LA-UR-95- 437

Conf-950439--9

Los Alamos National Laboratory is operated by the University of California for the United States Department of Energy under contract W-7405-ENG-36.

TITLE: POLYNOMIAL APPROXIMATIONS FOR MATERIALS SIMULATIONS

AUTHOR(S): R. N. Silver, T-11
A. F. Voter, T-12
J. D. Kress, T-12
H. Röder*

SUBMITTED TO: Invited Talk Presented at Simulation Multi-Conference: High Performance Computing '95, Phoenix, AZ, April 9-13, 1995

*Dept. of Physics, Universitat Bayreuth,
W-8850 Bayreuth, Germany

Los Alamos
Los Alamos National Laboratory
Los Alamos, New Mexico 87545

FORM NO. 836 R4
ST. NO. 2629 5/81

POLYNOMIAL APPROXIMATIONS FOR MATERIALS SIMULATIONS *

R. N. Silver[1], A. F. Voter[1], J. D. Kress[1], H. Röder[2]
[1] Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545
[2] Dept. of Physics, Universitat Bayreuth, W-8850 Bayreuth, Germany
e-mail correspondence: rns@loke.lanl.gov

Keywords: Physics, Statistical Analysis, Electronic Structure, Molecular Dynamics, Materials

Abstract

Chebyshev polynomial approximations are an efficient and numerically stable way to calculate properties of the very large Hamiltonians important in computational materials science. We describe kernel polynomial methods (KPM) producing estimates for densities-of-states (DOS) which are strictly positive and simple convolutions of known broadening functions, or kernels, with true DOS. The methods are demonstrated for tight binding electronic structure calculations of Si, yielding rapid convergence of cohesive and vacancy formation energies. KPM are also applicable to dynamical spectral functions, statistical mechanics, and density matrices.

1 Introduction

In recent papers (Silver and Röder 1994; Silver *et al.* 1995), we proposed kernel polynomial methods (KPM) for estimating properties of very large Hamiltonian matrices. The density of states (DOS) is approximated by an expansion in Chebyshev polynomials, using moments calculated by Hamiltonian matrix-on-vector multiplications (MVMs) according to the Chebyshev recursion relation. The resulting DOS estimate equals the true DOS convoluted with a broadening function, or *kernel*. The energy resolution is uniform in a definable sense and inversely proportional to the number of moments. Applications of KPM to computational materials science cited in (Silver *et al.* 1995) include the thermodynamics of Heisenberg antiferromagnets, the many body DOS of the Holstein t-J Model, the optical spectra of quantum dots, the dynamical magnetic susceptibility of disordered magnets, etc. KPM are numerically stable, algorithmically simple, and can take advantage of the same optimized MVM algorithms used Lanczos recursion methods (LRM). The cpu and memory required can scale linearly in the number of states for sparse Hamiltonians provided finite energy resolution and statistical accuracy are acceptable. KPM have been demonstrated for Hamiltonian dimensions as large as $2^{26} \times 2^{26}$.

This paper focuses on a single application of KPM to the tight binding electronic structure of Si. We demonstrate how rapid convergence and high precision can be obtained for the cohesive energy and vacancy formation energy. We conclude with some comparisons of KPM to other methods for calculating DOS and spectra.

2 The Kernel Polynomial Method

Consider the calculation of the DOS of an $N \times N$ Hamiltonian $\mathbf{H}$, defined as

$$D(\varepsilon) = \frac{1}{N} \sum_{n=1}^{N} \delta(\varepsilon - \varepsilon_n) \quad , \tag{1}$$

where ε_n are eigenenergies. Scale the Hamiltonian by $\mathbf{H} = a\mathbf{X} + b$ so that all the eigenvalues of $\mathbf{X}$ satisfy $-1 \leq x_n \leq +1$. The scaled DOS can be represented by a polynomial expansion,

$$D(x) = \frac{1}{N} \sum_{n=1}^{N} \delta(x - x_n)$$

$$= \frac{1}{\pi\sqrt{1-x^2}} \left[\mu_0 + 2 \sum_{m=1}^{\infty} \mu_m T_m(x) \right] \quad . \tag{2}$$

The $T_m(x)$ are Chebyshev polynomials of the first kind defined by $T_0(x) = 1$, $T_1(x) = x$ and the recursion relations,

$$T_{m+1}(x) = 2xT_m(x) - T_{m-1}(x) \quad . \tag{3}$$

*Published in *Simulation MultiConference '95 Proceedings, High Performance Computing '95*

They are orthogonal satisfying

$$\int_{-1}^{1} \frac{1}{\sqrt{1-x^2}} T_m(x) T_n(x) dx = \frac{\pi}{2} \delta_{m,n} \quad \{m, n \geq 1\} \quad . \tag{4}$$

This expansion may be reexpressed in terms of trigonometric functions using $x = \cos(\phi)$ and $T_m(x) = \cos(m\phi)$, so that Eq. (2) is analogous to a Fourier expansion. The Chebyshev polynomial moments of the DOS are

$$\mu_m \equiv \int_{-1}^{1} T_m(x) D(x) dx \quad . \tag{5}$$

They may be generated using the Chebyshev recursion relations and MVMs,

$$T_m(\mathbf{X})|i> = 2\mathbf{X} T_{m-1}(\mathbf{X})|i> - T_{m-2}(\mathbf{X})|i> \quad , \tag{6}$$

where $|i>$ are basis states. Then moments are constructed using

$$\mu_{2m} = \sum_{i=1}^{N} [2 < i|T_m(\mathbf{X}) T_m(\mathbf{X})|i> - 1] \quad , \tag{7}$$

and a similar relation for μ_{2m-1}, such that evaluating M moments requires $M/2$ MVMs.

Practical calculations will yield only a finite number of moments, but abrupt truncation of Eq. (2) would result in unwanted Gibbs oscillations. KPM consider instead smooth truncations of the form

$$D_K(x) = \frac{1}{\pi\sqrt{1-x^2}} \left[\mu_0 g_0 + 2 \sum_{m=1}^{M} \mu_m g_m T_m(x) \right] \quad . \tag{8}$$

The g_m are *Gibbs damping factors* which depend implicitly on M. The relation of this estimate to the true DOS is

$$D_K(x) = \int_{-1}^{1} K(x, x_o) D(x_o) dx_o \quad , \tag{9}$$

where the *kernel polynomial* is

$$K(x, x_o) = \frac{1}{\pi\sqrt{1-x^2}} \left[g_0 + 2 \sum_{m=1}^{M} g_m T_m(x) T_m(x_o) \right] \quad . \tag{10}$$

More generally, the *kernel polynomial approximation* to any function $F(x)$ defined on $x \in [-1, +1]$ may be defined as follows. Switch variables to $\phi = \arccos(x)$ such that $\widetilde{F}(\phi) \equiv F(x)\sin(\phi)$. Extend the domain to $\phi \in [0, 2\pi]$ by invoking π-antiperiodic boundary conditions, $\widetilde{F}(2\pi - \phi) = \pm\widetilde{F}(\phi)$. Define the π-antiperiodic function

$$\delta_K(\phi) \equiv \frac{1}{2\pi} \left[g_0 + 2 \sum_{m=1}^{M} g_m \cos(m\phi) \right] \quad . \tag{11}$$

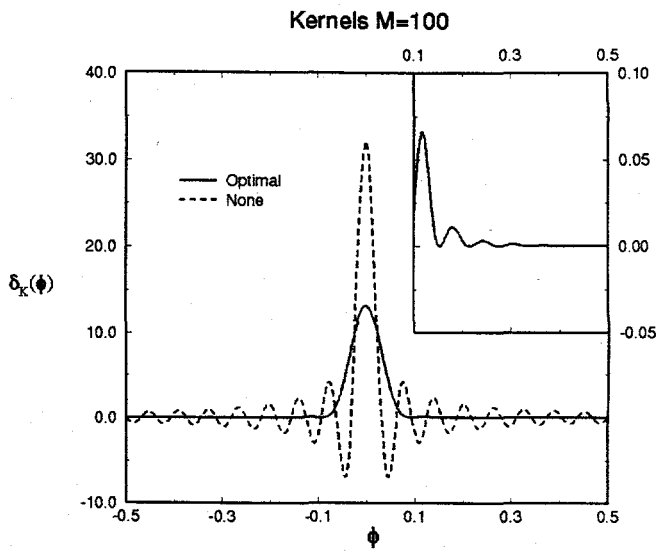

Figure 1: **Kernel Polynomials** - Dashed curve (*None*) is the kernel $\delta_K(\phi)$ for no Gibbs damping. Solid curve (*Optimal*) is the kernel for the uniform approximation by polynomials. The inset uses an expanded vertical scale to display the leakage and damping of Gibbs oscillations at large $|\phi|$. Calculations are for 100 moments.

This should be regarded as a π-antiperiodic polynomial approximation to a Dirac delta function. It is normalized to $\int_{-\pi}^{\pi} \delta_K(\phi) d\phi = g_0$ so that we choose $g_0 = 1$. It is peaked at $\phi = 2\pi n$ where n is integer. The width of the peak is proportional to M^{-1}. The kernel polynomial approximation to $F(x)$ is then

$$\widetilde{F}_K(\phi) = \int_0^{2\pi} \delta_K(\phi - \phi_o) \widetilde{F}(\phi_o) d\phi_o \quad . \tag{12}$$

The choice of Gibbs damping factor, g_m, determines the quality of kernel polynomial approximations. For applications to DOS and spectra we propose four reasonable criteria, which result in the same Gibbs damping factor originally derived in the mathematical theory of uniform approximation by polynomials (Jackson 1930; Meinardus 1967; Rivlin 1969). The first criterion is that the kernel should be a polynomial of degree M, if only M Chebyshev moments of the DOS are available. The second criterion is that the kernel estimates of DOS and spectra should be strictly positive as required by physics. These two criteria can be met uniquely by the representation

$$\delta_K(\phi) = \frac{1}{2\pi} \left| \sum_{\nu=0}^{M} a_\nu e^{i\nu\phi} \right|^2 \quad , \tag{13}$$

where the a_ν are real. Upon comparison with Eq. (11)

$$g_m = \sum_{\nu=0}^{M-m} a_\nu a_{\nu+m} \quad . \tag{14}$$

The third criterion is that the kernel should be normalized so that the total number of states is preserved, which is met by constraining $g_0 = 1$. The fourth criterion is that the energy resolution should be the best achievable for M Chebyshev moments subject to these other constraints. This corresponds to minimizing the variance

$$\Delta\phi^2 \equiv \int_{-\pi}^{\pi} \phi^2 \delta_K(\phi) d\phi \simeq 2g_0 - 2g_1 \quad . \tag{15}$$

Combining this with Eq.(14) is equivalent to maximizing

$$Q = g_1 - \lambda g_0 = \sum_{\nu=0}^{M-1} a_\nu a_{\nu+1} - \lambda \sum_{\nu=0}^{M} a_\nu a_\nu \quad , \tag{16}$$

where λ is a Lagrange multiplier to enforce the third criterion on normalization, $g_0 = 1$. The variational condition $\delta Q/\delta a_\nu = 0$ yields

$$\begin{aligned} a_1 - 2\lambda a_0 &= 0 \quad , \\ 0 \le \nu \le M-2 \quad a_{\nu+2} - 2\lambda a_{\nu+1} + a_\nu &= 0 \quad , \\ -2\lambda a_M + a_{M-1} &= 0 \quad . \end{aligned} \tag{17}$$

But these are just the recursion relations for Chebyshev polynomials of the second kind. Hence,

$$a_\nu = \frac{U_\nu(\lambda)}{\sqrt{\sum_{\nu=0}^{M} U_\nu^2(\lambda)}} \quad ; \quad U_\nu(\lambda) = \frac{\sin((\nu+1)\phi_\lambda)}{\sin(\phi_\lambda)} \quad . \tag{18}$$

with $\phi_\lambda \equiv \arccos(\lambda)$. The last line in Eq. (17) corresponds to $U_{M+1}(\lambda) = 0$, which is equivalent to $\phi_\lambda = \frac{\pi n}{M+2}$. The maximum of Q, Eq. (16), is obtained for $n = 1$. This choice satisfies our criteria for an optimal kernel.

Figure 1 exhibits kernel polynomials for no Gibbs damping (labeled *None*) and for the optimal Gibbs damping factor (labeled *Optimal*). Without Gibbs damping, the kernel is not positive and has oscillations extending to large $|\phi|$. With optimal Gibbs damping, the kernel is strictly positive, its width $\Delta\phi \propto M^{-1}$, and Gibbs oscillations at large $|\phi|$ are rapidly damped (see the inset with a magnified vertical scale). This rapid damping minimizes leakage of information from one energy to another. Small leakage is especially important for many-body physics applications where the scientific interest is often focused on a few low energy states near to orders of magnitude more higher energy states; see (Silver and Röder 1994) for an example using Heisenberg antiferromagnets.

3 Application to Tight-Binding Electronic Structure

The total-energy, tight-binding method is gaining popularity for atomistic simulations in materials science (Sankey and Allen 1986; Wang, Chan and Ho 1989). This semiempirical electronic structure approach is based on a one-electron Hamiltonian with on-site energies and interatomic-distance-dependent hopping integrals. The atom-centered basis set is chosen to be appropriate for the valence orbitals of the system under study. The band (electronic) energy E_B, defined as the sum of the energies of all occupied states up to the Fermi energy, is augmented by a short-ranged pair potential. The computational bottleneck in tight-binding calculations is the diagonalization of $\mathbf{H}$, requiring $O(N^3)$ work. There has been much interest recently in approximations that improve on this scaling without undue loss of accuracy. In this section we test an $O(N^2)$ implementation of KPM for this problem.

Silicon is chosen as a working example, and we employ the tight-binding parameters of (Goodwin, Skinner and Pettifor 1989). The basis set consists of an s function and the 3 Cartesian p functions on each atom. For the present calculations, only the electronic energy is considered. To compute the DOS and cohesive energy of a system approximating bulk Si, a Hamiltonian matrix is constructed for a block of 216 Si atoms arranged in the diamond structure. This block is placed in a cubic box (supercell) with periodic boundary conditions. [Only the $\vec{k} = \vec{0}$ (Γ) point in the Brillouin zone is needed, since the 216 atom supercell system is equivalent to a nearly converged sample of $\vec{k}$-points for the primitive diamond-structure cell.] We also consider a 215-atom system, in which one atom has been removed while the remaining atoms are held fixed. This allows computation of the unrelaxed vacancy formation energy (E_V), defined as $E_V \equiv E_B^{215} - \frac{215}{216} E_B^{216}$, which has been shown (Kress and Voter 1995) to provide a much more stringent test of moment-based approximations than the cohesive energy.

Let $D(\varepsilon)$ be the DOS, $C(\varepsilon) = \int_{-\infty}^{\varepsilon} D(\varepsilon')d\varepsilon'$ be the cumulative DOS, and $E(\varepsilon) = \int_{-\infty}^{\varepsilon} \varepsilon' D(\varepsilon')d\varepsilon'$ be the cumulative energy. For Si, the Fermi energy ε_F is defined by the condition that the number of occupied states corresponds to two electrons per atom, $C(\varepsilon_F) = 2N_{atom}$. The band energy is then defined as $E_B \equiv E(\varepsilon_F)$. The Hamiltonian matrices considered here are sufficiently small that we may compare KPM results to those of exact diagonalization by $O(N^3)$ methods.

The cumulative DOS and cumulative energy may

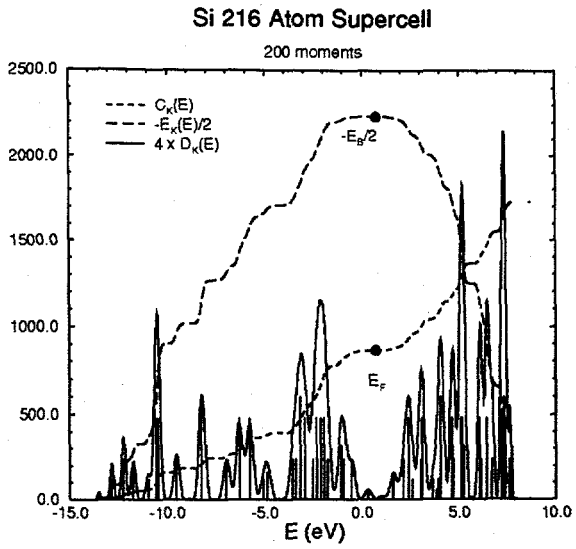

Figure 2: **KPM for Si Supercell** - The vertical lines indicate eigenenergies of a 216 Atom Si supercell calculated by exact $O(N^3)$ diagonalization, with the height proportional to the degeneracy ($\times$ 20). Also displayed are results of the kernel polynomial method for the density of states (solid), cumulative DOS (dashed), and cumulative energy (long dash). The Fermi energy E_F is the energy at which the cumulative DOS equals the number of electrons. For this example, E_F is positioned in a large gap in the DOS. The band energy E_B is the cumulative energy at the Fermi energy.

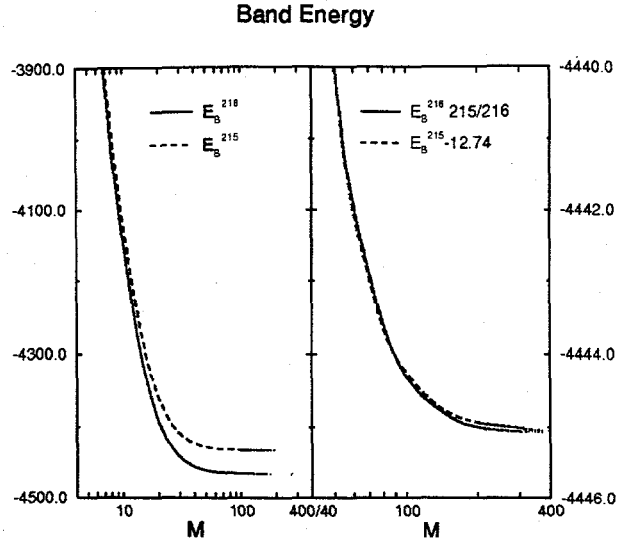

Figure 3: **Band Energies** - The left side shows band energies E_B as a functions of the number of Chebyshev moments M (on a logarithmic scale) for the 216 and 215 atom Si cells. The right side uses a two orders of magnitude expanded vertical scale to display the asymptotic convergence of band energies at large M. The band energies are scaled in the way they enter the vacancy energy, $E_V \equiv E_B^{215} - \frac{215}{216}E_B^{216}$, and E_B^{215} is shifted by the vacancy formation energy obtained from exact diagonalization, 12.74 eV.

be calculated from

$$\widetilde{C}(\phi) = \int_{\phi}^{\pi} \widetilde{D}(\phi')d\phi' \quad ,$$

$$\widetilde{E}(\phi) = \int_{\phi}^{\pi} \cos(\phi')\widetilde{D}(\phi')d\phi' \quad , \qquad (19)$$

using $\widetilde{C}(\phi) = C(\varepsilon)$ and $E(\varepsilon) = a\widetilde{E}(\phi) + b$. Define $c_m(\phi) \equiv \frac{\cos(m\phi)}{\pi}$ and $s_m(\phi) \equiv \frac{\sin(m\phi)}{\pi m}$. Then expansions for these quantities in exact Chebyshev moments are

$$\widetilde{D}(\phi) = \frac{\mu_0}{\pi} + 2\sum_{m=1}^{\infty} \mu_m c_m(\phi) \quad ,$$

$$\widetilde{C}(\phi) = \mu_0\left(1 - \frac{\phi}{\pi}\right) - 2\sum_{m=1}^{\infty} \mu_m s_m(\phi) \quad ,$$

$$\widetilde{E}(\phi) = \mu_1\left(1 - \frac{\phi}{\pi}\right) - \sum_{m=1}^{\infty} (\mu_{m-1} + \mu_{m+1})\, s_m(\phi) \quad . \qquad (20)$$

$\widetilde{E}(\phi)$ may be continued to $0 \le \phi \le 2\pi$ using $\widetilde{E}(2\pi - \phi) = -\widetilde{E}(\phi)$.

Practical calculations will yield only a finite number of moments. We form the kernel polynomial approximations to Eqs. (20) using Eq. (12), yielding

$$\widetilde{C}(\phi) = 2\sum_{m=1}^{M} (\mu_o - \mu_m)\, g_m s_m(\phi) \quad ,$$

$$\widetilde{E}(\phi) = \sum_{m=1}^{M+1} g_m s_m(\phi)\,(2\mu_1 - \mu_{m-1} - \mu_{m+1}) \quad . \qquad (21)$$

Figure 2 shows the DOS, cumulative DOS and cumulative energy for the 216 atom supercell obtained from 200 Chebyshev moments using KPM. For comparison, the DOS obtained by $O(N^3)$ exact diagonalization are displayed as vertical lines at the eigenenergies with a height equal to 20 times their degeneracy. The KPM DOS equals the true DOS convolved with the kernel polynomial. The Fermi energy ε_F is the point where the cumulative DOS equals the number of electrons. The band energy E_B is then the cumulative energy evaluated at the Fermi energy, $E_B = E(\varepsilon_F)$.

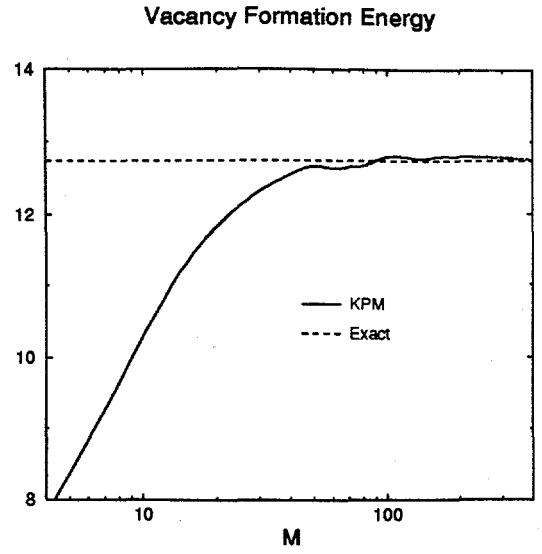

Figure 4: **Vacancy Formation Energy** - Solid curve is the vacancy formation energy E_V vs. the number of Chebyshev moments M on a logarithmic scale. The dashed curve indicates the exact diagonalization result. E_V has converged to within 0.1 eV at about 40 moments.

The left side of Figure 3 shows the band energies for the 216 atom and 215 atom supercells as functions of the number of moments M. The right side shows the asymptotic convergence of these quantities at large M. Plotted are the quantities which enter the calculation of the vacancy energy, i.e. $E_B^{216}\frac{215}{216}$ and $E_B^{215} - 12.74$ where 12.74 eV is the vacancy energy obtained by exact $O(N^3)$ diagonalization. Figure 4 shows the vacancy formation energy, E_V vs. M. The dashed line is the exact result obtained by diagonalization of the Hamiltonian, an $O(N^3)$ process. It is remarkable that E_V has converged on the exact E_V to within 0.1 eV for $M \geq 40$, even though E_B does not asymptote to the exact E_B until $M \geq 150$. Since E_V is formed from the difference of large numbers, the relative accuracy is $\frac{0.1}{4465} \approx 0.0024\%$. Note also that the Chebyshev recurrence procedure requires only $M/2$ MVMs to yield M moments. So E_V has converged after only 20 MVMs.

Details of these calculations may be found in (Silver et al. 1995). For example, Eqs. (21) differ subtly but significantly from cumulative distributions obtained by integrating over the kernel polynomial approximation to the DOS, a difference which removes systematic bias and significantly improves the rate of convergence with M (Silver *et al.* 1995). This change in the order of approximation can dramatically affect the convergence rate of E_V with increasing M, because E_V is formed as a difference of large numbers. The same paper compares the task performance of the optimal Gibbs damping to that of many other Gibbs damping factors proposed in the applied math and physics literature. The other Gibbs damping factors do not enforce positivity of DOS and montonicity of cumulative DOS, so they yield multiple solutions for the Fermi energy which slows convergence for this application.

4 Conclusion

Taking into account the variety of KPM applications discussed in (Silver et al. 1995), we present some comparisons of KPM with competing Lanczos recursion methods (LRM) and maximum entropy methods (MEM) for estimating DOS and spectra:

- Uniformity: KPM are 'uniform' in two respects. They produce estimates for DOS and spectra which are simple linear convolutions of a kernel figure function with the true DOS or true spectrum, with resolution uniform in ϕ and inversely proportional to the number of moments. They also produce estimates corresponding to "uniform approximations by polynomials" for which there exists a well-developed mathematical theory (Jackson 1930; Meinardus 1967; Rivlin 1969) including error bounds. LRM should be preferred for determining exact eigenvalues of a few well-separated or band edge states, but LRM converge slowly for complete DOS or spectra. MEM yield better resolution than KPM, because MEM fit all the moment data. Moments of KPM estimates are related to exact moments by Gibbs damping factors. MEM may be preferred when there exists significant prior knowledge about the DOS or spectrum such as default models and other kinds of data. But the resolution of MEM is not uniform because MEM are non-linear.

- Positivity: An important argument commonly advanced in favor of MEM is they enforce positivity of DOS and spectral estimates, and monotonicity of cumulative DOS. With the introduction of the optimized kernel in Section 2, KPM also enforce positivity and monotonicity. Monotonicity is important in our electronic structure example to eliminate instability caused by the possible multiple solutions for the Fermi energy.

- Numerical Stability: The Chebyshev recurrence relations used in the KPM are numerically stable to extremely large number of moments, with negligible susceptibility to numerical roundoff error. They can be used to generate data for both KPM and MEM. KPM estimates are a simple linear function of the moments. MEM estimates are a non-linear function of the moments, and finding them requires a convex optimization algorithm which can be unstable. LRM algorithms are notoriously unstable in the presence of numerical roundoff error, with the Lanczos phenomenon causing a loss of orthogonality. LRM can be corrected only be reorthogonalization or by a complex labeling scheme producing error bars.

- Algorithmic simplicity: KPM algorithms typically need only a small fraction of number of lines of code of MEM or LRM algorithms.

- Scaling: For the electronic structure application described in the previous section, the cpu required scales like $O(N^2M)$, where N is the dimension of the sparse Hamiltonian and M is the number of moments. The memory required scales like N for sparse Hamiltonians. However, Chebyshev recursion algorithms with cpu scaling as $O(NM)$ can be generated in a variety of ways, which produce moments usable by both KPM and MEM. For example, (Silver and Röder 1994) introduce a stochastic KPM to produce estimates of DOS and thermodynamic functions subject to statistical error. Other linear scaling KPM algorithms avoid statistical errors for physical systems satisfying extensivity, i.e. where the range of off-diagonal elements of density matrices is limited.

- Flexibility: KPM can be simply interpreted as polynomial approximations to Dirac delta functions, wherever they occur in physical problems. For example, kernel polynomial approximations to Fermi projection operators (defined as operators which project unoccupied states in a random vector to zero) can be formed by integrating over kernel polynomials in the same sense that step functions are integrals over delta functions. As another example, replacing delta functions by kernels in statistical mechanics calculations leads to rapidly converging modified Bessel - Chebyshev moment expansions for thermodynamic functions.

We conclude that KPM should become a useful complement to LRM and MEM for computations involving large sparse Hamiltonians.

Acknowledgements

Supported in part by the Office of Basic Energy Sciences of the U. S. Department of Energy.

References

Goodwin, L.; A. J. Skinner; D. G. Pettifor 1989. "Generating Transferable Tight-Binding Parameters: Application to Silicon", Europhysics Letters **9**, 701-706.

Jackson, D. 1930, *Theory of Approximation*, American Mathematical Society Colloquium Publications, Vol XI, Providence, R. I.

Kress, J. D.; A. F. Voter 1995. "Interatomic Potentials from Tight Binding Moment Expansions", submitted to Physical Review B.

Meinardus, G. 1967. *Approximation of Functions: Theory and Numerical Methods*, Springer-Verlag, New York.

Rivlin, T. J. 1969. *An Introduction to the Approximation of Functions*, Blaisdell, Waltham.

Silver, R. N.; H. Röder 1994. "Densities of States of Mega-Dimensional Hamiltonian Matrices", International Journal of Modern Physics C**5**, 735-753.

Silver, R. N.; H. Röder; A. F. Voter; J. D. Kress 1995. "Kernel Polynomial Approximations for Densities of State and Spectra," submitted to Physical Review B.

Sankey, O. F.; R. E. Allen 1986. "Atomic Forces from Electronic Energies via the Hellmann-Feynman Theorem," Physical Review B**33**, 7164-7171.

Wang, C. Z.; C. T. Chan; K. M. Ho 1989. "Empirical Tight-Binding Force Model for Molecular Dynamics Simulations of Si," Physical Review B**39**, 8586.