ELSEVIER

Parallel Computing 22 (1997) 1493–1502

PARALLEL COMPUTING

Short communication

A conjugate gradient method for the spectral partitioning of graphs

N.P. Kruyt [1]

Department of Mechanical Engineering, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands

Received 20 April 1995; revised 7 July 1996

Abstract

The partitioning of graphs is a frequently occurring problem in science and engineering. The spectral graph partitioning method is a promising heuristic method for this class of problems. Its main disadvantage is the large computing time required to solve a special eigenproblem. Here a simple and efficient method is proposed to reduce this computing time. This method is based on the conjugate gradient minimization method. The convergence properties of the new method are studied for the case of regular one-, two-, and three-dimensional grids. The influence of the aspect ratio of the graph on the convergence rate is also investigated.

Keywords: Graph partitioning; Spectral bisection; Ordering algorithms; Eigenvalue problems; Conjugate gradient minimization

1. Introduction

Many important scientific and engineering problems can be formulated as *graph partitioning* problems, i.e. the division of the vertices of a graph into sets of specified sizes in such a way that the sum of the weights of the edges that cross between the sets is minimized. Examples of such graph partitioning problems are the mapping of parallel computations [12,23], the ordering of sparse matrix computations [1,22] and the laying out of VLSI-circuits [4].

It has been shown that the graph partitioning problem is NP-hard [8]. Therefore a great variety of heuristic methods has been developed (e.g. [18]). A promising heuristic

[1] Email: n.p.kruyt@wb.utwente.nl.

PII S0167-8191(96)00059-2

method is the spectral partitioning technique that was recently proposed by Pothen et al. [23]. With this method a partition is obtained from the eigenvector corresponding to the second-smallest eigenvalue of the Laplacian matrix of the graph [5,6].

Generally the method produces good quality partitions, but its disadvantage is formed by the large computing time required for the solution of the eigenproblem. This paper describes a simple method to reduce this computing time.

The outline of this paper is as follows. Firstly, the basic ideas of the spectral method are briefly described. This leads to an eigenvalue problem of which a specific solution is required. Then the conjugate gradient method for eigenvalue problems is given and its application to graph partitioning is described in detail. The method is applied to some test cases in order to study its convergence properties. Finally, conclusions and recommendations for further research are formulated.

2. Spectral graph partitioning method

An undirected connected graph $G = (V,E)$ is defined by its vertex set V and its edge set E. A positive scalar weight $w(e_k)$ is associated with each edge e_k. The graph partitioning problem (in its simplest form) is the division of the vertices into two disjoint sets with an equal number of vertices such that the sum of the weights of the edges that connect vertices in the two different sets is minimized.

The mathematical formulation of the problem is as follows. Let each vertex be assigned a number x_i $(i = 1,\ldots,|V|)$ equal to -1 or 1. The graph partitioning problem can then be formulated as a constrained discrete minimization problem [17,23]:

$$\min \boldsymbol{x} \cdot \boldsymbol{L} \cdot \boldsymbol{x} \tag{1}$$

$$\text{sub } \boldsymbol{e} \cdot \boldsymbol{x} = 0 \tag{2}$$

$$\text{sub } x_i = \pm 1 \tag{3}$$

Herein the Laplacian matrix $\boldsymbol{L}$ of the graph is defined as

$$\boldsymbol{L} = \boldsymbol{D} - \boldsymbol{A} \tag{4}$$

where the adjacency matrix $\boldsymbol{A}$ is given by:

$$A_{ij} = \begin{cases} 1 & \text{if } \{v_i \leftrightarrow v_j\} \in E, \\ 0 & \text{otherwise,} \end{cases} \tag{5}$$

and the vertex degree matrix $\boldsymbol{D}$ is defined by:

$$D_{ij} = \begin{cases} \sum_k A_{ik} & \text{if } i = j, \\ 0 & \text{otherwise,} \end{cases} \tag{6}$$

and the vector $\boldsymbol{e}$ in (2) is the vector consisting of all ones: $\boldsymbol{e} = (1,\ldots,1)^T$.

It can be shown that the Laplacian matrix $\boldsymbol{L}$ is symmetric and non-negative definite. For the case of connected graphs considered here, the multiplicity of the smallest eigenvalue $\lambda = 0$ is one and the corresponding eigenvector is $\boldsymbol{e}$.

The bold step made in the spectral graph partitioning method is to replace the discrete optimization problem by an approximate continuous problem: the condition that $x_i = \pm 1$ is replaced by the normalizing constraint:

$$\boldsymbol{x} \cdot \boldsymbol{x} = |V|. \tag{7}$$

The solution of the minimization problem (1), (2) and (7) is the eigenvector corresponding to the *second-smallest* eigenvalue of the Laplacian matrix $\boldsymbol{L}$ [23]:

$$\boldsymbol{L} \cdot \boldsymbol{x} = \lambda \boldsymbol{x}. \tag{8}$$

Properties of the second-smallest eigenvalue of the Laplacian matrix $\boldsymbol{L}$ and the corresponding eigenvector have been studied in [5,6].

Recently an alternative spectral partitioning method has been proposed [27]. This method leads to the problem of determining the second-smallest eigenvalue of the generalized eigenvalue problem:

$$\boldsymbol{L} \cdot \boldsymbol{x} = \lambda \boldsymbol{D} \cdot \boldsymbol{x}. \tag{9}$$

In the sequel only the eigenproblem (8) is considered.

The partition of the graph is determined by computing the median μ of the elements of the eigenvector $\boldsymbol{x}$ corresponding to the second-smallest eigenvalue and assigning vertices to the sets by comparing the vector elements x_i with μ. Alternative partitions that are derived from the eigenvector $\boldsymbol{x}$ are proposed in [16].

The problem of computing eigenvalues of matrices is a standard problem in mathematics and engineering. A vast literature exists on this subject (e.g. [11,29,30]). The Laplacian matrix $\boldsymbol{L}$ generally will be sparse, since it is determined by a large graph. Therefore iterative methods that take advantage of this sparsity are attractive. Almost all previous studies of the spectral graph partitioning method have used the Lanczos algorithm [19] for this purpose.

The solution of the eigenproblem by standard techniques takes up the major part of the computing time of the spectral graph partitioning method. Since only a single eigenvalue and the corresponding eigenvector have to be determined, a more efficient and specialized method is appropriate. This method could use the additional knowledge that the smallest eigenvalue is zero and its eigenvector is $\boldsymbol{e}$. The remainder of this paper deals with a method to accomplish this.

3. Conjugate gradient method for spectral graph partitioning

A method that is especially suited to the determination of the smallest (or largest) eigenvalue and its corresponding eigenvector is the conjugate gradient method for symmetric eigenvalue problems [3,20,25]. The basic idea of this method is to minimize the Rayleigh quotient by means of the conjugate gradient method. The well-known conjugate gradient method is a general method for minimizing functions; a lucid account of the method is given in [24]. For symmetric eigenproblems the function to be minimized is the Rayleigh quotient (e.g. [11,29]):

$$F(\boldsymbol{x}) = \frac{\boldsymbol{x} \cdot \boldsymbol{L} \cdot \boldsymbol{x}}{\boldsymbol{x} \cdot \boldsymbol{x}} \tag{10}$$

The gradient of the Rayleigh quotient, which is required in the conjugate gradient method, is given by

$$\nabla F(x) = \frac{2}{x \cdot x}(L \cdot x - F(x)x). \tag{11}$$

3.1. Basic conjugate gradient minimization algorithm

The algorithm for the Fletcher–Reeves variant of the conjugate gradient method [7] for the minimization of a function $F(x)$ is summarized by:

$$g_0 := -\nabla F(x_0) \tag{12}$$

$$h_0 := g_0 \tag{13}$$

for $k := 1$ **to** K **do**

$$\alpha_k := \min_{\alpha} F(x_{k-1} + \alpha h_{k-1}) \tag{14}$$

$$x_k := x_{k-1} + \alpha_k h_{k-1} \tag{15}$$

$$g_k := -\nabla F(x_k) \tag{16}$$

$$h_k := g_k + \frac{g_k \cdot g_k}{g_{k-1} \cdot g_{k-1}} h_{k-1} \tag{17}$$

endfor

$$x := x_k \tag{18}$$

3.2. Conjugate gradient minimization method for eigenproblems

The smallest eigenvalue of a symmetric matrix can be found by minimizing the Rayleigh quotient over all x [11]. The second-smallest eigenvalue can be obtained by minimizing the Rayleigh quotient over all x that are perpendicular to the eigenvector corresponding to the smallest eigenvalue.

Since this eigenvector equals e, this means that the minimization has to be performed over all x with $x \cdot e = 0$. A key observation is that if the initial estimate x_0 for the eigenvector is perpendicular to e, then all estimates x_k produced by the conjugate gradient minimization method are perpendicular to e. This is shown in the sequel. Therefore the conjugate gradient minimization method for the determination of the smallest eigenvalue can be used to determine the second-smallest eigenvalue! This is simply accomplished by using an initial estimate that is perpendicular to e.

Here the main modifications of the basic conjugate gradient algorithm of the preceding section are given for the case that the function to be minimized is the Rayleigh quotient. These modification concern the line minimization in (14) and the efficient computation of the gradient in (16). Given the specific form of the Rayleigh quotient, the line minimization can be written as:

$$\min_{\alpha} F(x + \alpha h) \equiv \min_{\alpha} \frac{x \cdot L \cdot x + 2(x \cdot L \cdot h)\alpha + (h \cdot L \cdot h)\alpha^2}{x \cdot x + 2(x \cdot h)\alpha + (h \cdot h)\alpha^2}. \tag{19}$$

Hence the unknown α for which the Rayleigh quotient attains its minimum can be determined from a quadratic equation. Note that the term $\boldsymbol{x} \cdot \boldsymbol{L} \cdot \boldsymbol{x}$ equals $F(\boldsymbol{x})\boldsymbol{x} \cdot \boldsymbol{x}$. Therefore only the single matrix–vector product $\boldsymbol{L} \cdot \boldsymbol{h}$ has to be computed to perform the line minimization.

The direct evaluation of the gradient from (11) would require an extra matrix–vector multiplication. An efficient way to compute the gradient is obtained by taking into account the update relation (15). This leads to the following update relation for the vector $\boldsymbol{g}$:

$$\boldsymbol{g}_k := \frac{2}{\boldsymbol{x}_k \cdot \boldsymbol{x}_k}\left(F_k \boldsymbol{x}_k - F_{k-1}\boldsymbol{x}_{k-1} + \frac{1}{2}(\boldsymbol{x}_{k-1} \cdot \boldsymbol{x}_{k-1})\boldsymbol{g}_{k-1} - \alpha_k \boldsymbol{L} \cdot \boldsymbol{h}_{k-1} \right). \tag{20}$$

If $\boldsymbol{x}_0 \cdot \boldsymbol{e} = 0$, then it follows from (11–13) and $\boldsymbol{L} \cdot \boldsymbol{e} = 0$ that $\boldsymbol{g}_0 \cdot \boldsymbol{e} = 0$ and $\boldsymbol{h}_0 \cdot \boldsymbol{e} = 0$. This implies, using (15,17,20), that $\boldsymbol{x}_k \cdot \boldsymbol{e} = 0$, $\boldsymbol{g}_k \cdot \boldsymbol{e} = 0$ and $\boldsymbol{h}_k \cdot \boldsymbol{e} = 0$. This proves the key observation mentioned before that if the initial estimate for the eigenvector corresponding to the second-smallest eigenvalue is perpendicular to $\boldsymbol{e}$, then all other estimates produced by the conjugate gradient minimization method are also perpendicular to $\boldsymbol{e}$.

The complete algorithm of the conjugate gradient method for the spectral partitioning of graphs is given in the appendix. Three vector updates, one matrix–vector and 6 vector–vector multiplications have to be performed per iteration. The memory requirements are low: only four vectors need to be stored.

4. Test cases

In this section some test cases are studied to determine the convergence properties of the method.

Attention is focused on graphs associated with regular one-dimensional, two-dimensional and three-dimensional grids, and on a more general three-dimensional finite-element mesh.

The regular grid cases were chosen since analytical solutions were given in [23] and they enabled an investigation of the dependence of the convergence rate on the size of the problem. Furthermore, the aspect ratio of the grids can easily be modified in order to study its influence on the convergence rate.

The test cases considered are that of the M three-point grid, $M \times M$ five-point grid and $M \times M \times M$ seven-point grid. The dimension of the grid was varied: for the one-dimensional grid $M = 100$, 625, 5625; for the two-dimensional grids $M = 10$, 25, 75, 200, 500, 1000; for the three-dimensional grids $M = 5$, 10, 20, 35, 65, 100. Unless stated otherwise, the initial estimate recommended in [23] was used: $x_i = i - (|V| + 1)/2$. A typical plot of the convergence history is shown in Fig. 1 for the 500×500 five-point grid, which shows the *relative error* in the second-smallest eigenvalue as a function of the number of iterations. For a relative error of 10^{-6}, the method converged in 313 iterations with the standard initial estimate; with three random initial estimates the method converged in 1817 iterations on average.

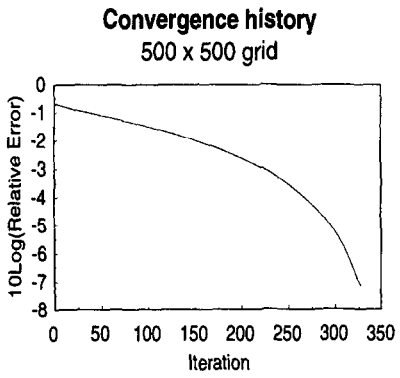

Fig. 1. Convergence history for the 500 × 500 five-point grid.

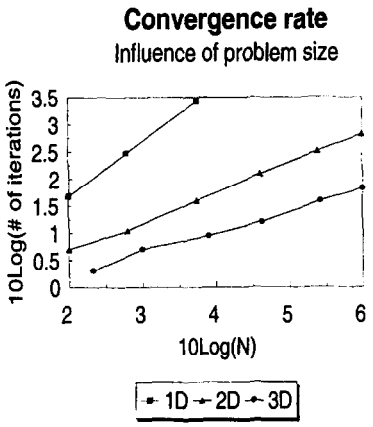

Fig. 2. Influence of the problem size on the convergence rate for 1D, 2D and 3D regular grids.

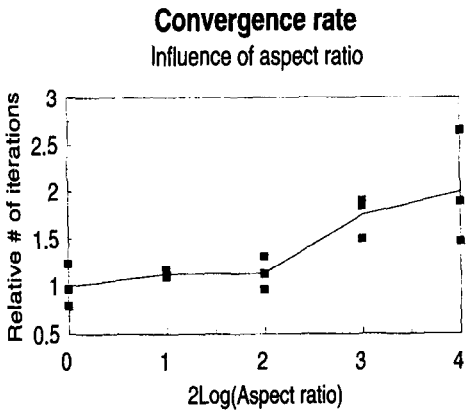

Fig. 3. Influence of the aspect ratio of the grid on the convergence rate, relative to the average of the case of a square grid, for two-dimensional grids of circa 25000 points.

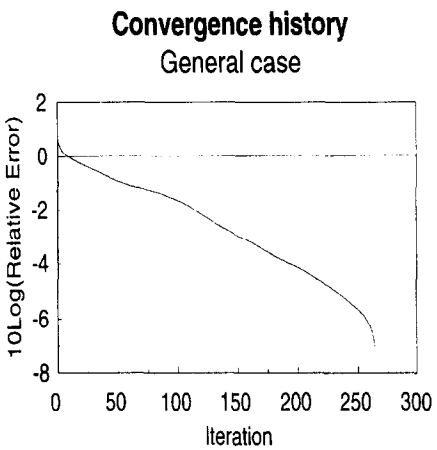

Fig. 4. Convergence history for the general "brack2" mesh.

An analysis of the influence of the size of the problem on the convergence rate is given in Fig. 2. A (very strict) relative error of 10^{-7} was used. Fig. 2 indicates that the number of iterations that is required to obtain the second-smallest eigenvalue to a specified accuracy is proportional to $N^{1/D} \equiv M$, where N is the dimension of the matrix and D is the spatial dimension of the problem.

So far only square grids have been considered. In order to study the influence of the aspect ratio of the grids on the convergence rate, the second-smallest eigenvalues of $M_1 \times M_2$ five-point grids have been determined for grids consisting of a total of circa 25000 points with varying aspect ratio M_1/M_2. Three random initial estimates were used. The result of the individual runs as well as the average is shown in Fig. 3. For grids that do not deviate significantly from the square shape, the required number of iterations increases only by a moderate factor.

The more general test case considered corresponds to the three-dimensional finite-element mesh (62631 nodes and 366554 edges) "brack2" from the Harwell–Boeing collection. A plot of the convergence history is shown in Fig. 4.

5. Conclusions and recommendations

The use of conjugate gradient minimization methods is proposed in conjunction with the spectral graph partitioning method. This leads to a simple and efficient method to determine the second-smallest eigenvalue of the Laplacian matrix of a graph and the corresponding eigenvector.

For the regular grids that were studied, the required number of iterations is proportional to $N^{1/D}$, where N is the dimension of the matrix and D is the spatial dimension of the problem.

The effect of the aspect ratio of the grids on the convergence rate was investigated. For grids that do not deviate significantly from the square shape, the required number of iterations increases only by a moderate factor.

As mentioned before, most previous studies of the spectral bisection method use the Lanczos method for determining the second-smallest eigenvalue. For this reason the conjugate gradient minimization and the Lanczos method are briefly compared. Intricacies of the Lanczos algorithm in the context of spectral partitioning, such as the choice of full or (what sort of) selective orthogonalization, are discussed in [15] and [23]. The Lanczos algorithm can suffer from the phenomenon of "misconvergence" [21], which appears not to be present with the conjugate gradient minimization (see Figs. 1 and 4). A significant advantage of the conjugate gradient minimization is that it directly gives the eigenvector associated with the second-smallest eigenvalue. Only this eigenvector is used in bisecting the graph. Furthermore, the memory requirements of the conjugate gradient minimization are very low in comparison with the Lanczos method.

The following recommendations are made for future research:

- The application of the spectral partitioning method to nodal and element resequencing methods deserves further research. An advantage of the spectral resequencing method over other heuristic methods (e.g. [9,10,26]) is that it generalizes simply to the case that the nodes do not have equal weights [12].
- It is suggested to adapt the present method to the generalized eigenproblems that arise in some formulations of the spectral technique [27,28].
- This paper does not deal with the spectral quadrisection and octasection methods [12,13]. The extension of the present method to these cases of spectral graph partitioning deserves further study.
- A critical comparison of the merits of recently proposed multilevel methods for graph partitioning [2,14] in comparison with conjugate gradient methods for spectral graph partitioning is recommended. Presently it is not clear whether the (theoretical) convergence rate of these methods outweigh their overhead and complexity in comparison with the simple conjugate gradient methods, especially for three-dimensional graphs. The conjugate gradient methods could be further accelerated by appropriate preconditioning techniques (e.g. [25]).

Appendix A. Algorithm of the conjugate gradient method for spectral partitioning of graphs

The algorithm of the conjugate gradient method for the spectral partitioning of graphs is summarized by

$$x_0 := x_o - \left(\frac{e \cdot x_0}{e \cdot e}\right) e$$

$$\rho_0 := x_0 \cdot x_0$$

$$y_0 := L \cdot x_0$$

$$F_0 := \frac{x_0 \cdot y_0}{\rho_0}$$

$$g_0 := \frac{2}{\rho_0}(F_0 x_0 - y_0)$$

$$h_0 := g_0$$

$$\beta_0 := g_0 \cdot g_0$$

for $k := 1$ **to** K **do**

$$y_k := L \cdot h_{k-1}$$

$$p_k := \rho_{k-1} \qquad q_k := 2x_{k-1} \cdot h_{k-1} \qquad r_k := h_{k-1} \cdot h_{k-1}$$

$$s_k := \rho_{k-1} F_k \qquad t_k := 2x_{k-1} \cdot y_k \qquad u_k := h_{k-1} \cdot y_k$$

$$\alpha_k := \frac{-(u_k p_k - s_k r_k) + \sqrt{(u_k p_k - s_k r_k)^2 - (u_k q_k - t_k r_k)(t_k p_k - s_k q_k)}}{u_k q_k - t_k r_k}$$

$$F_k := \frac{s_k + t_k \alpha_k + u_k \alpha_k^2}{p_k + q_k \alpha_k + r_k \alpha_k^2}$$

$$x_k := x_{k-1} + \alpha_k h_{k-1}$$

$$\rho_k := x_k \cdot x_k$$

$$g_k := \frac{2}{\rho_k}\left(F_k x_k - F_{k-1} x_{k-1} + \frac{1}{2}\rho_{k-1} g_{k-1} - \alpha_k y_k\right)$$

$$\beta_k := g_k \cdot g_k$$

$$h_k := g_k + \frac{\beta_k}{\beta_{k-1}} h_{k-1}$$

endfor

$$x := x_k$$

References

[1] S.T. Barnard, A. Pothen and H.D. Simon, A spectral algorithm for envelope reduction of sparse matrices, Tech. Rept. CS-93-49, Dept. of Computer Science, University of Waterloo, Waterloo, Ontario, Canada, 1993.

[2] S.T. Barnard and H.D. Simon, A fast multilevel implementation of recursive spectral bisection for partitioning unstructured problems, in: R.F. Sincovec, D.E. Keyes, M.R. Leuze, L.R. Petzold and D.A. Reed, eds., *Proc. 6th SIAM Conf. on Parallel Processing for Scientific Computing* (SIAM, Philadelphia, PA, 1993) 711–718.

[3] W.W. Bradbury and R. Fletcher, New iterative methods for the solution of the eigenproblem, *Numer. Math.* 9 (1966) 259–267.

[4] A.E. Dunlop and B.W. Kernigan, A procedure for placement of standard-cell VLSI-circuits, *IEEE Trans. Computer-Aided Design* 4 (1985) 92–98.

[5] M. Fiedler, Algebraic connectivity of graphs, *Czechoslovak Math. J.* 23 (1973) 298–305.
[6] M. Fiedler, A property of eigenvectors of nonnegative symmetric matrices and its application to graph theory, *Czechoslovak Math. J.* 25 (1975) 619–633.
[7] R. Fletcher and C.M. Reeves, Function minimization by conjugate gradients, *Comput. J.* 14 (1974) 149–153.
[8] M.R. Garey and D.S. Johnson, *Computers and intructability: A Guide to NP-Completeness* (Freeman, San Francisco, CA, 1979).
[9] A. George and J.W.H. Liu, *Computer Solution of Large Sparse Positive-Definite Systems* (Prentice-Hall, Englewood Cliffs, NJ, 1981).
[10] N.E. Gibbs, N.G. Poole and P.K. Stockmeyer, An algorithm for reducing the bandwidth and profile of a sparse matrix, *SIAM J. Numer. Anal.* 13 (1976) 236–250.
[11] G.H. Golub and C.F. Van Loan, *Matrix Computations* (The Johns Hopkins University Press, Baltimore, MD, 1983).
[12] B. Hendrickson and R. Leland, An improved spectral graph partitioning for mapping parallel computations, *SIAM J. Sci. Comput.* 16 (1995) 452–469.
[13] B. Hendrickson and R. Leland, Multidimensional spectral load balancing, Tech. Rept. SAND93-0074, Sandia National Laboratories, Albuquerque, NM, 1993.
[14] B. Hendrickson and R. Leland, A multilevel algorithm for partitioning graphs, Tech. Rept. SAND93-1301, Sandia National Laboratories, Albuquerque, NM, 1993.
[15] B. Hendrickson and R. Leland, The Chaco user's guide version 2.0, Tech. Rept. SAND95-2344, Sandia National Laboratories, Albuquerque, NM, 1995.
[16] S.S. Hsieh, G.H. Paulino and J.F. Abel, Recursive spectral algorithms for automatic domain partitioning in parallel finite element analysis, *Comput. Meth. Appl. Mech. Engineering* 121 (1995) 137–162.
[17] Y.F. Hu and R.J. Blake, Numerical experiences with partitioning of unstructured meshes, *Parallel Computing* 20 (1994) 815–829.
[18] B. Kernigan and S. Lin, An efficient heuristic procedure for partitioning graphs, *Bell System Tech. J.* 29 (1970) 291–307.
[19] C. Lanczos, An iteration method for the solution of the eigenvalue problem of linear differential equations and integral operators, *J. Res. Nat. Bur. Stand.* 45 (1950) 255–282.
[20] D.E. Longsine and S.F. McCormick, Simultaneous Rayleigh-quotient minimization methods for $Ax = \lambda Bx$, in: A. Björck, R.J. Plemmons and H. Schneider, eds., *Large Scale Matrix Computations* (Elsevier, New York, 1981) 195–234.
[21] B.N. Parlett, H. Simon and L.M. Stringer, On estimating the largest eigenvalue with the Lanczos algorithm, *Math. Comp.* 38 (1982) 153–165.
[22] G.H. Paulino, I.V.M. Menezes, M. Gattass and S. Mukherjee, Node and clement resequencing using the Laplacian of a finite element graph, *Internat. J. Numer. Methods Engineering* 37 (1994) 1511–1555.
[23] A. Pothen and H.D. Simon and K.P. Liou, Partitioning sparse matrices with eigenvectors of graphs, *SIAM J. Matrix Anal.* 11 (1990) 430–452.
[24] W.H. Press, B.P. Flannery, S.A. Teukolsky and W.T. Vetterling, *Numerical Recipes in C* (Cambridge University Press, Cambridge, UK, 1988).
[25] F. Sartoretto, G. Pini and G. Gambolati, Accelerated simultaneous iterations for large finite element eigenproblems, *J. Comp. Physics* 81 (1989) 53–69.
[26] S.W. Sloan, An algorithm for profile and wavefront reduction of sparse matrices, *Internat. J. Numer. Methods Engineering* 23 (1986) 239–251.
[27] R. Van Driessche and D. Roos, An improved spectral bisection algorithm and its application to dynamic load balancing problems, *Parallel Computing* 21 (1995) 29–48.
[28] R. Van Driessche and D. Roos, A spectral bisection algorithm for constrained graph partitioning I: The bisection case, Rept. TW216, Dept. of Computer Science, Katholieke Universiteit Leuven, Leuven, Belgium, 1994.
[29] J.H. Wilkinson, *The Algebraic Eigenvalue Problem* (Oxford University Press, Oxford, UK, 1965).
[30] J.H. Wilkinson and C. Reinsch, *Handbook for Automatic Computation, Vol. II. Linear Algebra* (Springer, New York, NY, 1971).