UCLA
COMPUTATIONAL AND APPLIED MATHEMATICS

On the Optimality of the Median Cut Spectral Bisection Graph Partitioning Method

Tony F. Chan

P. Ciarlet, Jr.

W. K. Szeto

June 1993

(Revised January 1994)

CAM Report 93-14

Department of Mathematics
University of California, Los Angeles
Los Angeles, CA. 90024-1555

ON THE OPTIMALITY OF THE MEDIAN CUT SPECTRAL BISECTION GRAPH PARTITIONING METHOD

TONY F. CHAN *, P. CIARLET, JR † AND W. K. SZETO ‡

Abstract. Recursive Spectral Bisection is a heuristic technique for finding a minimum cut graph bisection. In it, the second eigenvector of the Laplacian of the graph is computed and from it a bisection is obtained. The most common method is to use the *median* of the components of the second eigenvector to induce a bisection. We prove here that this median cut method is optimal in the sense that the partition vector induced by it is the closest partition vector, in any l_s norm, for $s \geq 1$, to the second eigenvector. Moreover, we prove that the same result also holds for any m-partition, that is a partition into m and $(n-m)$ vertices, when using the mth largest or smallest components of the second eigenvector.

Key Words. Graph partitioning, Recursive Spectral Bisection, graph Laplacian, Fiedler vector, parallel computing.

AMS Subject Classifications. 65N20, 65F10.

1. Introduction. A key problem in parallel computing is how to partition a computational problem into smaller pieces so that each piece can be mapped onto an individual processor. The objective is to have the pieces to be load balanced and the communication between the pieces as small as possible. If we use an undirected graph as a computational model, then the simplest version of this problem corresponds to finding a minimum cut bisection of the graph, which is a well known NP-complete problem.

Among the many heuristics proposed for approximately solving this problem, one of the most successful is the Recursive Spectral Bisection method first proposed by Pothen, Simon and Liou [9]. In it, the NP-hard combinatorial minimization problem of finding a partition vector with components equal to $+1$ or -1 is approximated by minimizing a quadratic form involving the *Laplacian* of the graph over a larger search space of real numbers, the solution of which reduces to finding the eigenvector associated with the smallest nonzero eigenvalue (also known as the Fiedler vector) of the Laplacian. The remaining step is then to map the Fiedler vector onto a "nearby" partition vector. The most common method is to use the *median* of the components of the Fiedler eigenvector to induce a bisection. We shall call this the *median cut* method. It has been widely used in practice [11, 13], especially for unstructured finite element meshes, and further improvements have since then been proposed [1, 4, 5]. However, a complete theoretical justification for using the median of the Fiedler vector still seems to be lacking in the literature. Barnard and Simon [1] did mention that the median cut method "chooses a partition based on the partition vector closest

* Dept of Mathematics, Univ. of Calif. at Los Angeles, Los Angeles, CA 90024. E-mail: chan@math.ucla.edu. The work of this author was partially supported by the National Science Foundation under contract ASC 92-01266, the Army Research Office under contract DAAL03-91-G-0150, and ONR under contract ONR-N00014-92-J-1890. Part of this work was completed during a visit to the Dept of Computer Science of the Chinese University of Hong Kong. Their support is also acknowledged here.

† Commissariat à l'Energie Atomique, CEL-V, D.MA, MCN, 94195 Villeneuve-St-Georges Cedex, France and Dept of Mathematics, Univ. of Calif. at Los Angeles, 405, Hilgard Avenue, Los Angeles CA 90024. E-mail: ciarlet@limeil.cea.fr, ciarlet@math.ucla.edu. The work of this author was partially supported by the DGA/DRET under contract 93-1192.

‡ Dept of Computer Science, Chinese University of Hong Kong, Shatin, Hong Kong.

to the second eigenvector" but no precise statement or proof was given. Henrickson and Leland [5] considered a multi-dimensional extension of the RSB method, in which they define the partition as the solution to a minimum cost assignment problem and solve it by an algorithm due to Tokuyama and Nakano [12]. However, it is not clear from the general algorithm in [12] that, for the bisection problem, the solution is indeed the median cut partition. On the other hand, Rendl and Wolkowicz [10] chose to formulate the bisection problem as a quadratic assignment problem in which the partition corresponds to an n by 2 matrix. A feasible solution is obtained by linearizing the objective function at the solution X of a relaxed version of the problem. Recently, Pothen [8] showed that, *up to this linearization*, a closest partition matrix to X is obtained by using a technique similar to the median cut, and extended the result to any m-partition, that is a partition into m and $(n-m)$ vertices.

In this short note, we prove that the median cut method is in fact *optimal* in the sense that the partition vector induced by it is the closest partition vector, in any l_s norm, for $s \geq 1$, to the Fiedler vector. Moreover, we extend the result to any m-partition: we prove that the partition vector obtained with respect to the mth smallest component of the Fiedler vector is also, in this more general case, the closest one to the Fiedler vector in any l_s norm, for $s \geq 1$.

2. Recursive Spectral Bisection. Suppose we are given an undirected, connected graph $G = (V, E)$, with $n = |V|$ even. Let L be the set of lattice vectors with components equal to ± 1, i.e. $L = \{l \in R^n | l_i \in \{\pm 1\}\}$. Let B be the set of load balanced vectors, defined as $B = \{b \in R^n | \sum_{i=1}^n b_i = 0\}$. We shall denote the set of bisection (i.e. load balanced partition) vectors by $P \equiv L \cap B$.

To follow the spectral strategy, we express the size of the cut set algebraically by associating a variable x_i, with each node of the graph, which may be $+1$ or -1 corresponding the two sides of the cut. The size of the cut set corresponding to a partition vector x may then be expressed as

$$|C|_{x \in P} = \frac{1}{4} \sum_{(v,w) \in E} (x_v - x_w)^2 \tag{1}$$

where the sum is over all edges (v, w) connecting vertices v and w of the graph.

Now, we define the $n \times n$ Laplacian matrix $Q = (q_{ij})$ of the graph G as

$$q_{ij} = \begin{cases} -1 & \text{if } (v_i, v_j) \in E \\ deg(v_i) & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases}$$

It is well-known that (via summation by parts)

$$x^T Q x = \sum_{(v,w) \in E} (x_v - x_w)^2$$

By relaxing the discreteness constraint of $x \in L$, the minimum cut set problem may be approximated as follows

$$\min_{x \in P} |C| = \frac{1}{4} \min_{x \in P} x^T Q x \geq \frac{1}{4} \min_{x \in B, ||x||_2^2 = n} x^T Q x = \frac{1}{4} x_2^T Q x_2 = \frac{n}{4} \lambda_2,$$

where λ_2 is the smallest nonzero eigenvalue of Q (which is positive semi-definite) and x_2 is the corresponding eigenvector (the Fiedler vector), normalized as $||x_2||^2 = n$. As x_2 is in B (because it is orthogonal to $x_1 = (1 \cdots 1)^T$), but not in L, we need to map it to a nearby vector in P. The median cut RSB method chooses the median cut vector p^m as the approximation by finding the median value of the components of x_2 and mapping values above the median to $+1$ and values below to -1. The partitions are then further partitioned by recursive application of the same procedure. Note that if the graph is not connected, then the same procedure can be applied to each of its connected components. It is well known that the multiplicity of zero as an eigenvalue of Q is equal to the number of connected components of G [7]. In this case, the Fiedler vector is the eigenvector which corresponds to the smallest positive eigenvalue.

3. Optimality of Median Cut Partitioning. The median cut vector, p^m, which is obviously feasible by construction (i.e. $p^m \in P$), has some additional favourable properties, first proved by Fiedler [3]:

1. The median cut vector guarantees that at least one of the two partitions of a connected graph is also connected.
2. Nevertheless, if x_2 has $\frac{n}{2}$ positive components and $\frac{n}{2}$ negative components, both partitions are connected.

Although p^m has the above properties, in general it does not minimize the cut $|C|_{x \in P}$. However, as x_2 minimizes the continuous version $|C|_{x \in B}$, it would be reasonable to expect that the vector in P which is closest to x_2 in some norm would be near the true optimum in P. We will show next that p^m is indeed the nearest (in any l_s norm) vector in P to the Fiedler vector x_2. Note that if we relax the condition of the partition being load-balanced, then the minimum distance partition can easily be shown to be $p_s = sign(x_2)$, i.e. p_s solves $\min_{p \in L} ||x_2 - p||_s$, where $||.||_s$ denotes the l_s norm. A partition strategy based on p_s can be found in [2].

THEOREM 3.1. *Given any $v \in R^n$, n even, let $p^m \in P$ be any median cut partition vector induced by v. Then $p^m = arg \min_{p \in P} ||v - p||_s$.*

Proof. We shall prove the theorem by showing that, for all $p \in P$, $||v-p||_s^s - ||v - p^m||_s^s \geq 0$. Without loss of generality, let $v \in R^n$, with components in non-decreasing order, i.e. $v_1 \leq v_2 \leq \cdots \leq v_n$. Then, by definition, $p^m = (-1, \cdots, -1, +1, \cdots, +1)^T$. Now, for a given p, let us define S_- as the set of indices i lower than or equal to $\frac{n}{2}$ and such that $p_i = +1$. Accordingly, define S_+ as the set of indices j greater than $\frac{n}{2}$ and such that $p_j = -1$. As $\sum p_i = 0$ and $|p_i| = 1$, it is clear that $|S_-| = |S_+|$. We now define a set S consisting of $|S_-|$ pairs (i,j) with $i \in S_-$ and $j \in S_+$, where each element of $S_- \cup S_+$ appears only once in S, and the pairing is arbitrary. Note that $p_k = p_k^m$ if $k \notin S_- \cup S_+$ and that for all $(i,j) \in S$, $v_i \leq v_j$. Then we have

$$\begin{aligned}
||v-p||_s^s - ||v-p^m||_s^s &= \sum_k (|v_k - p_k|^s - |v_k - p_k^m|^s) \\
&= \sum_{k \leq \frac{n}{2}} (|v_k - p_k|^s - |v_k - p_k^m|^s) + \sum_{k > \frac{n}{2}} (|v_k - p_k|^s - |v_k - p_k^m|^s) \\
&= \sum_{i \in S_-} (|v_i - p_i|^s - |v_i - p_i^m|^s) + \sum_{j \in S_+} (|v_j - p_j|^s - |v_j - p_j^m|^s) \\
&= \sum_{(i,j) \in S} (|v_i - 1|^s - |v_i + 1|^s + |v_j + 1|^s - |v_j - 1|^s).
\end{aligned}$$

To prove the theorem, we will show that, for all $(i,j) \in S$, $t_{ij} = |v_i - 1|^s - |v_i + 1|^s + |v_j + 1|^s - |v_j - 1|^s$ is non-negative. In the following, C_k^s denotes the binomial coefficient. We divide the proof into two parts, depending on whether s is even or odd. If s is even, then we eliminate the absolute values to find

$$\begin{aligned} t_{ij} &= (v_i - 1)^s - (v_i + 1)^s + (v_j + 1)^s - (v_j - 1)^s \\ &= \sum_{k=1}^{s}(-1)^k C_k^s v_i^{s-k} - \sum_{k=1}^{s} C_k^s v_i^{s-k} + \sum_{k=1}^{s} C_k^s v_j^{s-k} - \sum_{k=1}^{s}(-1)^k C_k^s v_j^{s-k} \\ &= 2 \sum_{k=1, k\ odd}^{s} C_k^s (v_j^{s-k} - v_i^{s-k}) \geq 0 \end{aligned}$$

because $v_i \leq v_j$ implies that $v_i^{s-k} \leq v_j^{s-k}$ for all odd values of $s-k$. For s odd, we have to be careful before we remove the absolute values. Actually, six cases have to be considered, depending on the positions of -1 and 1 with respect to v_i and v_j.
Case 1. $1 \leq v_i \leq v_j$

$$\begin{aligned} t_{ij} &= (v_i - 1)^s - (v_i + 1)^s + (v_j + 1)^s - (v_j - 1)^s \\ &= 2 \sum_{k=1, k\ odd}^{s} C_k^s (v_j^{s-k} - v_i^{s-k}) \geq 0 \end{aligned}$$

because $s-k$ is even and $|v_i| \leq |v_j|$.

Case 2. $-1 \leq v_i \leq 1 \leq v_j$

$$\begin{aligned} t_{ij} &= -(v_i - 1)^s - (v_i + 1)^s + (v_j + 1)^s - (v_j - 1)^s \\ &= -2 \sum_{k=1, k\ even}^{s} C_k^s v_i^{s-k} + 2 \sum_{k=1, k\ odd}^{s} C_k^s v_j^{s-k} \\ &= -2 \sum_{k=1, k\ odd}^{s} C_{s-k}^s v_i^{k} + 2 \sum_{k=1, k\ odd}^{s} C_k^s v_j^{s-k} \\ &= 2 \sum_{k=1, k\ odd}^{s} C_k^s (v_j^{s-k} - v_i^{k}) \geq 0 \end{aligned}$$

because $v_j^{s-k} \geq 1$ and $-v_i^k \geq -1$.

Case 3. $v_i \leq -1$ and $1 \leq v_j$

$$\begin{aligned} t_{ij} &= -(v_i - 1)^s + (v_i + 1)^s + (v_j + 1)^s - (v_j - 1)^s \\ &= 2 \sum_{k=1, k\ odd}^{s} C_k^s (v_i^{s-k} + v_j^{s-k}) \geq 0 \end{aligned}$$

because $s-k$ is even.

Case 4. $-1 \leq v_i \leq v_j \leq 1$

$$
\begin{aligned}
t_{ij} &= -(v_i-1)^s-(v_i+1)^s+(v_j+1)^s+(v_j-1)^s \\
&= 2\sum_{k=1,k\ even}^{s} C_k^s(v_j^{s-k}-v_i^{s-k})\geq 0
\end{aligned}
$$

because $v_i \leq v_j$ and $s-k$ is odd.

Case 5. $v_i \leq -1 \leq v_j \leq 1$

$$
\begin{aligned}
t_{ij} &= -(v_i-1)^s+(v_i+1)^s+(v_j+1)^s+(v_j-1)^s \\
&= 2\sum_{k=1,k\ odd}^{s} C_k^s v_i^{s-k}+2\sum_{k=1,k\ even}^{s} C_k^s v_j^{s-k} \\
&= 2\sum_{k=1,k\ even}^{s} C_k^s(v_j^{s-k}+v_i^{k})\geq 0
\end{aligned}
$$

because $v_i^k \geq 1$ and $|v_j^{s-k}| \leq 1$.

Case 6. $v_i \leq v_j \leq -1$

$$
\begin{aligned}
t_{ij} &= -(v_i-1)^s+(v_i+1)^s-(v_j+1)^s+(v_j-1)^s \\
&= 2\sum_{k=1,k\ odd}^{s} C_k^s(v_i^{s-k}-v_j^{s-k})\geq 0
\end{aligned}
$$

because $s-k$ is even and $|v_j| \leq |v_i|$. □

4. The General $m-$partition Case. We now consider the general case of an m-partition, namely a partition into m and $(n-m)$ vertices. In this case, we consider partition vectors p such that $p_i = -1$ m times and $p_i = +1$ $(n-m)$ times. Or, equivalently, p belongs to $P^m = L \cap B^m$ where $B^m = \{b \in R^n \,|\, \sum_{i=1}^n b_i = n-2m\}$. In order to get a formulation which is similar to the median cut case, we define q^m of R^n whose components are all equal to $\alpha = \frac{n-2m}{n}$ and let $\tilde{p} = p - q^m$. Then

$$
p \in P^m \Longleftrightarrow \tilde{p} \in \tilde{P}^m,
$$

where $\tilde{P}^m = \tilde{L} \cap B$, and $\tilde{L} = \{l \in R^n | l_i = \pm 1 - \alpha\}$. Then the size of the cut set is

$$
\min_{x\in P}|C| = \frac{1}{4}\min_{x\in P^m} x^TQx = \frac{1}{4}\min_{\tilde{x}\in\tilde{P}^m}\tilde{x}^TQ\tilde{x}, \text{ as } Qq^m = 0.
$$

By relaxing the discreteness constraint, $\tilde{x} \in \tilde{L}$ now becomes $||\tilde{x}||_2^2 = \frac{4m(n-m)}{n}$ and the solution to the problem in $\tilde{x}$ is the Fiedler vector x_2 normalized in the proper way. Then the solution in terms of x is $x_2 + q^m$. Here, it is crucial to note that the components of $x_2 + q^m$ and x_2 are ordered identically. From there, we can define the *m-cut* bisection method: we map the m smallest components of x_2 to -1 and the $(n-m)$ largest to $+1$ to define p^m and call it the *m-cut* vector. As in the median cut case, we prove that p^m is the closest m-partition vector to x_2 in any l_s norm, for $s \geq 1$.

THEOREM 4.1. *Given any $v \in R^n$, let $p^m \in P^m$ be any m-partition vector induced by v. Then $p^m = arg\min_{p \in P^m} ||v - p||_s$.*

Proof. Again, without loss of generality, let $v \in R^n$, with components in non-decreasing order. Then $p^m = (-1, \cdots, -1, +1, \cdots, +1)^T$. Now, for a given p of P^m, let us define S_- as the set of indices i lower than or equal to m and such that $p_i = +1$. Accordingly, define S_+ as the set of indices j greater than m and such that $p_j = -1$. It is easy to check that $|S_-| = |S_+|$. Then we can define a set S as before and compute

$$\begin{aligned} ||v-p||_s^s - ||v-p^m||_s^s &= \sum_{k \le m} \left(|v_k - p_k|^s - |v_k - p_k^m|^s\right) + \sum_{k>m} \left(|v_k - p_k|^s - |v_k - p_k^m|^s\right) \\ &= \sum_{i \in S_-} \left(|v_i - p_i|^s - |v_i - p_i^m|^s\right) + \sum_{j \in S_+} \left(|v_j - p_j|^s - |v_j - p_j^m|^s\right) \\ &= \sum_{(i,j)\in S} t_{ij}. \end{aligned}$$

From there, the remainder of the proof is elementary and identical to that of Theorem 3.1. □

5. Conclusion. This result shows that the bisection method, and more generally the m-partition, can be viewed as embedding an NP-hard minimization problem over a discrete set into a larger search space in R^n, finding the solution there by continuous (rather than discrete) algorithms and then using the nearest member of the discrete feasible set as an approximate solution to the original discrete optimization problem. The use of the Fiedler vector can be viewed as an efficient way of getting into the neighborhood of the global minimum, avoiding being trapped by local minimums. The median cut and m-cut methods are then seen as staying as close to the global minimum as possible while remaining feasible for the discrete problem. Seen in this light, it is also clear that a local search method (e.g. the Kernighan-Lin method [6]) following the partition (see e.g. [4, 5, 9]) could be much more effective than using it to find the global minimum directly.

6. Acknowledgements. We wish to thank an anonymous referee for making a remark about the m−partition case which led to the generalizations in Section 4.

REFERENCES

[1] S.T. Barnard and H. Simon. A fast multilevel implementation of recursive spectral bisection for partitioning unstructured problems. Technical Report RNR-92-033, NASA Ames, NAS Systems Division, Applied Research Branch, NASA Ames Research Center, Mail Stop T045-1, Moffett Field, CA 94035, November 1992. A short version is in Proc. 6th SIAM Conf. on Parallel Processing for Scientific Computing, 1993, pp.711-718.

[2] T. F. Chan and W.K. Szeto. A sign cut version of the recursive spectral bisection graph partitioning algorithm. In J.G. Lewis, editor, *Proc. of the 5th SIAM Conf. on Applied Linear Algebra, Snowbird, UT.*, pages 562–566. SIAM, Philadelphia, June 1994.

[3] M. Fiedler. A property of eigenvectors of nonnegative symmetric matrices and its application to graph theory. *Czechoslovak Math. J.*, 25(100):619–633, 1975.

[4] B. Hendrickson and R. Leland. An improved spectral graph partitioning algorithm for mapping parallel computations. Technical Report SAND92-1460, Sandia Nat'l Lab., Albuquerque, N.M., September 1992.

[5] B. Hendrickson and R. Leland. Multidimensional spectral load balancing. Technical Report SAND93-0074, Sandia Nat'l Lab., Albuquerque, N.M., January 1993.

[6] B. Kernighan and S. Lin. An efficient heuristic procedure for partitioning graphs. *Bell Systems Tech. Journal*, 29:291–307, 1970.

[7] B. Mohar. The laplacian spectrum of graphs. In *6th International Conference on Theory and Applications of graphs*, pages 871–898, 1988.

[8] A. Pothen. An analysis of spectral graph partitioning via quadratic assignment problems. In *Domain decomposition methods in science and engineering, Proc. of the 7th Int'l Conf. on Domain Decomposition*. The American Mathematical Society, Providence, RI, 1995.

[9] A. Pothen, H. Simon, and K.P. Liou. Partitioning sparse matrices with eigenvector of graphs. *SIAM J. Mat. Anal. Appl.*, 11(3):430–452, 1990.

[10] F. Rendl and H. Wolkowicz. A projection technique for partitioning the nodes of a graph. Technical Report CORR 90-20, University of Waterloo, Canada, 1990. To appear in Annals of Operations Research.

[11] H. Simon. Partitioning of unstructured problems for parallel processing. *Computing Systems in Engineering*, 2(2/3):135–148, 1991.

[12] T. Tokuyama and J. Nakano. Geometric algorithms for a minimum cost assignment problem. *Proc. 7th Annual Symp. on Comp. Geometry, ACM*, pages 262–271, 1991.

[13] R. Williams. Performance of dynamic load balancing algorithms for unstructured mesh calculations. *Concurrency*, 3:457–481, 1991.