sum of its distances to the other vertices. This was circumvented in [FK01] by dividing each vertex into an appropriate number of "clones" and doing standard sampling on the set of clones. Then one could easily conclude from the fact that the clones of each fixed vertex go together in a maximum cut. This does not hold anymore for MIN-BISECTION (and also MAX-BISECTION) where we cannot use this cloning procedure. This is circumvented in this paper by a new method of *guessing* the placement of the outliers and a new technique of biased hybrid placements.

As mentioned before our result on existence of a PTAS for metric MIN-BISECTION is in a sense optimal, as it is easily seen, following an argument of Theorem 1 of [FK01], to be NP-hard in exact setting even if restricted to instances with weghts 1 and 2.

# 2 Organization of the Paper

The rest of the paper is organized as follows.
In Section 3, we formulate some metric lemmas which we need later. In Section 4, we give an algorithm for the Euclidean case and the analysis of its correctness. Finally, in Section 5, we construct two new PTASs for the general metric MIN-BISECTION problem.

## 3 Preliminary Results

Given a finite metric $(V, d)$, we define

$$w_x = \sum_{y \in V} d(x, y)$$

for each $x \in V$, and

$$W = \sum_{x \in V} w_x.$$

Thus, $W$ is twice the sum of all distances in $V$.

We define also for $U \subseteq V$,

$$W_U = \sum_{v \in U} w_v.$$

First, a couple of metric lemmas.

**Lemma 1**

$$d(v, u) \leq \frac{4 w_v w_u}{W}$$

**Proof**: See [FK01].

■

**Lemma 2**

$$\forall u \;\; \max_v d(u, v) \leq W/n$$

**Proof**: See [FK01].

■

The following lemma is crucial here. It shows that it suffices to obtain an additive approximation (within $\epsilon W$) to get a PTAS for metric MIN-BISECTION.

**Lemma 3** *In the metric case, the optimal value of* MIN-BISECTION *satisfies* OPT $\geq W/5$.

**Proof**: Let $X = L \cup R$ be the optimal min bisection, of value $OPT$. Let $W = \sum_{X \times X} d(x, y)$, $W_L = \sum_{L \times L} d(x, y)$, and $W_R = \sum_{R \times R} d(x, y)$. Take 2 points $x_1$ and $x_2$ at random uniformly with replacement from $L$ and 2 points $x_3$ and $x_4$ at random uniformly with replacement from $R$, and conside r the 6 edges of their induced subgraph. Then the contribution to the bisection is $a = d(x_1, x_3) + d(x_1, x_4) + d(x_2, x_3) + d(x_2, x_4)$, with expectation $4OPT/(n^2/4)$, and the contribution to $W_L + W_R$ is $d(x_1, x_2) + d(x_3, x_4)$, with expectation $(W_L + W_R)/(n^2/4)$, and satisfies:

$$\begin{aligned}
d(x_1, x_2) &\leq d(x_1, x_3) + d(x_3, x_2) \\
d(x_1, x_2) &\leq d(x_1, x_4) + d(x_4, x_2) \\
d(x_3, x_4) &\leq d(x_3, x_1) + d(x_1, x_4) \\
d(x_3, x_4) &\leq d(x_3, x_2) + d(x_2, x_4) \\
d(x_1, x_2) + d(x_3, x_4) &\leq a
\end{aligned}$$

Hence $W_L + W_R \leq 4OPT$, and so $W \leq 5OPT$.

■

# 4 A Fixed Dimension Case

We describe first the Euclidean case, when the dimension of the underlying space is fixed. Here, we describe the PTAS for MIN-BISECTION on the plane. The cases of higher but fixed dimension are similar (replacing polar coordinates by spherical coordinates).

## 4.1 The Algorithm

The algorithm is the following.

**Input:** A set $V$ of $n$ points on the Euclidean plane.

1. Scale the problem so that the average interpoint distance is equal to 1.
2. Compute $g = \sum_{x \in V} x/n$, the center of gravity of $V$.
3. If $(d(x,g), \theta(x))$ denote the polar coordinates of $x$ w.r.t. to $g$, define the domains

   $$D_{r,k} = \left\{ x \in \mathrm{R}^2 : \begin{array}{c} {\scriptstyle \epsilon(1+\epsilon)^{r-1} \leq d(x,g) < \epsilon(1+\epsilon)^r} \\ \text{and} \\ {\scriptstyle k\pi\epsilon \leq \theta(x) < (k+1)\pi\epsilon} \end{array} \right\},$$

   where $r \geq 1$ and $0 \leq k < 2\pi/\epsilon$. Let

   $$D_0 = \{x \in \mathrm{R}^2 : d(x,g) < \epsilon\}.$$

4. Construct a point (multi)set $V'$ obtained by replacing each element of $V \cap D_{r,k}$ by $y_{r,k}$, the point with polar coordinates $d(y_{r,k}, g) = \epsilon(1+\epsilon)^{r-1}$ and $\theta(y_{r,k}) = k\pi\epsilon$. Hence $y_{r,k}$ has multiplicity $m_{r,k}$ equal to the number of points in $V \cap D_{r,k}$. Moreover, each element of $V \cap D_0$ is replaced by $g$.
5. Let $s = 1 + \log_{1+\epsilon}(n/2\epsilon)$. Let $\omega_{r,k}$ denote the weighted distance from $y_{r,k}$ to $X'$:

   $$w_{r,k} = \sum_{0 \leq j \leq 2\pi/\epsilon} \sum_{0 \leq \ell \leq s} m_{j,l} d(y_{r,k}, y_{j,\ell}).$$

   Note that a partition $(L, X' \backslash L)$ of $X'$ is defined by the set of pairs of integers $(p_{r,k}, q_{r,k})$ with $q_{r,k} = m_{r,k} - p_{r,k})$ where for each $0 \leq k < 2\pi/\epsilon$ and $0 \leq r \leq s$, $p_{r,k}$ denotes the number of points in $D_{r,k}$ which belong to $L$. We do exhaustive search on all the partitions $P = \{(p_{r,k}, q_{r,k} : (r,k) \in Q\}$ with $\sum_k p_{r,k} = \frac{n}{2}(1 \pm \epsilon^2)$ and with $p_{r,k}$ satisfying $0 \leq p_{r,k} \leq m_{r,k}$ when $m_{r,k} \leq 1/\epsilon^2$, and with $p_{r,k} \in \{j \lfloor \epsilon^2 m_{r,k} \rfloor : 0 \leq j \leq 1/\epsilon^2 - 1\}$. for $m_{r,k} > 1/\epsilon^2$. Let us refer to these partitions as special partitions. Let $BP$ be the best partition found. We output the bisection obtained from $BP$ by moving from the largest side to the smallest side the necessary number of points with the smallest weights.

Note first that the last step of the algorithm can increase the value of the cut defined by $BP$ by at most $\epsilon^2 n^2$. It is thus enough to analyse $BP$.
Note also that there are $O(\log n)$ domains $D_{r,k}$. Thus the exhaustive search tests at most $(1/\epsilon^2)^{O(\log n)} = n^{O(\log(1/\epsilon))}$ distinct bisections.

### 4.2 Analysis of Correctness

Let us analyse the effect of the restrictions of the sizes of the possible intersections of each domain with each side of the cut.

Let $J$ denote the set of admissible pairs $(r, k)$. Given an optimum special partition OPT=$(p_{r,k}, m_{r,k} - p_{r,k})_{r\leq s,k\leq\pi/\epsilon}$ of $V'$, we are guaranteed that our exhaustive search tests a special partition OPT'=$(p'_{r,k}, q'_{r,k})_{r\leq s,k\leq\pi/\epsilon}$ with $|p'_{r,k} - p_{r,k}| \leq \epsilon^2 m_{r,k}$. Denote by $Q$ the set of pairs $(r, k)$ for which the inequality $\epsilon m_{r,k} \leq p_{r,k} \leq (1-\epsilon)m_{r,k}$ is satisfied. Clearly, the $y_{r,k}$ for $(r, k) \notin Q$ contribute at most $\epsilon W$ to the partition. We have thus

$$Val(\mathrm{OPT}') - Val(\mathrm{OPT}) \leq \sum_{(r,k)\in Q}\sum_{(s,\ell)\in Q} \left(p'_{r,k}q'_{s,\ell} - p_{r,k}q_{s,\ell}\right) d(y_{r,k}, y_{s,l}) + \epsilon W$$

For $(r, k) \in Q$, $(s, \ell) \in Q$ we have $|p'_{r,k} - p_{r,k}| \leq \epsilon p_{r,k}$, $|q'_{s,l} - q_{s,l}| \leq \epsilon q_{s,l}$, and so $Val(\mathrm{OPT}') - Val(\mathrm{OPT})$

$$\begin{aligned}
&\leq \sum_{(r,k)\in Q}\sum_{(s,\ell)\in Q} \left((1+\epsilon)^2 p_{r,k}q_{s,\ell} - p_{r,k}q_{s,\ell}\right) d(y_{r,k}, y_{s,l}) + \epsilon W \\
&\leq (1+3\epsilon)\sum_{(r,k)\in Q}\sum_{(s,\ell)\in Q} p_{r,k}q_{s,\ell}d(y_{r,k}, y_{s,l}) + \epsilon W \\
&\leq (1+3\epsilon)\mathrm{OPT} + \epsilon W \\
&\leq (1+8\epsilon)\mathrm{OPT},
\end{aligned}$$

the last because we know that OPT $\geq W/5$ by Lemma 3. The proof that the preliminary grouping of the vertices does not change the value of the optimum bisection by more than $O(\epsilon W)$ is similar to the proof given in [FK01] and is omitted. Thus we have a PTAS for MIN-BISECTION on the Euclidean plane.

## 5 PTASs for General Metric MIN-BISECTION

We move now to the case of arbitrary metric spaces covering all geometric spaces of arbitrary dimension. The methods of this section will generalized vastly over the methods of Section 4.

Our PTASs for metric MIN-BISECTION will make essential use of the following lemma.

**Lemma 4** *Let $t = 4\log n/\epsilon^2$. Let $V$ be a finite metric space and let $U \subseteq V$, $W_U = \sum_{u\in U} w_u$. Let $T$ be a random sample of $U$ obtained by picking each point $u \in U$ with probability $tw_u/W$. Let $v \in V$. Then,*

$$\left|\sum_{u\in U} d(v, u) - \frac{W}{t}\sum_{u\in T}\frac{d(v, u)}{w_u}\right| \leq \epsilon w_v \qquad (1)$$

*with probability at least $1 - n^{-2}$.*

**Proof**: Let $d(v,U) = \sum_{u\in U} d(v,u)$ and consider the random variable $D(v,T) = \sum_{u\in T} d(v,u)/w_u$. We have:

$$D(v,T) = \sum_{i=1}^{t} Y_i$$

where the $Y_i$ are pairwise independent and each distributed as $Y_1$ with

$$\Pr(Y_1 = \frac{d(v,u)}{w_u}) = \frac{w_u}{W} \; \forall u \in U.$$

We have that $\mathrm{E}D(v,T) = \frac{t}{W} d(v,U)$. Also $|Y_1| \leq \frac{w_v}{W}$. Thus Azuma's inequality [AZU67] gives $\Pr\left(|D(v,T) - \frac{t}{W} d(v,U)| \geq \frac{\lambda w_v}{W}\sqrt{t}\right)$

$$\leq e^{-\lambda^2/2}$$

or,

$$\Pr\left(|d(v,U) - \frac{W}{t} D(v,T)| \geq \frac{\lambda w_v}{\sqrt{t}}\right) \leq e^{-\lambda^2/2}.$$

Take $\lambda = 2\sqrt{\log n}$, $t = 4\log n/\epsilon^2$ to get

$$\Pr\left(|d(v,U) - \frac{W}{t} D(v,T)| \geq \epsilon w_v\right) \leq n^{-2}$$

■

We will also need the two following simple lemmas.

**Lemma 5** *Let $V'$ denote the set of vertices of weight less than $\epsilon^2 W$. If $V_j$ is a random subset of $V'$ obtained by picking each vertex $v \in V$ with probability $\epsilon$, then with probability at least $1-\epsilon^2$ we have: $\sum_{v\in V_j} w_v \leq 2\epsilon W$.*

**Proof**: The sum $\sum_{v\in V_j} w_v$ is dominated by the product $\epsilon^2 W.B$ where $B$ has the Binomial distribution $\mathrm{BIN}(n,\epsilon)$. The result follows by applying a Chernoff-Hoeffding Bound.

■

**Lemma 6** *Let $V'$ and $V_j$ be defined as above. Then with probability at least $1-\epsilon/10$ we have $\sum_{u,v\in V_j} d(u,v) \leq 11\epsilon^2 W$.*

**Proof**: The sum $\sum_{u,v\in V_j} d(u,v)$. has expectation bounded above by $\frac{|V_j||V_j-1|W}{2n^2} \leq 1.1\epsilon^2 W$. The result follows by using Markov inequality.

■

## 5.1 The First PTAS for the General Metric MIN-BISECTION

In this section we design and analyse our first PTAS for the general metric MIN-BISECTION. As mentioned in the introduction it builds on absolute approximation sampling methods of [AKK95], [F96], [GGR96], [FVK00], [FK01], and [AFKK01]. The crucial point here is however an introduction of a new technique combining biased metric sampling with some novel hybrid placements and partitioning method.

The following definition will be crucial in the development of this section.

**Definition 1** *Consider a partition $(L,R)$ of $V$. A multiset $T$ of vertices with multiplicities $\mu(u)$, $u \in V$ is called $(\delta,\epsilon)$-representative with respect to $P$, if for every vertex $v$ except perhaps for a subset of* **exceptional** *vertices of weight at most $\delta W$, we have, with $t=|T|$,*

$$\left|\frac{W}{t}\sum_{u\in T\cap L}\frac{\mu(u)d(u,v)}{w_u}-\sum_{u\in L}d(u,v)\right|\leq \epsilon w_v,$$

*and*

$$\left|\frac{W}{t}\sum_{u\in T\cap R}\frac{d(u,v)}{w_u}-\sum_{u\in R}d(u,v)\right|\leq \epsilon w_v.$$

*A vertex which is not exceptional is called* **normal**.

We take $t=1/\epsilon^3$, $\lambda=1/\sqrt{\epsilon}$ in Lemma 4 to get

$$\Pr\left(|d(v,U)-\frac{W}{t}D(v,T)|\geq \epsilon w_v\right)\leq e^{-1/2\epsilon}$$

Thus, for this choice of $t$, the expectation of the number of vertices for which at least one of the inequalities in definition 1 does not hold is at most $2ne^{-1/2\epsilon}$. This implies that the expectation of the weights of the corresponding vertices does not exceed $2We^{-1/2\epsilon}$ and, by Markov Inequality, with probability at least $1-\epsilon/10$, this weight does not exceed $(20/\epsilon)We^{-1/2\epsilon}\leq \epsilon^2 W$ for sufficiently small $\epsilon$. This proves the following lemma.

**Lemma 7** *Let $T$ be a random sample of $V$ with size $|T|=t$, defined as in Lemma 4 and let $(L,R)$ be an arbitrary bipartition of $V$. Let $\epsilon>0$ be sufficiently small. If $t\geq 1/\epsilon^3$, then with probability at least $1-\epsilon/10$, $T$ is $(\epsilon^2,\epsilon)$-representative with respect to $(L,R)$ and moreover, the total weight of the exceptional vertices does not exceed $\epsilon^2 W$.*

**Proof**: See above.

■

We need the following lemma.

**Lemma 8** *Let $(L,R)$ be an optimum bisection of $V$ Let $\ell=1/\epsilon$ and define a partition $V_1,V_2,...,V_\ell$ of $V$ by placing each vertex in a randomly chosen $V_j$. With probability $1-o(1)$, there exists a partition $(A,B)$ whose cost is within an additive error at most $\epsilon W$ from the optimum bisection and such that for each $j$ it satisfies*

$$||A\cap V_j|-|B\cap V_j||\leq 1. \tag{2}$$

**Proof**: Let $L_j=V_j\cap L$, $R_j=V_j\cap R$. For each $j$ we do the following:

- If $|L_j|>|R_j|$, we set $\delta_j=\lfloor\frac{|L_j|-|R_j|}{2}\rfloor$ and we move from $R$ to $L$ $\delta_j$ vertices randomly chosen in $L_j$
- If $|L_j|<|R_j|$, we set $\delta_j=\lfloor\frac{|R_j|-|L_j|}{2}\rfloor$ and we move from $L$ to $R$ $\delta_j$ vertices randomly chosen in $R_j$.

Clearly, the resulting partition satisfies to 2. Let $MV$ be the set of vertices whose positions have been changed according to the above rules and let $\Delta$ be the resulting loss in the objective function. Clearly,

$$\Delta\leq\sum_{x\in MV}w_x,$$

and so

$$\mathbf{E}(\Delta) \leq \sum_{x \in V} p_x w_x, \tag{3}$$

where $p_x$ is the probability that $x$ is moved. Fix now attention on a particular $x \in L$. (The case of $x \in R$ is exactly similar.) Assume without loss of generality that $x \in V_1$. We have that $|L_1| - 1$ has the binomial distribution with parameters $n/2 - 1$ and $p = \epsilon$. Also, $|R_1|$ has the binomial distribution with parameters $n/2$ and $p = \epsilon$. Hoeffding-Chernoff gives (see [HO63]),

$$\Pr(|L_1 - EL_1| \geq \sqrt{n \log n} \leq 2n^{-\epsilon}$$

and the analogue for $|R_1|$. Thus,

$$\Pr(|L_1 - EL_1| + |R_1 - ER_1| \geq 2\sqrt{n \log n}) \leq 4n^{-\epsilon}$$

For fixed $|L_1|$ and $|R_1|$, we have $p_x \leq \frac{|L_1 - EL_1| + |R_1 - ER_1|}{\epsilon n}$ and thus

$$\begin{aligned} p_x &\leq \frac{4}{\epsilon}\sqrt{\frac{\log n}{n}} + 4n^{-\epsilon} \\ &\leq 5n^{-\epsilon} \end{aligned}$$

and using (3),

$$\begin{aligned} \mathbf{E}(\Delta) &\leq W \max_x p_x \\ &\leq 5Wn^{-\epsilon}. \end{aligned}$$

The lemma follows now by using Markov Inequality.

■

### 5.1.1 The Algorithm

The algorithm takes as input a finite metric space $(V, d)$. It makes a series of guesses and returns, when all these guesses are correct and with probability at least 3/4 a bisection of $V$ whose cost is within $O(\epsilon W)$ from the optimum.

1. Compute vertex weights $w_v = \sum_u d(u, v)$ and total weight $W = \sum_v w_v$.
2. Let $X$ denote the set of vertices with weight $> \epsilon^2 W$ and let $V' = V \backslash X$.

   Let $\ell = 1/\epsilon$ and define a partition $V_1, V_2, ..., V_\ell$ of $V'$ by placing each vertex in a randomly chosen $V_j$.
3. Let $P_o = (L, R)$ be a bisection $(L, R)$ with value at most $\epsilon W$ from the optimum and with the property that it induces on each $V_i$ a partition whose parts sizes differ by at most one. (The existence of such a partition is guaranteed by Lemma 8.) By exhaustive search, find the partition $(X_L, X_R)$ of $X$ induced by $P_0$. Let $(L_j, R_j)$ be the partition of $V_j$ induced by $P_o$. In the next phase the algorithm will construct inductively a sequence of "hybrid" partitions $P_0, P_1, ..., P_j, ..., P_\ell$ where the first hybrid is $P_0$, the last partition $P_\ell$ is the output, and such that, for each fixed $j$, $P_j$ coincides with $P_0$ on each of the sets $V_{j+1}, V_{j+2}, ..., V_\ell$:
4. For each $j = 1, 2, \ldots, \ell$, do the following:

(a) Let $T_{j-1}$ denote a random multiset of $V$ obtained by picking $t$ times a vertex $v \in V$ according to the probabilities $tw_v/W$, $v \in V$, where $t$ is defined as in Lemma 7.

(b) By exhaustive search, guess the partition $(T'_{j-1}, T''_{j-1})$ induced on $T_{j-1}$ by $(X_L, X_R)$, $(A_1, B_1), \ldots, (A_{j-1}, B_{j-1})$, $(L_j, R_j), \ldots, (L_\ell, R_\ell)$. That is, classify the vertices of $T_{j-1}$ which are in $X, V_1, V_2, \ldots, V_{j-1}$ according to the partition being built by the algorithm, and classify the remaining vertices of $T_{j-1}$ according to the optimal partition guessed by exhaustive search.

(c) For $v \in V_j$, let

$$\hat{b}(v) = \sum_{u \in T'_{j-1}} \frac{d(u,v)}{w_u} - \sum_{u \in T''_{j-1}} \frac{d(u,v)}{w_u}.$$

(d) Construct a partition $(A_j, B_j)$ of $V_j$ by placing the $|V_j|/2$ vertices with smallest value of $\hat{b}(v)$ in $A_j$ and placing the other $|V_j|/2$ vertices in $B_j$.

Let $A = \cup_j A_j$ and $B = \cup_j B_j$.

5. Output the best of the bisections $(A, B)$ thus constructed.

### 5.1.2 The Analysis

Recall that for each $j \in \{0, ..., \ell\}$ $P_j$ is the partition which agrees with the partitions $(A_1, B_1), ..., (A_j, B_j)$ constructed by the algorithm in $V_1, ..., V_j$, and which agrees with the optimal partition $(L, R)$ in $V_{j+1}, ..., V_\ell$

We will prove that when the algorithm correctly guesses for each $j$ the partition $(T'_j, T''_j)$ induced on a random sample $T_j$ by $P_j$, then the bisection $(A, B)$ is optimal within at most $16\epsilon W$ with probability at least 3/4. The analysis will consist in showing that the increase of the objective function when changing one hybrid bisection into the next is small. We will need the following definition.

**Definition 2** *Consider a partition $P = (L, R)$ of $V$. The unbalance of a vertex $v \in V$ with respect to $P$ is the quantity*

$$\widehat{ub}(v) = \sum_{u \in L} d(u,v) - \sum_{u \in R} d(u,v).$$

**Lemma 9** *If $T_{j-1}$ is representative with respect to $P_{j-1}$, and if Lemma 5 holds, then* $\mathrm{COST}(P_j) - \mathrm{COST}(P_{j-1}) \le 16\epsilon^2 W$.

Before proving the lemma, let us first see how to use it to complete the analysis. By Lemma 7 the set $T_{j-1}$ has probability at least $1 - \epsilon/10$ of being representative with respect to $P_{j-1}$. Thus, with probability at least $1 - \ell\epsilon/10 = 9/10$, $T_{j-1}$ is representative for every $j$ and Lemma 5 holds for every $j$. Summing over $j$, we then deduce that in that case:

$$\mathrm{COST}((A,B)) - \mathrm{OPT} = \mathrm{COST}(P_\ell) - \mathrm{COST}(P_0) \le 16\epsilon W.$$

This implies with Lemma 3 a relative approximation ratio $1+95\epsilon$. To conclude the proof, it remains to verify that the result holds with the probability at least 3/4 as claimed,when all the guesses are correct, The probability that the result does *not* hold is bounded above by the sum of:

- the probability that Lemma 8 does not hold which is $o(1)$
- the probability that at least one of the samples $T_1, T_2, ..., T_\ell$ is not $(\epsilon^2, \epsilon)$-representative which is bounded above by 1/10
- the probability that Lemma 5 does not hold for at least one $j$ which is bounded above by 1/9.

The sum of these bounds is smaller than 0.25 and the claim follows

The running time is $2^{O(1/\epsilon^4)}n^2$ where the first factor accounts for the required number of exhaustive searches and $n^2$ is, within a constant factor, an upper bound for the number of operations needed for any fixed sequence of guesses. Hence, the algorithm is a PTAS for MIN-BISECTION on metric spaces.

We now proceed to prove Lemma 9. We will prove that when the algorithm correctly guesses for each $j$ the partition $P'_j$ induced on $T_j$ by $P_j$, then the partition returned by the algorithm is near-optimal with high probability. The analysis will compare, for each $j$, the cost of partition $P_j$ to the cost of partition $P_{j-1}$.

**Lemma 10** *If $T_{j-1}$ is representative with respect to $P_{j-1}$, and if Lemma 5 holds, then* $\mathrm{COST}(P_j) - \mathrm{COST}(P_{j-1}) \leq 16\epsilon^2 W$.

**Proof**: The only vertices which are classified differently in $P_{j-1}$ and in $P_j$ are vertices in $V_j$ : say, $x$ vertices are in the left side of $P_{j-1}$ and in the right side of $P_j$, and the same number $x$ of vertices are in the left side of $P_j$ and in the right side of $P_{j-1}$. Pair up these vertices in a matching $M$. For each such pair $(u, v)$, such that $P_{j-1}$ places $v$ on the right side and $u$ on the left side, let $P_{j-1}(u, v)$ denote the partition obtained from $P_{j-1}$ by swiching the sides of vertices $u$ and $v$. Note that by definition of the algorithm, $\hat{b}(u) \geq \hat{b}(v)$.
Note that the overall probability that for each $j$, $T_j$ is representative is at least 9/10, so we can assume that this is the case. Then,

$$\begin{aligned}
&\mathrm{COST}(P_j(u,v)) - \mathrm{COST}(P_{j-1}) \\
&\leq \widehat{\mathrm{ub}}(u) - \widehat{\mathrm{ub}}(v) \\
&= (\widehat{\mathrm{ub}}(u) - \widehat{\mathrm{ub}}(v)) - \frac{W}{t}(\hat{b}(u) - \hat{b}(v)) + \\
&\quad \frac{W}{t}(\hat{b}(u) - \hat{b}(v)) \\
&\leq (\widehat{\mathrm{ub}}(u) - \widehat{\mathrm{ub}}(v)) - \frac{W}{t}(\hat{b}(u) - \hat{b}(v)) \\
&\leq |\widehat{\mathrm{ub}}(u) - \frac{W}{t}\hat{b}(u)| + |\widehat{\mathrm{ub}}(v) - \frac{W}{t}\hat{b}(v)|.
\end{aligned}$$

There are two cases.
(i) If $u$ and $v$ are normal, then we use the upper bounds $|\widehat{\mathrm{ub}}(u) - \frac{W}{t}\hat{b}(u)| \leq \epsilon w_u$, $|\widehat{\mathrm{ub}}(v) - \frac{W}{t}\hat{b}(v)| \leq \epsilon w_v$.
(ii) For the total contribution of the exceptional vertices, we use the overall bound $\epsilon^2 W$ of Lemma 7. Also

$$\begin{aligned}
&\mathrm{COST}(P_j) - \mathrm{COST}(P_{j-1}) \\
&\leq \sum_{(u,v)\in M} (\mathrm{COST}(P_{j-1}) - \mathrm{COST}(P_{j-1}(u,v)) \\
&\quad +2 \sum_{u,v\in V_j} d(u,v).
\end{aligned}$$

Thus,

$$\begin{aligned}
&\text{COST}(P_j) - \text{COST}(P_{j-1}) \\
&\le 2\epsilon \sum_{u\in V_j} w_u + \epsilon^2 W + \sum_{u,v\in V_j} d(u,v) \\
&\le 4\epsilon^2 W + \epsilon^2 W + 11\epsilon^2 W,
\end{aligned}$$

the last by using Lemma 5. Thus

$$\text{COST}(P_j) - \text{COST}(P_{j-1}) \le 16\epsilon^2 W.$$

■

This completes the correctness proof of our first PTAS for the general metric MIN-BISECTION.

### 5.2 The Second PTAS for General METRIC BISECTION: The Use of Linear Programs

Our second algorithm will use an extension of linearization method for quadratic integer programs introduced in [AKK95]. For this we need a new analysis of the rounding procedure which works for arbitrary metric.

The input is finite metric space $(V, d)$ with $n$ points. We represent a bi-partition $(S, T)$ of $V$ by the vector $(x_v)$ where $x_v = 0$ if $v \in S$,$x_v = 0$ if $v \in T$. We write $d(u, B_L) = \sum_{v\in B_L} d(u, v)$. We denote by $(L, R)$ an optimum bisection. We denote by $B$ the set of vertices $v$ with $w_v \ge n^2/\log n$. The algorithm is the following

1. We guess the partition $(B_L, B_R)$ induced on $B$ by $(L, R)$. We set $U = V\backslash B$. (Guessing the placement of the vertices in $B$ is needed for the success of the rounding procedure.) and we let $t = 4\log n/\epsilon^2$. For conciseness, we write $\Delta = \sum_{u\in B_L}\sum_{v\in B_R} d(u, v)$. ($\Delta$ will be constant in most of our computations.)
2. We take first a random sample $S \subseteq V\backslash B$ by picking independently $t$ points $u_i \in V\backslash B$ according to the distribution defined by $\Pr(u_i = u) = w_u/W$. By exhaustive search, we can assume that $S \subseteq L\backslash B$
3. We define for each $v \in V$,
$$e_v = \frac{W}{t}\sum_{u\in S}\frac{d(v,u)}{w_u} + d(v, B_L) \tag{4}$$
4. We introduce the program LP(U) with variables $x_v, v \in U$,
$$\text{Minimize} \sum_{v\in U} x_v e_v + \sum_{v\in U}(1 - x_v) d(v, B_R) + \Delta$$
subject to the constrains
$$\sum_{u\in U}(1 - x_u)d(u, v) \le e_v(1 + \epsilon)$$
$$\sum_{v\in U} x_v = \frac{n}{2} - |B_L|$$
(equal sides condition) and to
$$x_v \in \{0, 1\}\ \forall v \in U$$

5. We solve the fractional relaxation of LP(U) and then use randomized rounding adopted to the unbounded metric situation to obtain an integer solution to LP(U).

The crucial difference here compared to [AKK95] is that the underlying programs are non-smooth and we must deal necessarily with the unboundeness of their coefficients. We remedy this difficulty in two ways:

(i) We use weighted sampling,

(ii) We use exhaustive search to guess the placement of the points with highest weights so that we can successfully perform the rounding within the rest.

### 5.2.1 Analysis

Assume that $S \subseteq L' = L \backslash B_L$. Then by lemma 4 the first term in the right-hand side of 4 is a good estimate of the sum of the distances from $v$ to the points in $L'$. The second term is the sum of the distances to $B_L$. Thus the $e_v$ estimate well the $d(v, L)$ in the sense that we have [assuming $S \subseteq L'$],

$$\Pr(|d(v, L) - e_v| \leq \epsilon w_v) \leq 1 - 2n^{-2}.$$

for each fixed $v$ and

$$\Pr(|d(v, L) - e_v| \leq \epsilon w_v \; \forall v \in V \backslash B) \leq 1 - 2n^{-1}.$$

Thus the $e_v$ are simultaneously all good estimates of the $d(v, L)$

We prove now the following two assertions:

(i)The optimum of the integer program $LP(U)$ gives a nearly optimum solution for MIN-BISECTION, namely to the value of the optimum solution of the following program $LPO(U)$ with variables $x_v, v \in U$,

$$\text{Minimize} \sum_{v \in U} x_v d(v, L) + \sum_{v \in U} (1 - x_v) d(v, B_R) + \Delta$$

subject to the constrains

$$\sum_{v \in U} x_v = \frac{n}{2} - |B_L|$$

(equal sides condition) and to

$$x_v \in \{0, 1\} \; \forall v \in U$$

(ii)Randomized rounding of an optimum fractional solution to $LP(U)$ provides a nearly optimal (integer) solution to $LP(U)$.

**Proof of (i)** Recall that we have, for each $v \in V$,

$$d(v, L) - \epsilon w_v \leq e_v \leq d(v, L) + \epsilon w_v$$

Recall that OPT, the value of an optimum bisection, is the value of an optimum solution to $LPO(U)$. Let VAL denote the value of a n optimum solution to $LP(U)$. Thus, with all sums being taken over $v \in U$, VAL $\leq$

$$\begin{aligned}
&\leq \min_x \left(\sum x_v e_v + \sum (1-x_v) d(v, B_R)\right) + \Delta \\
&\leq \min_x \left(\sum x_v d(v, L) + \sum (1-x_v) d(v, B_R)\right) \\
&+ \Delta + \epsilon W \\
&\leq \text{OPT} + \epsilon W \\
&\leq \text{OPT}(1+6\epsilon)
\end{aligned}$$

the last because of Lemma 3. □

**Analysis of Randomized Rounding (ii)** Let $(x_v), v \in U$ denote an optimal fractional solution to $LP(U)$ Denote by $y = (y_v), v \in U$ the result of randomized rounding applied to the $x_v$. The $y_v$ are pairwise independent random variables with

$$\Pr(y_v = 1) = x_v, \qquad \Pr(y_v = 0) = 1 - x_v, \ v \in U.$$

Let $Y = \sum_{v \in U} y_v$. Then, $\mathbf{E}(Y) = \frac{n}{2} - |B_L|$ and $\mathbf{Var}(Y) = \sum_{v \in U} x_v(1-x_v) \leq \frac{n}{4}$. Thus using Chebyshev's inequality,

$$\Pr\left(|Y - n/2 + |B_L|| \leq \delta\sqrt{n}\right) \leq 1 - \frac{1}{4\delta^2},$$

i.e., we can assume that the unbalance of the partition defined by the $y_v$ is $O(n^{1/2})$.

Let us analyse now the value of the cut $y$ given by the rounding. Let us write

$$Z = \sum_v x_v e_v - \sum_v y_v e_v = \sum_v Z_v,$$

say, with $Z_v = (x_v - y_v) e_v, \ v \in U$. Then $\mathbf{E}(Z_v) = 0$ and

$$\begin{aligned}
\mathbf{Var}(Z_v) &= e_v^2 x_v (1 - x_v) \\
&\leq e_v^2 \leq w_v^2 \\
\mathbf{Var}(Z) &= \sum_v \mathbf{Var}(Z_v) \leq \sum_v w_v^2 \\
&\leq \frac{W^2}{\log n},
\end{aligned}$$

the last because each $w_v$ is upperbounded by $\frac{W}{\log n}$ and the $w_v$ sum up to less than $W$.
We use now the following Proposition (Theorem 2.7, pp.203 in [MCD98], see also [B62]).

**Proposition** Let the random variables $X_1, X_2, \ldots X_n$ with $X_k - \mathbf{E}(X_k) \leq b$ for each $k$. Let $S_n = \sum X_k$ and let $S_n$ have expectation $\mu$ and variance $V$ (the sum of the variances of the $X_k$). Then for any $t \leq 0$,

$$\Pr(S_n - \mu) \leq t) \leq \exp \frac{-t^2}{2V(1 + \frac{bt}{3V})}$$

□

We apply this Proposition with $\mathbf{E}(Z) = 0$, $\mathbf{Var}(Z) \leq \frac{W^2}{\log n}$, and $|Z_v - \mathbf{E}(Z_v)| \leq \frac{W}{\log n}$ with probability 1 for each $v$. This gives

$$\Pr(|Z| \leq t) \quad \leq \quad 2 \exp\left(-\frac{t^2 \log n}{2W^2\,(1 + \frac{t}{3W})}\right)$$

$$\leq \quad 2\exp\left(-\frac{\epsilon^2 \log n}{3}\right)$$
$$\leq \quad 2n^{-\epsilon^2/3}$$

for $t = \epsilon W$. Thus we have that, for any fixed $\epsilon \leq 0$, the value of the cut given by the rounding is w.h.p. not smaller than the value $\sum_v x_v e_v$ of the fractional relaxation of LP(U) by more than $\epsilon W$. □

Noting that the unbalance created by the ronding can be repaired with an increase of the objective function by $O(n^{3/2})$, we see that we have a PTAS for MIN-BISECTION.

**Acknowledgements.** We thank Mark Jerrum, Uri Feige, Alan Frieze, and Ravi Kannan for stimulating remarks and discussions on the subject of this paper.

## References

[AFKK01] N. Alon, W. Fernandez de la Vega, R. Kannan, and M. Karpinski, *Random Sampling and MAX-CSP Problems*, Proc. 34th ACM STOC (2002), pp. 232-239.

[AKK95] S. Arora, D. Karger, and M. Karpinski, *Polynomial Time Approximation Schemes for Dense Instances of NP-Hard Problems*, Proc. 27th STOC (1995), pp. 284-293; the full version appeared in J. Comput. System Sciences 58 (1999) 193-210.

[AZU67] K. Azuma, *Weighted sums of certain dependent random variables*, Tôhoku Math. J. 19, 357-367.

[B62] G. Bennet, *Probability inequalities for sums of independent random variables*, Journal of the American Statistical Association 57 (1962) 33-45

[BK01] P. Berman and M. Karpinski, *Approximation Hardness of Bounded Degree MIN-CSP and MIN-BISECTION*, ECCC Technical Report, TR01-026, 2001, also in Proc. 29th ICALP (2002), LNCS 2380, Springer, 2002, pp. 623-632.

[FK00] U. Feige and R. Krauthgamer, *A Polylogarithmic Approximation of the Minimum Bisection*, Proc. 41st IEEE FOCS (2000), pp. 105–115.

[F96] W. Fernandez de la Vega, *MAX-CUT has a Randomized Approximation Scheme in Dense Graphs,* Random Structures and Algorithms 8 (1996), pp.187-198

[FVK00] W. Fernandez de la Vega and Marek Karpinski, *Polynomial time approximation of dense weighted instances of MAX-CUT*, Random Structures and Algorithms 16 (2000), pp. 314-332.

[FK01] W. Fernandez de la Vega and Claire Kenyon, *A randomized approximation scheme for metric MAX-CUT*, Journal of Computer and System Sciences 63 (2001), pp. 531-541.

[FK96] A.M. Frieze and R. Kannan, *The Regularity Lemma and Approximation Schemes for Dense Problems,* Proc. 37th FOCS (1996)i, pp. 12-20.

[FK97] A.M. Frieze and R. Kannan, *Quick Approximation to Matrices and Applications,* Combinatorica 19 (2) (1999), pp. 175-120.

[GGR96] O. Goldreich, S. Goldwasser and D. Ron, *Property Testing and its Connection to Learning and Approximation,* Proc. 37th IEEE FOCS (1996), pp. 339-348 , the full paper appeared in J. ACM 45 (1998), pp. 653-750.

[H97] D. S. Hochbaum (ed.), Approximation Algorithms for NP-Hard Problems, PWS, 1997.

[HO63] W. Hoeffding, *Probability inequalities for sums of bounded random variables*, Journal of the American Statistical Association 58 (1963) 13-30.

[I99] P. Indyk, *A Sublinear Time Approximation Scheme for Clustering in Metric Spaces,* Proc. 40th IEEE FOCS (1999), 154-159.

[MCD98] C. McDiarmid, *Concentration*, Probabilistic Methods for Algorithmic Disctrete Mathematics, M.Habib, C.McDiarmid, J. Ramirez Alfonsin, B.Reed (Eds.), Springer 1998.

[PY91] C. H. Papadimitriou and M. Yannakakis. *Optimization, Approximation and Complexity Classes*, Journal of Computer and System Sciences, 43 (1991), pp. 425-440.