

Dr. Constantine Bekas  
Postdoctoral Associate

Computer Science & Engineering Dept., University of Minnesota, Twin Cities

## RESEARCH EXPERIENCE AND STATEMENT

### General Interests

My general research interests are in the area of algorithms in numerical linear algebra, for the solution of large scale problems in computational materials science and in mining of high dimensional data. The computation of eigenvalues and singular values of very large and sparse matrices holds central role in the above setting. The computational framework for these problems involves high performance parallel systems as well as highly and widely distributed networks of computers. I strongly believe that it is necessary to incorporate advances in algorithms for these problems as well as recent innovations in computing hardware and software techniques in user friendly software interfaces. Thus, I consider research in problem solving environments to be of significant importance.

### Large Scale Eigenvalue & Singular Value Problems

Eigenvalue and Singular value calculations hold a central role in an exceptionally broad spectrum of scientific disciplines. Large scale eigenvalue problems that involve very sparse and large matrices are very demanding. Even though that the field has enjoyed significant progress in the last twenty years, we are still far from “black-box” algorithms that can be safely and efficiently used as their counterpart algorithms for smaller, dense matrices (as in the widely used LAPACK library).

My research efforts in this field involve Krylov subspace methods in the framework of implicitly restarted techniques [9]. The immense difficulty of large scale problems, especially when we are interested in the eigenvalues in the interior of the spectrum, call for combining several different techniques. I strongly believe in the importance of unifying algorithmic frameworks, that seamlessly allow for the integration of advances in our theoretical understanding of the problem.

A very important new technique, that is totally different from traditional Krylov methods, is the Automated Multilevel Substructuring method (AMLS) [8]. AMLS is a multilevel extension of Component Mode Synthesis methods and has been demonstrated to solve extremely large eigenproblems in Structural Engineering, allowing for unprecedented simulations potentially revolutionizing this scientific field. In the recent paper [5], together with Y. Saad we were able to conduct a purely algebraic analysis of the approximation mechanism of AMLS. This novel understanding has lead to important improvements that greatly increase the convergence properties of the method, thus allowing its use in other applications, such as electronic structure calculations, with more stringent accuracy demands than Structural Engineering. Currently, AMLS is a one-shot algorithm in the sense that certain approximate eigenvectors are built and no further refinements are made. The current framework does not iteratively refine these approximations. I am currently exploring the feasibility of an iterative scheme based on AMLS. Preliminary results ([6], [3]) indicate that one can devise an AMLS based method that adopts shifting strategies of shift-and-invert Lanczos, resulting in a highly robust method. I firmly believe that substructuring techniques for the eigenvalue problem is a very promising line of research with significant potential impact on many application areas.

### Computational Materials Science

Perhaps one of the greatest achievements of modern science and technology is the ability to master the properties of materials. Indeed, innovative materials have revolutionized many aspects of modern life. Undoubtedly, this dramatic progress could not have happened without the corresponding revolution of computing systems. However, the ever growing needs of humanity for lighter, stronger and cheaper materials will surely push our computing capabilities to their limits. Therefore, new algorithmic techniques that will allow us to attack larger problems in computational materials science are much needed.

My main goals in this line of research is to exploit high performance computers for solving large scale problems that arise in modelling of real-life materials, i.e., materials beyond simple crystals or artificial mod-

els. During my post-doc collaboration with prof. Y. Saad (CS Dept.) and prof. J. Chelikowsky (Chem. Eng. & Materials Science Dept.) we are investigating new algorithms for excited states of various nanostructures. We have focused primarily on methods based on time-dependent density functional theory (TDDFT). My contribution involves a) better eigenvalue algorithms (such as **AMLS**, see above) and b) methods which avoid eigenvalue problems altogether. Concerning the first approach, apart from **AMLS** we are proposing to exploit the particular structure of density (or occupation) matrices, both in real space and in the Fourier space as well [7]. Our preliminary results indicate considerable improvement over standard methods. Concerning the second approach, our research has focused on using expressions of the density matrix in bases other than the eigenbasis, which are much easier to compute than eigenvectors [4]. This methodology is similar to the popular Order-N methods but has the potential to yield much better accuracy while at the same time having no applicability restrictions as Order-N methods appear to have.

It is very important to stress that research in computational materials science is most likely to benefit from advances in algorithms that are suited for large scale distributed computation. Thus, I consider that in the (near) future enabling wide area parallel calculations (i.e. computational **GRIDS**) in materials science research will be of crucial importance and thus one of my central research goals.

### **Numerical Methods in Mining of High Dimensional Data**

The revolution in communications and data acquisition has dramatically increased the volume of available data. However, this wealth of information is useless unless it is properly analyzed. In recent years we have witnessed a significant extension of the application of linear algebra algorithms and techniques in areas such as Information Retrieval, Pattern Recognition and Data Mining. Latent semantic indexing, support vector machines and principal components analysis are only a few of the many methodologies based on linear algebra. My research interests in this field involve new (numerical) algorithms for clustering high dimensional data as well as new efficient techniques for classification. I consider equally important the scalability of the methods in distributed computational environments as well as the development of software environments (problem solving tools—see below) that will provide a user-friendly interface to the back end algorithms, so that practitioners will be able to achieve a better understanding of the data.

### **Problem Solving Environments and High Performance Parallel & Distributed Computing**

Advances in algorithmic techniques as well as in mathematical modelling are indeed essential in solving demanding problems. However, in recent years researchers in computer science have recognized the need for software that will allow transparent access to the above theoretical advances, combined with equally transparent access to computational platforms, especially ones that facilitate distributed computation. These problem solving (software) environments (PSEs) enable practitioners to describe problems in an abstract framework. The goal is to free researchers from programming details and thus allow them to concentrate on the problem itself. In addition, such tools should be able to efficiently harness the full potential of the computational infrastructure available. Automatically ensuring scalability and efficiency in a wide range of applications is a particularly exciting and significant problem which is central in modern high performance computing research around the world.

My research interests in this important area has so far concentrated on the development of a parallel problem solving environment for pseudospectra computations [1], [2]. This tool is based on **MATLAB** and recent advances that allow Message Passing parallel computing in **MATLAB**.

Through my collaboration with researchers from the Department of Chemical Eng. & Materials Science at the University of Minnesota, I was able to appreciate the potential impact of PSEs in materials science research. The development of such a tool that will combine graphic illustration and molecule design, wealth of models and algorithms with direct access to highly (and widely) distributed computation resources is bound to enjoy great success.

### **PhD Thesis Research**

My PhD thesis work was in the area of pseudospectra calculation, that is a problem with a well documented extreme computational cost, which is closely related to eigenvalue and singular value calculations of large and sparse (non-normal) matrices. My advisor was prof. Dr. E. Gallopoulos in the University of Patras, Greece. In order to address the problem I have developed parallel geometric path following algorithms as

well as parallel iterative methods for the efficient approximation of the smallest singular value of large and sparse shifted matrices  $(A - z_k I)$  for many complex shifts  $z_k$ . (see my C.V. for a complete list of relevant publications). My overall goal was to combine the above approaches in a unifying algorithmic and software framework that will help pseudospectra to become a useful tool for researchers and engineers.

### **Interest in Collaborations & Interdisciplinary Research**

I am particularly interested in continuing my research efforts at an institution that encourages interdisciplinary collaboration for the solution of a variety of problems of science and engineering. I am especially looking forward to working in an environment that nurtures and rewards scientific achievement. Furthermore, I find particularly appealing the prospect of working on real world problems that originate from the industry as well as from the life sciences.

I plan to establish a research group in the areas of my interest. My goal is to train graduate students in developing state-of-the art (numerical) algorithms and linear algebra techniques on high performance computing platforms for the solution of challenging problems in these research areas. It is my conviction that the computational challenges of our time demand a new breed of scientists that will be able to combine rigorous mathematical modelling with advanced computer science skills.

### **Contact Information:**

Dr. Constantine Bekas  
4-192, EE/CS BLD.  
Computer Science & Engineering Dept.  
University of Minnesota, Twin Cities  
200 Union St. SE, 55455, Minneapolis, USA  
email: [bekas@cs.umn.edu](mailto:bekas@cs.umn.edu)  
WWW: [www.cs.umn.edu/~bekas](http://www.cs.umn.edu/~bekas)  
Tel: +1 612 626 7510  
Fax: +1 612 625 0572

## **References**

- [1] C. Bekas, E. Kokiopoulou, and E. Gallopoulos. The design of a distributed PSE for computing Pseudospectra. *To appear in Future Generation Computer Systems, Elsevier, 2004.*
- [2] C. Bekas, E. Kokiopoulou, I. Koutis, and E. Gallopoulos. Developing a cluster based matlab environment for computing pseudospectra. In *GRACM Congress on Computational Mechanics, 2002.*
- [3] C. Bekas and Y. Saad. Amls, spectral schur complements and iterative computation of eigenvalues. *PMAA 04, Marseilles, France, 2004.*
- [4] C. Bekas and Y. Saad. Avoiding large scale eigenvalue problems in electronic structure calculations. *In preparation., 2004.*
- [5] C. Bekas and Y. Saad. Computation of smallest eigenvalues using spectral schur complements. *To appear in SIAM J. Scient. Comput., 2004.*
- [6] C. Bekas and Y. Saad. Accurate computation of eigenvalues via multiple-shift amls. *In preparation., 2005.*
- [7] C. Bekas, Y. Saad, and J. Chelikowsky. Exploiting structure in large scale eigenvalue problems in electronic structure calculations. *In preparation., 2004.*
- [8] J. K. Bennighof and R. B. Lehoucq. An automated multilevel substructuring method for eigenspace computation in linear elastodynamics. *SIAM. J. Sci. Comput., 25(6), 2004.*

- [9] E. Kokiopoulou, C. Bekas, and E. Gallopoulos. Computing smallest singular triplets with implicitly restarted Lanczos bidiagonalization. *J. Appl. Num. Math.*, 49(1):39–61, 2004.