LA-UR-95- 3279

CONF-9507156--2

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: Chebyshev Recursion Methods: Kernel Polynomials and Maximum Entropy

AUTHOR(S): R. N. Silver
H. Roeder
A. F. Voter
J. D. Kress

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

SUBMITTED TO: Proceedings of the Hayashibara Forum '95, Kyoto, Japan
July, 1995, World Scientific Publishers

By acceptance of this article, the publisher recognized that the U S Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution or to allow others to do so for U S Gove National Laboratory requests that the publisher identify this article as work performe the auspices of the U S Department of Energy.

MASTER

Los Alamos National Laboratory
Los Alamos, New Mexico 87545

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

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

CHEBYSHEV RECURSION METHODS: KERNEL POLYNOMIALS AND MAXIMUM ENTROPY

R. N. SILVER, H. ROEDER, A. F. VOTER, J. D. KRESS
Theoretical Division, Los Alamos National Laboratory
Los Alamos, New Mexico 87545, USA
E-mail: rns@loke.lanl.gov

ABSTRACT

We describe two Chebyshev recursion methods for calculations with very large sparse Hamiltonians, the kernel polynomial method (KPM) and the maximum entropy method (MEM). They are especially applicable to physical properties involving large numbers of eigenstates, which include densities of states, spectral functions, thermodynamics, total energies, as well as forces for molecular dynamics and Monte Carlo simulations. We apply Chebyshev methods to the electronic structure of Si, the thermodynamics of Heisenberg antiferromagnets, and a polaron problem.

1. Introduction

In computational condensed matter physics we are often interested in calculating physical properties of sparse model Hamiltonians for finite systems. The number of states N is usually much too large to apply conventional eigenvalue methods scaling as $O(N^3)$. Efficient calculations of ground and isolated state properties usually employ Lanczos recursion methods, which scale as $O(N)$ and use only matrix-vector-multiplies (MVM) to minimize storage. Unfortunately, Lanczos methods are inefficient and statistically uncontrolled for properties involving large numbers of eigenstates. These include densities of states (DOS), spectral functions, thermodynamics, total energies, forces for molecular dynamics and Monte Carlo simulations, etc. Lanczos methods are also numerically unstable for large numbers of recursions without expensive reorthogonalizations. The present paper suggests that Chebyshev recursion methods[1,2] can overcome such difficulties. They scale as $O(N)$ for properties involving large numbers of states if finite energy resolution and statistical accuracy are acceptable. They are numerically stable for large numbers of recursions.

Consider the density of states (DOS) as representative of properties of interest. The first step in applying Chebyshev methods is to scale the Hamiltonian, $\mathbf{H} = a\mathbf{X}+b$ such that all eigenvalues X_n of $\mathbf{X}$ lie between -1 and $+1$. The DOS is then

$$D(X) = \frac{1}{N}\sum_{n=1}^{N}\delta(X - X_n) \quad . \tag{1}$$

The data about $D(X)$ consists of Chebyshev moments,

$$\mu_m = Tr\{T_m(\mathbf{X})\} = \int_{-1}^{1} T_m(X)D(X)dX \quad . \tag{2}$$

These are more informative than power moments, $Tr\{X^m\}$, at finite machine precision. Calculations use Chebyshev recursion,

$$T_{m+1}(\mathbf{X}) = 2\mathbf{X}T_m(\mathbf{X}) - T_{m-1}(\mathbf{X}) \quad , \tag{3}$$

with the same MVM algorithm used in Lanczos methods. Exact evaluation of M moments uses cpu time $\propto O(N^2M)$. A stochastic method[1], scaling as $O(NMN_r)$, uses estimators

$$\mu_m \approx \frac{1}{N_r}\sum_r < r|T_m(\mathbf{X})|r > \quad , \tag{4}$$

where $|r>$ are N_r Gaussian random vectors. Such data have calculable statistical variance proportional to $1/NN_r$. If the Hamiltonian has only local off-diagonal elements, as in tight-binding Hamiltonians, a non-stochastic locally truncated approximation to the Hamiltonian $\mathbf{H}_i$ may be adequate[4]. The estimator,

$$\mu_m \approx \sum_i < i|T_m(\mathbf{X}_i)|i > \quad , \tag{5}$$

generates data with a systematic error determined by the truncation range. Cpu scales as $O(NMJ)$, where J is the number of states in the truncation range. Exact energy derivatives (or forces) can also be calculated.

2. Methods

The goal is to make the best possible estimate of the DOS using the least cpu time and memory. The number of moments M will be be limited and subject to statistical and systematic errors.

KPM starts with an exact expansion of the DOS

$$D(X) = \frac{1}{\pi\sqrt{1-x^2}}\left[\mu_0 + 2\sum_{n=1}^{\infty}\mu_m T_m(X)\right] \quad , \tag{6}$$

which is then truncated at M moments. A factor g_m is also introduced to damp the Gibbs phenomenon,

$$D_K(X) = \frac{1}{\pi\sqrt{1-x^2}}\left[\mu_0 + 2\sum_{n=1}^{M}\mu_m g_m T_m(X)\right] \tag{7}$$

The label "kernel" becomes meaningful after rewriting in the variable $\phi = \cos^{-1}(x)$ in which $T_m(x) = \cos(m\phi)$. Then

$$D_K(\phi) = \int_0^{2\pi}\delta_K(\phi-\phi_o)D(\phi_o)d\phi_o \quad ; \quad \delta_K(\phi) = \frac{1}{2\pi}\left[g_0 + 2\sum_{m=1}^{M} g_m\cos(m\phi)\right] \quad , \tag{8}$$

is a simple convolution. $D_K(\phi)$ is a truncated Fourier series. The "kernel" $\delta_K(\phi)$ is a 2π-periodic polynomial approximation to a Dirac delta function, analogous to the

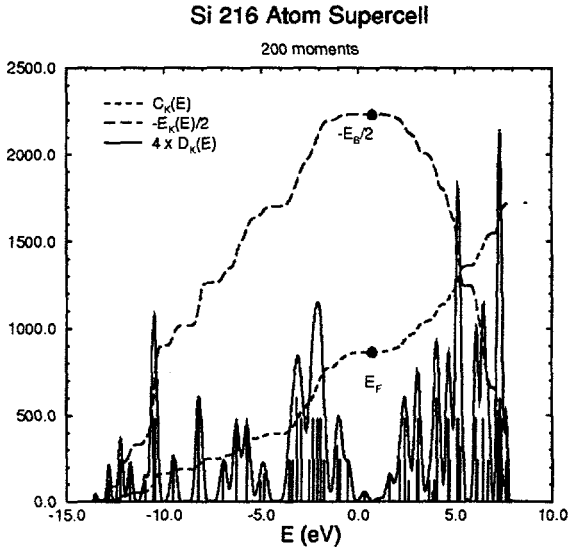

Figure 1: KPM DOS and band energy calculation for Si 256 atom supercell.

resolution function of a spectrometer. Resolution is uniform in ϕ with width $\Delta\phi \propto M^{-1}$. If $g_m = 1$, the kernel is oscillatory with a slowly decreasing envelope function at large $|\phi|$. The result is the Gibbs phenomenon, a lack of uniform convergence at discontinuities in DOS. An optimal g_m is determined variationally by minimizing the uniform norm[6]. The resulting kernel is strictly positive, normalized and has minimal variance in ϕ[3]. This choice guarantees positivity of the DOS and monotonicity of cumulative DOS required for electronic structure.

KPM can be applied to other properties such as spectral functions[3],

$$A(\omega) = \lim_{\eta \to 0+} \frac{1}{\pi} Im \left\{ < \Psi_0 | \mathbf{O}^\dagger \frac{1}{\omega - \mathbf{H} - i\eta} \mathbf{O} | \Psi_0 > \right\} \quad . \tag{9}$$

KPM approximations use moments $\mu_m^{\mathbf{O}} = < \Psi_0 | \mathbf{O}^\dagger T_m(\mathbf{X}) \mathbf{O} | \Psi_0 >$.

Applications to thermodynamics use a rapidly converging Fourier-Bessel expansion of the partition function[1],

$$Z = e^{-\beta b} \left[I_0(\beta a) + 2 \sum_{m=1}^{\infty} I_m(\beta a) \mu_m \right] \quad . \tag{10}$$

The $I_m(\beta a)$ are modified Bessel functions. The partition function involves integral rather than pointwise properties of the DOS, so the best convergence is achieved with $g_m = 1$.

Kernel polynomial approximations for finite temperature spectral functions can be calculated from double moments of the form $\mu_{mn} = Tr\{T_m(\mathbf{X}) \mathbf{O}^\dagger T_n(\mathbf{X}) \mathbf{O}\}$.

The maximum entropy method (MEM)[7,8] can use the same Chebyshev moment data as KPM. MEM achieves at least a factor of 4 better energy resolution than KPM.

MEM enforces prior knowledge such as positivity. It can take advantage of default models and other known constraints to improve convergence. It is readily extended to uncertain data. But MEM introduces a non-uniform resolution, added algorithmic and computational complexity, and some tendency toward artifacts. For exact data, the DOS is estimated by maximizing its entropy relative to a default model $M(X)$,

$$S(D) = -\int_{-1}^{1} \left[D(X) \ln \left(\frac{D(X)}{M(X)} \right) \right] dX \quad , \tag{11}$$

within the constraints of the data. That is, maximize

$$Q \equiv S(D) - \sum_{m=1}^{M} \lambda_m \int_{-1}^{+1} T_m(X) D(X) dX \tag{12}$$

where the λ_m are Lagrange multipliers. The solution is

$$D(X) = \frac{M(X)}{Z} \exp \left(- \sum_{m=1}^{M} \lambda_m T_m(X) \right) \tag{13}$$

Finding a set of $\{\lambda_m\}$ is a dual-space non-linear convex optimization problem, which can be solved using standard algorithms. Chebyshev moment data are advantageous over power moments for MEM because they permit FFT methods to evaluate integrals. Required cpu time scales as $O(M^2)$ and is negligible compared to data generation time. Efficient MEM algorithms are discussed elsewhere[9].

3. Applications

Chebyshev recursion methods have now been applied to a wide variety of condensed matter physics problems including the electronic structure and relaxation of Si and its defects[3,4], the dielectric functions of quantum dots[5], the many-body densities of states of the Holstein $t-J$ model [10], the thermodynamics of the Heisenberg model on various lattices[11], spectral functions of the disordered XXZ model[3], etc.

Figure 1 illustrates the application of KPM to the electronic structure of a 216 atom Si supercell using a tight binding Hamiltonian[3]. This system is small enough to be exactly diagonalized. Vertical lines are at the energies of the exact eigenstates and their height is proportional to their degeneracy. The solid line is the KPM approximation to the DOS obtained for 200 Chebyshev moments. A Fermi energy E_F is the energy at which the cumulative DOS $C_K(E)$ equals the number of electrons. The total band energy E_B is then the cumulative energy $E_K(E)$ at E_F. For band energies KPM converges in proportion to M^{-2} reaching 10^{-5} accuracy at about $M \approx 150$. MEM converges about a factor of 4 faster, reaching 10^{-5} accuracy at $M \approx 35$.

Figure 2 applies KPM to the DOS of a 26 site Heisenberg Hamiltonian with $N = 67,108,864$. The inset blows up the low energy region. Accurate thermodynamic

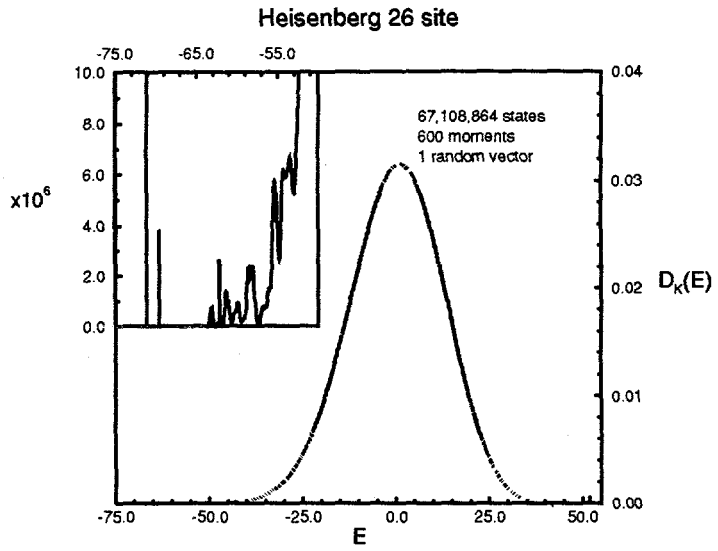

Figure 2: DOS of 26 site Heisenberg model on square lattice from stochastic KPM.

results require fewer random vectors as N increases because the variance scales as $1/NN_r$. In the present example, $N_r = 1$ is sufficient to achieve an accuracy less than 5% for the entropy down to temperatures $T = 0.5$. We have calculated the Heisenberg model on various Kagome lattices, revealing a surprising size dependence of thermodynamic quantities such as the specific heat and static susceptibility[11].

Figure 3 compares MEM and KPM for the DOS of a 1D polaron formation problem. The Hamiltonian consists of an electron placed into a $10,000$ atom chain with a Peierls distortion, which is then allowed to relax resulting in the polaron state at $E = 1.0$. MEM achieves dramatically better energy resolution than KPM for isolated states and band edges, but it tends to "ring" (or oscillate) when singular structures, such as Van Hove singularities, are nearby.

4. Conclusions

Both KPM and MEM are efficient N-scaling methods for computational many-body physics and electronic structure problems involving large numbers of eigenstates. They are based on well-developed concepts in analysis and statistics such as Chebyshev approximations, Fourier analysis, unbiased estimators, random sampling and non-linear optimization. They use the same MVM algorithm as Lanczos diagonalization minimizing storage requirements. KPM is a controlled approximation with known error bounds. MEM achieves significantly better resolution at the expense of computational complexity. Both are applicable to extremely large Hamiltonians, and complementary to Lanczos methods.

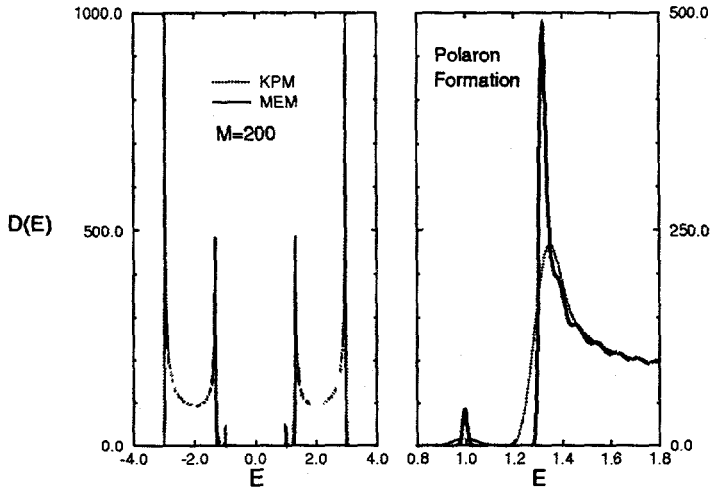

Figure 3: Comparison of KPM and MEM for the DOS of a polaron formation problem.

5. Acknowledgements

Work supported by the U. S. Dept. of Energy.

6. References

1. R. N. Silver, H. Roeder, Int. J. Mod. Phys. C**5**, 735-753 (1994)
2. L. W. Wang, Physical Review B**49**, 10154-10158 (1994).
3. R. N. Silver, H. Roeder, A. F. Voter, J. D. Kress, to appear in J. Comp. Phys. (1995).
4. A. F. Voter, J. D. Kress, R. N. Silver, at American Ceramics Meeting (1995) and submitted to Phys. Rev. B. (1995).
5. L. W. Wang, A. Zunger, Phys. Rev. Letts. **73**, 1039-1042 (1994).
6. D. Jackson, *Theory of Approximation*, Amer. Math. Soc. Colloq. **XI** (1930).
7. L. R. Mead, N. Papanicolaou, J. Math. Phys. **25**, 2404 (1984).
8. D. A. Drabold, O. F. Sankey, Phys. Rev. Lett. **70**, 3631 (1993).
9. R. N. Silver, H. Roeder, submitted to J. Comp. Phys. (1995).
10. H. Roeder, H. Feschke, R. N. Silver, Europhysics Letters **28**, 257-262 (1994).
11. H. Roeder, R. N. Silver, to appear in *Recent Progress in Many Body Theories* **4**, eds. E. Schachinger, et al. (1994).