HIERARCHICAL PROBING FOR GENERAL GRAPHS, A METHOD FOR COMPUTING diag($F(A)$).

JESSE LAEUCHLI * AND ANDREAS STATHOPOULOS *

One problem that occurs frequently in numerical linear algebra is the computation of the function of matrices that are too large to calculate directly. One such function is the trace of the inverse of A, which occurs in Lattice Quantum Chromodynamics,(LCQD), data mining, statistics, and uncertainty quantification. Several approaches have been proposed for this problem before. In the case of small matrices, a factorization approach can solve the problem exactly, but this becomes impractical for many matrices of interest, due to size. For larger matrices, purely statistical approaches are popular, since they allow for a bounded estimate which can be iteratively improved [1]. Another approach that can be taken is probing[2]. In many matrices, the magnitude of the elements of $f(A)_{i,j}$ decrease inversely with the graph theoretical distance between nodes i, j of A. Therefore, if we compute a k-distance coloring of A or its equivalent, a distance-1 coloring of A^k, we can recover the most important elements of diag($f(A)$). The value of the nodes that share the same color can then be recovered by creating a probing vector consisting of all ones for nodes sharing the same color, and zeros everywhere else. Because of the block structure induced by coloring, the diagonal can be recovered using only n vectors, where n is the number of colors used to color A^k. Since the structure of A^k approximates the structure of $f(A)$, if an iterative solver is used, these probing vectors can also be used to recover an approximation to the trace of $f(A)$.

A major shortcoming of this approach is the memory and computational time required to compute and store A^k which can increase very quickly as k grows, depending on the connectivity of the problem. Additionally, if the graph of the associated matrix is highly connected, as is the case in examples arising in social network analysis, A^k may quickly become completely dense. In this case the number of colors used will go from an amount insufficient to produce a good estimate, to assigning a unique color per node, which is equivalent to solving for every diagonal element individually. Finally, if after computing a trace approximation with a given k the accuracy of the trace computation is too low, a higher k must be selected, and the approximation recomputed. With classical probing, this means that the results of all the previously preformed solves must be discarded, since the intersection between sets of probing vectors for the two levels of colors is likely to be empty.

To address this problem we first consider approaches from clustering algorithms. Clustering turns out to be exactly the opposite problem to the one we are attempting to solve. Clustering seeks to find groups of nodes which have a high degree of intergroup communication. While there are several approaches to this, many of them are expensive and cannot be applied to very large graphs. One of the solutions to this is to merge nodes which are close together, as in the METIS algorithm[3], thus producing a smaller, coarser version of the graph. Once the graph is small enough, a more expensive clustering algorithm can be applied. One popular choice for this is spectral clustering, which works by taking the eigenvectors of the k smallest eigenvalues of the Laplacian of the graph, and using these as points in k-dimensional space, which are then clustered. This approach is popular since it provides bounds on how far the

*Department of Computer Science, College of William and Mary, Williamsburg, Virginia 23187-8795, U.S.A.(laeuchli@cs.wm.edu and andreas@cs.wm.edu)

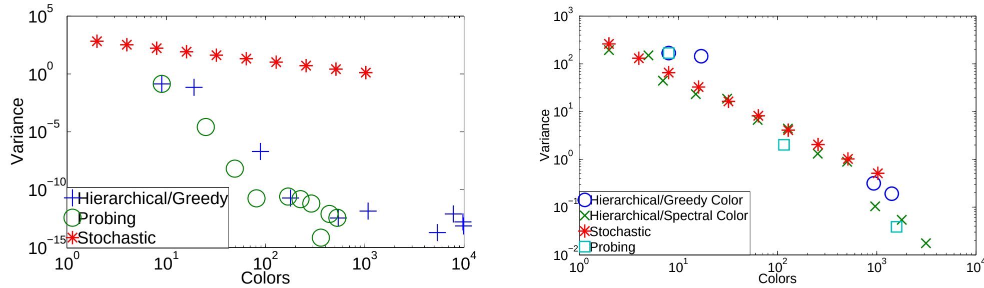

Fig. 0.1: Comparison of classical probing, multi-level probing, and stochastic methods on a synthetic Covariance matrix(where classical probing preforms very well), and on a p2p social network graph, where probing preforms poorly. Our algorithm is competitive with probing, while outperforming statistical methods in most cases

clustering can be from optimal.

The problem of finding a coloring for A^k is opposite to the problem of clustering. We want to find groups of nodes that have limited communication between themselves. However, the approaches used in clustering can be adapted to solve this problem. Our approach is similar to Walshaw's[5] multilevel algorithm. We build up a hierarchy of coarse graphs, and assign colors to each node at every level. The colors at each level are used to build up a final coloring for every node. This is done by treating the color at each level as a digit in a mixed radix representation of an integer, the integer being the final color the node receives, based on its membership in a color at each level of the hierarchy. We produce the coarse graph by merging nodes and their neighborhoods based on color. At each level we preform a distance-1 coloring. Then two nodes which are distance 2 away from each other and share a color are merged. This is done for all nodes, producing the next level of the graph. Not only does this avoid forming and storing the full A^k, it ensures that the colorings produced at each stage are hierarchical, meaning that the work done at one stage will be a subset of the work at the next, ensuring that no computation is wasted. The algorithm annihilates distance 2^k at the k-th level, corresponding to preforming probing on A^{2k} . In many of our experiments the number of colors needed by multi-level coloring is close to the same number needed by classical probing at the same level, while having the advantage of producing more intermediate steps than probing. This is useful, because it allows the user to terminate the algorithm early if the desired accuracy is achieved.

Finally, the algorithm can be improved in some cases by modifying the coloring strategy. While greedy coloring at each level is sufficient for many matrices for which probing works well, when probing works poorly other approaches can be applied. In particular, we may want to assign fewer colors than produced by greedy coloring. This can be achieved by grouping the nodes most weakly connected into the desired number of colors, information which can be obtained though spectral methods. We have observed that for some test cases arising from social networking graphs, this approach can provide an improvement over greedy coloring.

REFERENCES

[1] H. Avron and S. Toledo, *Randomized algorithms for estimating the trace of an implicit symmetric positive semi-denite matrix*, Journal of the ACM, 58 (2011)

[2] J. Tang and Y. Saad,*A probing method for computing the diagonal of the matrix inverse*, Numerical Lin. Alg. Appl. 19(3) 485-501 (2012)

[3] G. Karypis and V.Kumar, *A Fast and Highly Quality Multilevel Scheme for Partitioning Irregular Graphs*, 20(1) 359392 1999

[4] S. Liu, *Multi-way dual Cheeger constants and spectral bounds of graphs*, Advances in Mathematics 268(2) 306338 (2015)

[5] C. Walshaw, *A Multilevel Approach to the Graph Colouring Problem*, ISBN-13: 978-1899991709, 2001