ON THE OPTIMAL BISECTION OF A POLYGON

Elias Koutsoupias[1], Christos H. Papadimitriou[1,2] and Martha Sideri[3]

ABSTRACT: *We show that bisecting a polygon into two equal (possibly disconnected) parts with the smallest possible total perimeter is NP-complete, and it is in fact NP-hard to approximate within any ratio. In contrast, we give a dynamic programming algorithm which finds a subdivision into two parts with total perimeter at most that of the optimum bisection, such that the two parts have areas within ϵ of each other; the time is polynomial in the number of sides of the polygon, and $\frac{1}{\epsilon}$. When the polygon is convex, or if the parts are required to be connected, then the exact problem can be solved in quadratic time.*

1. INTRODUCTION

We study from an algorithmic standpoint the following natural geometric optimization problem: We are given a simple polygon P. We wish to subdivide P into two regions, not necessarily connected, but having the same total area, such that *the sum of perimeters of these regions is the smallest possible.* In other words, we wish to equitably cut a polygonal pie *with the smallest possible knife action.* For example, if the pie is an equilateral triangle, the optimal bisection is shown in Figure 1(a). For the irregular pie in Figure 1(b) we also show the optimal bisection. In fact, we can prove the following characterization of the optimum bisection of a polygon with c concave points:

Theorem 1. The optimal bisection consists of at most $c+1$ disjoint circural arcs orthogonal to the sides of the polygon, all of the same curvature. Moreover, at most one of them has both endpoints in the interior of the edges of the polygon and no two of them end at the same concave point. □

Here by *orthogonal* we mean something more specialized than usual: A concave point is orthogonal to any curve orthogonal to an internal tangent, a convex point to none (Figure 2). In the interiors of the sides, the definition of "orthogonal" is standard. We prove Theorem 1 in Section 2.

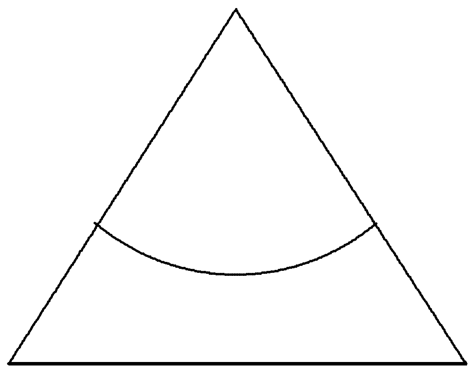
Our interest in this problem grew out of a related graph-theoretic problem, the *bisection width of solid grid graphs* [PSi]. The geometric problem has been looked at in the past, and has been recently studied intensely in the context of generalized planar separation theorems, see for example [LM]. The consensus seems to be that any such region (even a weighted one) can be bisected with total "knife action" (also appropriately weighted) proportional to the square root to its total weight. There are some very interesting results and open problems in this area. However, *there is very little on the obvious question, how to find the precise optimum in this problem.*

---

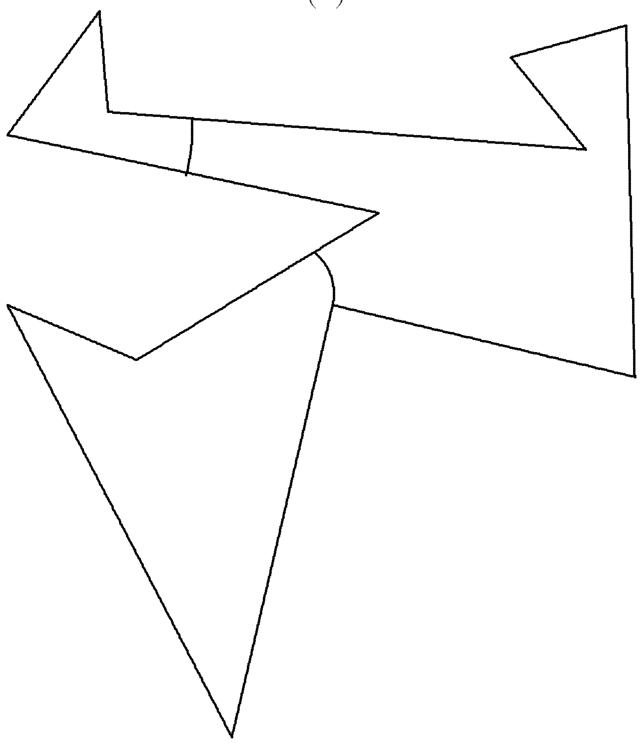
[1]University of California at San Diego.

[2]Research supported by the ESPRIT Basic Research Action No. 3075 ALCOM, a grant to the Universities of Patras and Bonn by the Volkswagen Foundation, and an NSF Grant.

[3]Computer Technology Institute, Patras, Greece.



(a)



(b)

Figure 1: Optimum Bisection of Polygons.

There is a very good reason for this:

Theorem 2. The optimal bisection problem for polygons is NP-hard. □

The proof is very easy: The NP-complete problem PARTITION can be reduced to optimal bisection as shown in Figure 3 ($a_1, \ldots, a_n$ are integers in the given instance of PARTITION). In fact, if we replace the dimension n^2 in the figure by $\lceil \frac{n}{\epsilon} \rceil^2$, we obtain:

Theorem 3: For *any* $\epsilon > 0$, it is NP-hard to obtain a solution to the optimal bisection problem (total length of curves subdividing the polygon) which is at most $1 + \epsilon$ times that optimum. □

However, we show that the only remaining hope can be realized, and in quite a strong way. Let us call a subdivision of the polygon into two regions an ϵ-*bisection* if the ratio of

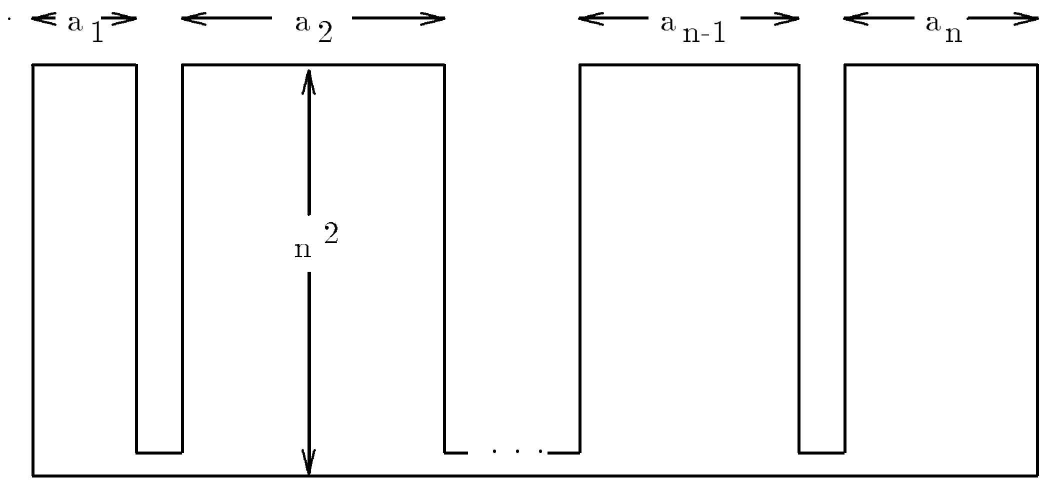


Figure 2: Reduction from PARTITION.

the areas is at most ϵ away from one.

Theorem 4: There is an algorithm which, for any given polygon with n sides, and $\epsilon > 0$, finds an ϵ-bisection which is at most as long as the optimum bisection, in time polynomial in n and $\frac{1}{\epsilon}$. $\square$

In other words, if we make the ratio of the two areas into a term of the objective function, with arbitrary weight, then a fully polynomial–time approximation scheme [PSt] is possible.

The method used to prove Theorem 4 is dynamic programming. Unfortunately, the polynomial is $O(\frac{n^5}{\epsilon^2})$. There appear to be easy ways to lower this by one or two factors of n and/or $\frac{1}{\epsilon}$, but no really practical algorithm is in sight. An additional insight into the structure of the optimal solutions (beyond Proposition 1, but weaker than Theorem 5 below, which is too strong for general polygons) will be needed in order to develop a fast approximate bisection algorithm. Naturally, our approach and its lemmata can be the basis of interesting heuristics.

For *convex polygons*, we can do much better:

Theorem 5: If the polygon is convex, the exact problem can be solved in time $O(n \log n)$ *after a preprocessing that computes the areas of the subpolygon cut off by each diagonal.* $\square$

That is, the total time is $O(n^2)$. The basic observation is that in this case, the bisection will be a single arc or segment (and thus the halves will be connected, recall Figure 1(a)). The technique is an (apparently required) binary search for each side of the polygon. Similarly for the general, non–convex problem in which the two parts are required to be connected.

2. THE OPTIMUM BISECTION.

In this section we prove the characterization of optimal bisections in Theorem 1. We first show this:

Lemma 1. The optimal bisection consists of disjoint circular arcs orthogonal to the sides of the polygon, all of the same curvature.

Proof. Take any two disjoint segments $A_1 - B_1$ and $A_2 - B_2$ of the optimum boundary of length L_1 and L_2, respectively, and consider the volumes V_1 and V_2 enclosed between these curves and the straight line segments A_1B_1 and A_2B_2 (volumes on the other side of the chords are counted as negative). By the well–known isoperimetric problem (see, for example, [BZ]), both curves must be circular arcs, say of curvature $\frac{1}{r_1}$ and $\frac{1}{r_2}$, repectively. We have, for $j = 1, 2$:

$$L_j = \frac{l\theta_j}{\sin\theta_j}$$

$$V_j = \frac{l^2(\theta_j - \sin\theta_j\cos\theta_j)}{4\sin^2\theta_j}$$

$$r_j = \frac{l}{2\sin\theta_j}$$

where $2\theta_j < \pi$ is the central angle associated with the arc α_j. It is easy to verify that

$$\frac{dL_j}{dV_j} = r_j^{-1} \tag{1}$$

and

$$\frac{d^2L_j}{dV_j^2} > 0 \tag{2}$$

that is $L_j(V_j)$ is convex. By the minimality of the optimal bisection, $L_1 + L_2$ is minimum subject to constant $V_1 + V_2$. By (1) and (2), this happens when $r_1 = r_2$.

We have shown that the boundary of the optimum bisection consists of circular arcs of the same curvature. That they are orthogonal to the perimeter follows easily by noticing that the reflection of the boundary by an edge that contains an endpoint must also be a circular arc. □

To conclude the proof of Theorem 1 we show:

Lemma 2. At most one of the arcs of the optimal bisection has both endpoints in the interior of edges of the polygon, and at every concave point of the polygon at most one arc ends.

Proof. Assume that two arcs, α_1 and α_2, have both endpoints in the interior of edges of the polygon. Let L_1,L_2 be the lengths of the arcs, let V_1,V_2 be the volumes included between the arcs and the edges of the polygon and let r_1^{-1},r_2^{-1} be the curvature of the arcs. Assume that the arcs are not straight line segments, that is r_1^{-1},$r_2^{-1} > 0$. An elementary calculation gives :

$$L_j = r_j\theta_j \ , \quad V_j = \frac{r_j^2\theta_j}{2}$$

where θ_j is the central angle associated with the arc α_j. It is easy to verify that $\frac{dL_j}{dV_j} = r_j^{-1}$ (1) and $\frac{d^2 L_j}{dV_j^2} < 0$ (2), that is $L_j(V_j)$ is concave. By the minimality of the optimal bisection, $L_1 + L_2$ is minimum subject to constant $V_1 + V_2$. By (1) and (2), this happens when $r_1 = 0$ or $r_2 = 0$, that is when one arc does not exist. A similar argument applies when one or both arcs are straight line segments.

If more than one arcs end at a concave point of the polygon, any two of them, say α_1 and α_2, form a continuous curve, which has to be a circular arc. It is obvious that one of the angles formed by α_1, α_2 and the edges of the polygon at the concave point has to be less than $\pi/2$, which contradicts the fact that α_1, α_2 are orthogonal to the concave point. $\square$

Notice that the above lemma suggests an upper bound of $c + 1$ of the number of the circular arcs of the optimal bisection, as stated in the Theorem.

3. A POLYNOMIAL ALGORITHM FOR ϵ-BISECTION

Our algorithm can be thought of as the geometric generalization of the well-known dynamic programming approximation algorithm for the PARTITION problem [PSt].

We assume, without loss of generality, that the polygon has unit area. Let $\delta > 0$ be a small parameter to be fixed later, and let $N = \lceil \frac{1}{\delta} \rceil$. Circular arcs that participate in a bisection (the optimal one, or our approximation) are of three kinds: Edge-to-edge (Figures 4(a) and 4(b)), edge-to-vertex (Figure 4(c)), or vertex-to-vertex (Figure 4(d)). Consider an ordered pair (x, y) of elements (edges or vertices, or one of each, not necessarily admissible). Each such pair has associated with it a part of the perimeter of P, namely the part traversed clockwise from x to y. We call a pair of elements (edges or concave vertices) *admissible* if there is at least one orthogonal curve between them. Each admissible pair has a *basic curve* assigned to it, namely, the one with smallest length (shown in bold lines in Figure 4). Each admissible pair has assigned to it a set of *standard curves*, of which the basic is one. These curves are linearly ordered, and the enclosed area between two consecutive ones is precisely δ (see Figure 4). So, there are $O(n^2 N)$ standard curves. Naturally, we exclude curves that intersect the boundary.

If (x, y) is admissible, consider a standard curve c from x to y. Suppose that $\alpha(c)$ is the area of the region enclosed between the curve c and the part of the boundary associated with the pair. We say that curve c has *virtual area* $A(c) = \lceil \frac{\alpha(c)}{\delta} \rceil$. Virtual area is an approximation of the true area, and is useful for our dynamic programming algorithm. Finally, we partially order pairs of edges or vertices by inclusion of the associated parts of the perimeter of P. This will be the order in which the dynamic programming recurrence will be computed.

For each pair (admissible or not) (x, y), and each integer $i \leq N$, we define $C(x, y, i)$ to be *the smallest total length of curves required in order to cut off a total virtual area of i between the curves and the part of the perimeter associated with* (x, y) (infinity if no such set of curves exists). By Lemma 2 we can write

$$C(x, y, i) = \min\{C_0(x, y, i), C_1(x, y, i),$$

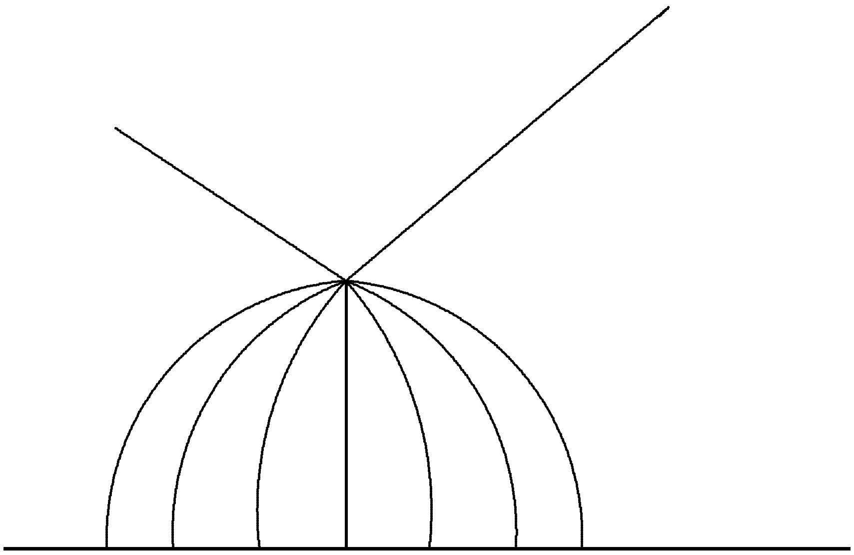


Figure 3: Standard Curves.

where $C_k(x, y, i)$, for $k = 0, 1$, is defined as the length of the shortest set of curves that isolate total virtual area of i between them and the part of the perimeter clockwise from x to y *and that include k curves between x and y*. In particular,

$$C_0(x, y, i) = \min_{j \le i, x \succ z \succ y} (C(x, z, j) + C(z, y, i - j)),$$

where z ranges over all elements (edges or concave vertices) between x and y on the part of the perimeter associated with (x, y) (Figure 5). Similarly,

$$C_1(x, y, i) = \min_{j \le N} (C_0(x, y, j) + \ell(x, y, i + j)),$$

where $\ell(x, y, i + j)$ is the length of the standard curve from x to y that has virtual area $i + j$ (Figure 6) (infinity if there is no such standard curve).

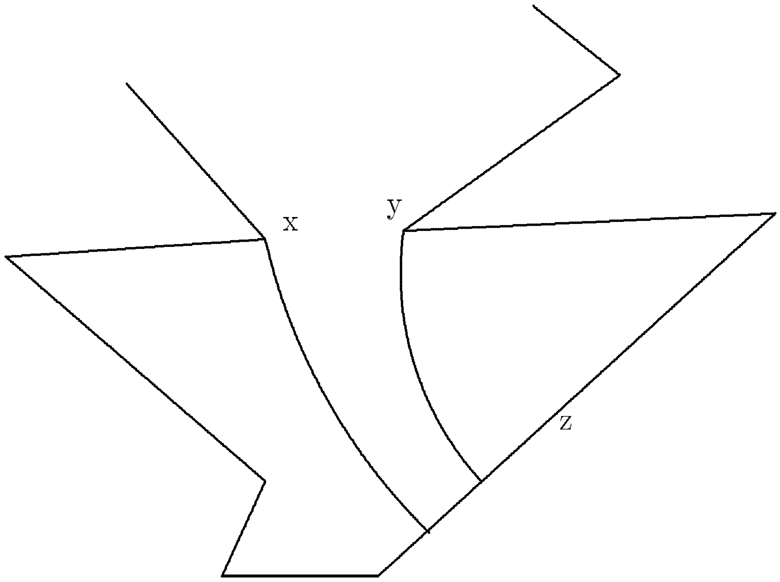


Figure 4: The calculation of C_0.

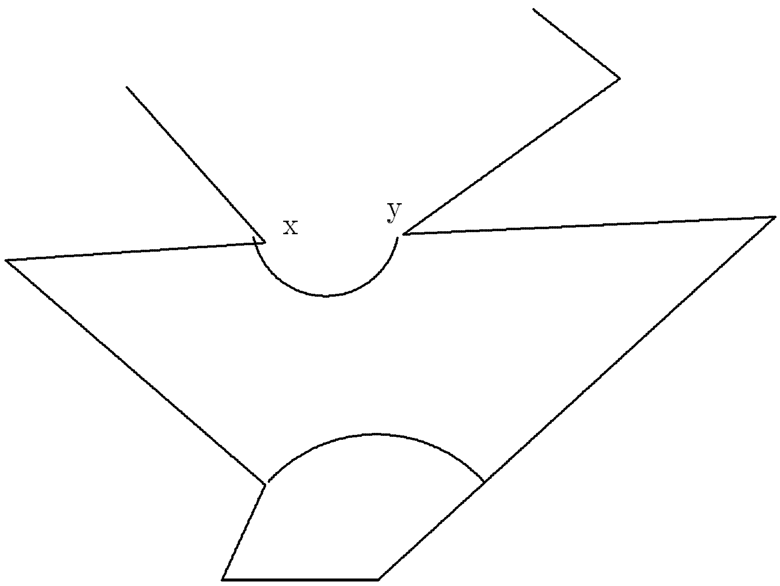


Figure 5: The calculation of C_1.

It is easy to see that the following is true:

Lemma 3. The values of C can be computed in time $O(n^3 N^2)$. $\square$

Having computed all C's, we now can find the desired approximation. It is not quite the best $C(x, y, \frac{N}{2})$, over all counterclockwise adjacent edges x, y, as one might have guessed. The reason is that we have miscalculated areas by using virtual areas. How big is this miscalculation? Since there are at most n nested curves in any solution (ours, or the optimal one), our virtual area estimates can be at most $n\delta$ away from the real area.

Consider now the optimal bisection. It contains at most n curves. We can perturb this solution, so that the curves become standard *and the overall solution is not deteriorated in length* (see Figure 7 for an example). The total deviation from a fair split (area equal to $\frac{1}{2}$) is at most $n\delta$. Hence, this perturbed solution has a virtual area value between $\frac{N}{2} - n$ and $\frac{N}{2} + n$. Hence, we pick as our solution the smallest $C(c, j)$ value for $\frac{N}{2} - n \leq j \leq \frac{N}{2} + n$. Our approximation constraint now requires that the area of the solution proposed, which is at most $n\delta$ away from half, be at most ϵ away from half. We achieve this by taking $\delta = \frac{\epsilon}{n}$. The proof of Theorem 4 is complete.

Finally, if a polygon is convex then by Theorem 1 the optimum bisection contains just one edge-to-edge circular arc. So, we are searching for the right admissible pair of edges (that is, a pair of edges that have a common orthogonal circular arc, see Figure 8). Let us suppose that we have fixed one of the two edges, e, and we are looking for the other one. This can be done by binary search. The basic observation is that the edges that make up an admissible pair are in at most four "runs" around the perimeter of the polygon (see Figure 7). A binary search determines the (at most two) edges e' such that there is a circular arc orthogonal to both e and e' bisecting the polygon. However, for this we need to have precomputed the area of any polygon cut off by a diagonal of the given one. Theorem 5 follows.

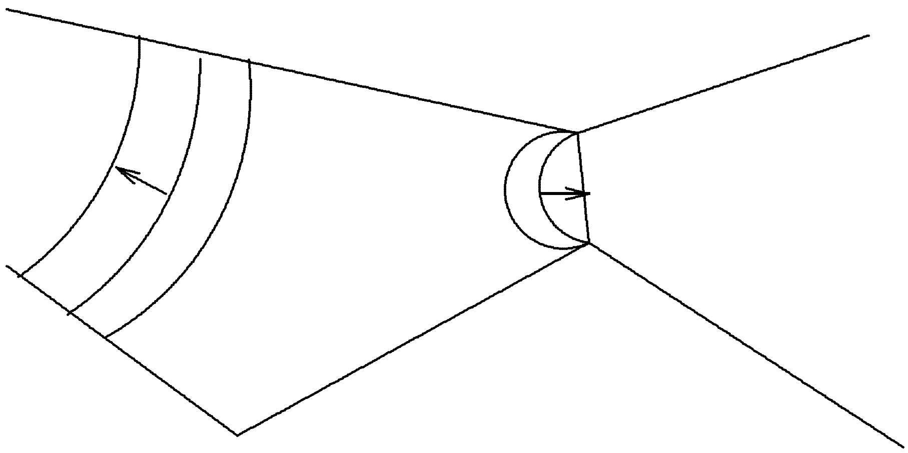


Figure 6: Perturbing the Optimal Solution.

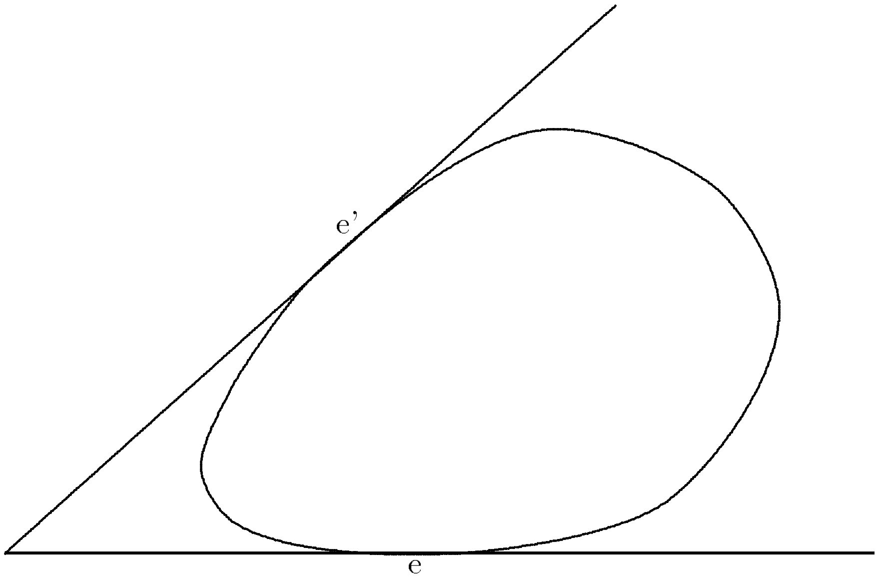


Figure 7: Proof of Proposition 1.

REFERENCES

[BZ] Vu. D. Burago, V. A. Zalgaller *Geometric Inequalities* (translated from Russian by A. B. Sosinsky) Springer-Verlag 1988.

[LM] T. Lengauer, K. Melhorn "Four Results on the Complexity of VLSI Computations," *Advances in Computing Research, vol. 2: VLSI Theory*, F. P. Preparate (ed.), JAI Press, 1984.

[PSi] C. Papadimitriou, M. Sideri "The Bisection Width of Solid Grid Graphs", *Proc. 1990 SODA*, to appear.

[PSt] C. Papadimitriou, K. Steiglitz "Combinatorial Optimization: Algorithms and Complexity," Prentice-Hall, 1982.