Projected random vectors and the recursion method in the electronic-structure problem

Otto F. Sankey
Department of Physics and Astronomy, Arizona State University, Tempe, Arizona 85287

David A. Drabold
Department of Physics and Astronomy, Condensed Matter and Surface Sciences Program, Ohio University, Athens, Ohio 45701

Andrew Gibson
Molecular Science Research Center, Pacific Northwest Laboratory, Richland, Washington 99352
(Received 8 February 1994)

We develop a technique to determine the occupied eigenstates in the matrix formulation of the electronic-structure problem. The theory uses a random vector projected onto the electron occupied subspace by use of a Fermi-Dirac projection operator. This random starting vector is inserted into the recursion scheme to generate all occupied eigenenergies and eigenvectors of the system. The method produces a tridiagonal Hamiltonian matrix, which unlike the original Hamiltonian matrix, can be diagonalized even for a very large system. Hellmann-Feynman forces are readily obtained because the eigenvectors can be efficiently computed. Care must be taken to correct for instabilities in the three-term recurrence which gives rise to spurious solutions.

I. INTRODUCTION

In an electronic-structure calculation of the relaxed geometry of a large system the band structure energy and the Hellmann-Feynman forces are the most difficult quantities which must be evaluated. This is because they depend upon the eigenvalue spectrum, which is usually calculated by diagonalization of the Hamiltonian matrix. In this paper, we describe a method for calculating the occupied electronic energy eigenvalues and eigenvectors for which the number of floating point operations is proportional to $(P \times M) + (N \times M \times M)$, where P is the number of operations to compute a single moment of the Hamiltonian, N is the dimension of the Hamiltonian and M is proportional to the number of occupied levels. The utility of this method over standard matrix diagonalization routines accrues when the number of the occupied levels is much smaller than the dimension of the Hamiltonian, N, and if P is linearly proportional to N.

Random vectors were introduced in a recent paper by Drabold and Sankey (DS) (Ref. 1) who proposed a method for the calculation of the electronic structure of an N-atom system with linear scaling with system size (a so-called order N method). That method was based upon the use of random vectors which mimic the effect of the impartial vector $|\xi\rangle$,

$$|\xi\rangle = \frac{1}{\sqrt{N}} \sum_i |\psi_i\rangle,$$

where the ψ_i are eigenstates of the single particle Hamiltonian $\hat{H}$, $\hat{H}\psi_i = \epsilon_i \psi_i$. DS used energy moments to construct the electronic density of states (DOS) using a maximum entropy scheme. From the DOS, the band structure energy can be obtained, $E_{\rm BS} = \int_{-\infty}^{E_F} ED(E)dE$.

A method similar to that of DS was used by Skilling[2] and Silver *et al.*[3] Silver *et al.* studied a two-dimensional (2D) 4×4 Heisenberg model. We have recently found that related ideas have been investigated in nuclear physics, primarily for computing level densities.[4] The use of random vectors in a recursion scheme was first used by Lanczos himself in the recursion method that now bears his name.[5] A comparison of Lanczos recursion and moments using random vectors has been made.[6] In addition, random vectors in a recursion scheme have been used to calculate the density of states of phonons in amorphous solids.[7]

In this work, we describe further applications of random vectors to the electronic structure problem. Here we use the recursion method to compute its spectral properties. A random vector is projected (using a Fermi-Dirac projection operator) onto the occupied subspace of eigenvectors and is used as a starting vector in the recursion method,[8] to generate a tridiagonal matrix. Because of instabilities of the recursion method, spurious states (described below) must be taken into consideration. As an example, we apply this theory to an empirical tight-binding Hamiltonian. With alterations, the method can be applied to plane wave basis calculations as well.[9] We will report this plane wave work elsewhere.[10]

II. METHODOLOGY

To illustrate these ideas we consider a tight-binding model with N states, specifically the Vogl *et al.*[11] sp^3s^*

nearest neighbor tight-binding Hamiltonian model of GaAs. Here, N is $5 \times n_A$, where n_A is the number of atoms. For demonstration purposes, we choose a 64 atom cubic supercell of GaAs in which we displace each atom in each of three directions by a random amount between -0.1 and $+0.1$ Å. This distortion breaks the symmetry of the cell. We use the Γ point ($\vec{k} = \vec{0}$) as the intercell wave vector. The hopping parameters are scaled with nearest neighbor distance as d^{-2}. The size of the Hamiltonian matrix is 320×320, which is readily exactly diagonalized, so that we can test the approximations and convergence of the ideas presented here.

The process is started by generating a normalized random vector $|x\rangle$ in the space of vectors of dimension N by use of a random number generator. This random vector will have components x_i on all eigenstates ψ_i (occupied and unoccupied),

$$|x\rangle = \sum_i x_i |\psi_i\rangle, \tag{1}$$

of the $N \times N$ Hamiltonian matrix. This vector is next projected onto the subspace of occupied states by applying the Fermi-Dirac projection operator to it, $\hat{f}(\hat{H})$, where

$$\hat{f}(\hat{H}) = (e^{\beta(\hat{H}-E_F)} + 1)^{-1}. \tag{2}$$

The utility of this operator has been noted in other contexts by Daw.[12] To apply $\hat{f}$ to $|x\rangle$, we first scale and shift the Hamiltonian $\hat{H}$ so that its spectrum lies in the range $[-1,1]$. Then $\hat{f}$ is expanded in a Tchebychev series of K terms in $\hat{H}$,

$$\hat{f}(\hat{H}) = \sum_{n=0}^{K} d_n(E_F) T_n(\hat{H}). \tag{3}$$

In this paper, we will choose the temperature parameter β large enough so that the action of the Fermi-Dirac operator on an eigenstate is to effectively multiply it by zero or 1. This requires a band gap between occupied and unoccupied states for this to be strictly true, and this is the only case we considered here. However, if we did wish to investigate some conduction states, we simply choose E_F to be in the conduction band, and the parameter β can be made larger. This approach should therefore be applicable to metallic systems which are a persistent challenge to *ab initio* methods.

The key utility of the projection methods is that it enables us to work entirely within the *occupied subspace* which has dimension $n_{\rm occ}$, the number of occupied electronic states, rather than N which can be much larger than $n_{\rm occ}$. Of course one can view the Car-Parrinello[13] and conjugate gradient[14] as variants on this idea; iteratively "filtering" out the lowest energy (*occupied*) eigenvectors. The present operator technique accomplished this *directly*—i.e., noniteratively, in one application of the Tchebychev approximant Fermi-Dirac operator.

The number of terms K in the expansion of Eq. (3) needed to obtain an accurate representation is between 100 and 200, and is independent of the system size. Notice that the expansion coefficients depend on E_F and implicitly on β and can be computed quite simply.[15] The projected random vector $|z\rangle$ obtained from Eq. (3) is

$$|z\rangle = \hat{f}(\hat{H})|x\rangle = \sum_n d_n(E_F)(T_n(\hat{H})|x\rangle). \tag{4}$$

This requires the application of $T_n(H)$ on $|x\rangle$ which is evaluated through a recurrence relation of Tchebychev polynomials,

$$T_n(\hat{H})|x\rangle = 2\hat{H}T_{n-1}(\hat{H})|x\rangle - T_{n-2}(\hat{H})|x\rangle, \tag{5}$$

with $T_0(\hat{H})|x\rangle = |x\rangle$ and $T_1(\hat{H})|x\rangle = \hat{H}|x\rangle$. The projected random vector z can then be written in terms of *occupied* states as

$$|z\rangle = \sum_{i \text{ occupied}} z_i |\psi_i\rangle. \tag{6}$$

We show in Fig. 1 the Fermi-Dirac function obtained from Eq. (3) using 150 terms in the Tchebychev expansion for our 64 atom GaAs supercell. The highest occupied valence state is at $+0.05$, and the lowest occupied conduction state is at $+0.14$. These two energies are indicated by short vertical lines in the figure. The Fermi level E_F was chosen half way between these, and the β parameter is 100. We see that this expansion accurately represents a Fermi-Dirac function, except for small oscillations near the conduction band edge. In the calculations reported in this paper these small wiggles cause no difficulty. However, they can be reduced if necessary simply by increasing the number of terms in the Tchebychev expansion.

We now attempt to exploit the fact that the vector $|z\rangle$ contains *only* occupied states. Therefore, any spectrum of energy eigenstates extracted from $|z\rangle$ must contain only occupied eigenvalues and no unoccupied eigen-

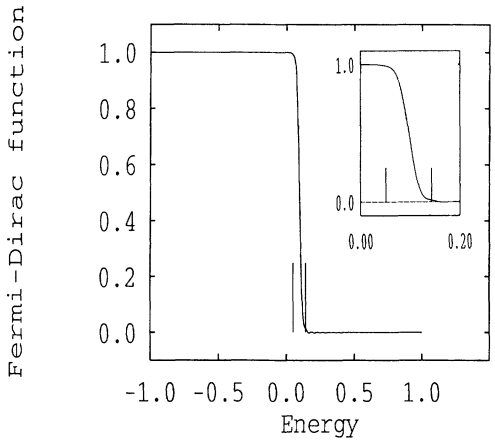

FIG. 1. Plot of the Tchebychev expansion of the Fermi-Dirac function using 150 terms, a β factor of 100, and the Fermi level E_F midway between the conduction and valence band extremum. The top of the valence band and bottom of the conduction band lie at the energies indicated by the small vertical lines. The energy scale has been adjusted so that the entire spectrum lies in the range -1 to $+1$. The inset shows a blow-up near the Fermi-level E_F.

values. An efficient method to extract the energy spectrum from a single vector is the recursion method[8] or the Lanczos method. We use the vector $|z\rangle$ as a starting vector $|z_0\rangle$ and generate the sequence of vectors $|z_0\rangle, |z_1\rangle, |z_2\rangle, ..., |z_M\rangle$. Since only occupied states are contained in $|z_0\rangle$, the sequence should theoretically terminate at $|z_M\rangle$, where $M = n_{\rm occ} - 1$, and $n_{\rm occ}$ is the number of occupied states. As discussed below, this termination is not realized due to instabilities in the three term recurrence. When, due to finite precision arithmetic, a small component of an energy eigenstate spuriously appears in $|z_n\rangle$, the error propagates and grows exponentially in the subsequent recursion vectors $|z_{n+1}\rangle$, $|z_{n+2}\rangle$ and so on. In order to prevent this, we amend the approach.

The sequence obtained from the recursion (Lanczos) method from $M+1$ applications of H is

$$H|z_0\rangle = a_0|z_0\rangle + b_1|z_1\rangle$$

$$H|z_1\rangle = a_1|z_1\rangle + b_1|z_0\rangle + b_2|z_2\rangle$$

$$H|z_2\rangle = a_2|z_2\rangle + b_2|z_1\rangle + b_3|z_3\rangle$$

$$\cdots$$

$$\cdots$$

$$H|z_{M-1}\rangle = a_{M-1}|z_{M-1}\rangle + b_{M-1}|z_{M-2}\rangle + b_M|z_M\rangle$$

$$H|z_M\rangle = a_M|z_M\rangle + b_M|z_{M-1}\rangle$$

The coefficients a_i, and b_i are easy to determine, as discussed in Ref. 8, and their evaluation is an order N operation if the matrix H is sparse. Unfortunately there are order M of these coefficients, so in total their complete evaluation is order $M \times N$.

The recursion sequence generates the tridiagonal matrix $H_{\rm Rec}$

$$H_{\rm Rec} = \begin{pmatrix} a_0 & b_1 & & & & & \\ b_1 & a_1 & b_2 & & & & \\ & b_2 & a_3 & b_4 & & & \\ & & & \cdots & \cdots & & \\ & & & \cdots & \cdots & & \\ & & & & b_{M-1} & a_{M-1} & b_M \\ & & & & & b_M & a_M \end{pmatrix}. \tag{7}$$

The eigenvalues are obtained from this tridiagonal matrix in order M operations.[16] Diagonalization of $H_{\rm Rec}$ bypasses the order N^3 diagonalization of the original $N \times N$ nontridiagonal tight-binding (or plane wave) matrix. The classical Householder tridiagonalization is an $O(N^3)$ procedure, and does not exploit sparseness of the Hamiltonian.[17] And of course the preconditioned starting vector for the Lanczos procedure allows us to work in the occupied subspace, whereas the Householder method does not.

The procedure just described can be used for either tight-binding models or for plane wave models. However, in either case we must face the complications of "ghosting" and "cloning." We define a "ghost" state as an energy eigenstate generated by the recursion method which is not contained in the original starting vector z_0. In the present case, since z_0 lies within the subspace of occupied states, a ghost is defined to be an unoccupied state generated by instabilities in the three term recurrence. Similarly, we define a "clone" as a replica of an occupied state, which is generated after the recursion method has already converged on this occupied state. We make a distinction between ghosts and clones in the electronic structure problem, but the generic phenomenon in Lanczos (recursion) methods is referred to as ghosting.[17]

We now illustrate the formation of ghost states. Start with the vector

$$|z_0\rangle = \frac{1}{\sqrt{N}} \sum_{i \text{ occupied}} |\psi_i\rangle, \tag{8}$$

which contains an equal admixture of each occupied state and no unoccupied states. This vector is artificially created by exactly diagonalizing the $N \times N$ Hamiltonian of our tight-binding GaAs example. After each recursion step, we compute $f_n^{\rm ghost}$,

$$f_n^{\rm ghost} = \sum_{i \text{ unoccupied}} |\langle \psi_i | z_n \rangle|^2 , \tag{9}$$

which is the projection of the recursion vector z_n onto any unoccupied states ψ_i. The logarithm of $f_n^{\rm ghost}$ is shown in Fig. 2. From this figure it is seen that initially the projection of ghost states is machine precision, and then grows exponentially with the recursion step number. After only twenty recursions the vector has nearly unit probability of being in the unoccupied subspace.

Cloning is a similar phenomenon, but refers to states in the occupied subspace. To illustrate this, we again

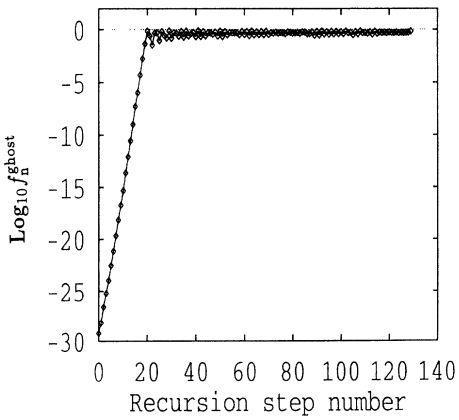

FIG. 2. The logarithm of $f_n^{\rm ghost}$ defined by Eq. (9) which gives the projection of the recursion vector z_n onto any unoccupied state. The starting vector z_0 was taken to be entirely within (machine accuracy) the occupied subspace. The projection initially grows exponentially with each recursive step due to instabilities of the recursion method.

start with z_0 of Eq. (8), and after every recursion project out "by hand" the occupied states from each of the states $|z_n\rangle$ as

$$|z_n\rangle_{\text{corrected}} = |z_n\rangle - \sum_{i \text{ unoccupied}} \langle\psi_i|z_n\rangle|\psi_i\rangle.$$

This removes all ghost states so that we can see the effect of cloning under ideal conditions. We then compute f_n^{clone} defined as

$$f_n^{\text{clone}} = \sum_{i\ \langle n} |\langle z_i|z_n\rangle|^2. \tag{10}$$

Successive states $|z_i\rangle$ generated by the recursion method theoretically are orthogonal to each other, but because of instabilities in the three term recurrence, they are not. The logarithm of f_n^{clone} is shown in Fig. 3. As the recursion process proceeds, the projection of each state $|z_n\rangle$ onto previous recursive states $|z_i\rangle$ $(i < n)$ increases; far less rapidly than for ghosts, but nevertheless by 70 recursive steps, z_n has nearly unit probability of being projected onto earlier states. Thus previous recursive vectors z_i are "cloned" into the current recursive vector z_n. Operationally this means that after many recursions a single eigenstate will appear multiple times.

Once these difficulties are understood, we can adjust the approach so that these factors do not destroy the convergence. We give two examples of approaches that surmount these difficulties.

III. PRACTICAL IMPLEMENTATIONS

We begin by describing the first approach. Ghost states can be removed by performing a recursion not with $\hat{H}$, but rather using the operator $\hat{H}\hat{f}(\hat{H})$. The spectrum of $\hat{H}\hat{f}(\hat{H})$ is $\epsilon_i f(\epsilon_i)$, which effectively places all *unoccupied* eigenvalues at zero energy and occupied states will exist at their correct energy. The states of zero energy do not appear in the calculated spectrum, since they are removed from the recursion vector when the operator is applied to that vector. This method therefore effectively removes ghosts, but requires that each step of the recursion method requires a computation of $\hat{f}(\hat{H})|z\rangle$. As explained earlier, the weight of a ghost state will grow exponentially with the number of recursion levels. By replacing the operator $\hat{H}$ with $\hat{H}\hat{f}(\hat{H})$ we are effectively removing them with exponential weighting.

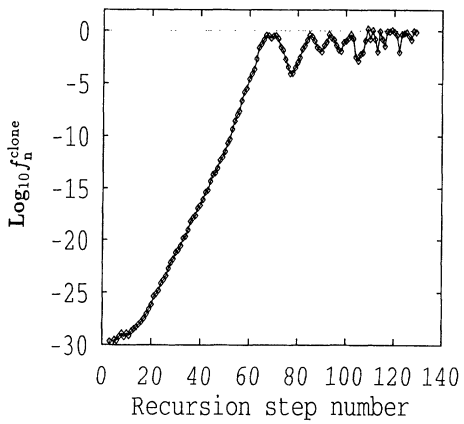

FIG. 3. The logarithm of f_n^{clone} defined by Eq. (10) which gives the projection of the recursion vector z_n onto all previous recursion vectors z_i. It grows due to instabilities of the recursion method.

Cloning is still present, and these extra cloned states makes it necessary to include more recursions M than we have occupied states. If we have n_{occ} occupied states, and perform M levels of recursion of $\hat{H}\hat{f}(\hat{H})$ $(M > n_{\text{occ}})$, then at least $M - n_{\text{occ}}$ states are clones. We increase M until all n_{occ} states are converged. The presence of converged states can be easily tested by the method of "orbital peeling."[18,19] In orbital peeling, the $M \times M$ tridiagonal Hamiltonian H_{Rec} is diagonalized as well as the $(M-1) \times (M-1)$ Hamiltonian H'_{Rec} obtained by removing the first row and column of H_{Rec}. Cloning arises from a loss of orthogonality between the computed recursion vectors. The central idea in orbital peeling is that the clone eigenvector contains the nonzero component of the starting vector, $|z_0\rangle$, in a recursion vector $|z_n\rangle$ which is not (but should be) orthogonal to $|z_0\rangle$. Hence, the clone eigenvalue is unaffected when the first row and column of the tridiagonal matrix are removed. The orbital peeling method works similarly for ghost eigenvalues, which, by definition have a zero component of $|z_0\rangle$, and so are unaffected when $|z_0\rangle$ is removed as a variational degree of freedom. To discard clones, we search for eigenstates of H_{Rec} *that are different* from those of H'_{Rec}—these are candidates for being converged occupied states. Such candidates can be further tested for convergence by comparing $\langle\psi_{\text{Rec}}|\hat{H}|\psi_{\text{Rec}}\rangle$ with ϵ_{Rec}, where $\hat{H}$ is the original full sparse Hamiltonian (not H_{Rec}) and ψ_{Rec} are the eigenfunctions obtained from the recursion method.

The results of this first method are shown in Table I.

TABLE I. Table showing the number of converged occupied states, n_c, as the number of recursive steps M increases. The recursion was performed using $\hat{H}\hat{f}(\hat{H})$ as the effective Hamiltonian and the starting vector was a computer generated random vector. A candidate for convergence was tested by orbital peeling (see text). The number of occupied states, n_{occ}, for this 64 atom GaAs supercell example is 128. The Fermi-Dirac projection operator was expanded exactly as in Fig. 1.

Order of recursion M	Number of converged eigenstates n_c
50	31
100	81
128	104
150	115
175	122
200	125
207	127
208	128
250	128
300	128

Here we show the number of converged states n_c as the order of recursion M is increased. Again, this example is for a 64 atom GaAs supercell which contains 128 occupied states. We use 150 terms in the Tchebychev expansion with $\beta = 100$ as in Fig. 1. We notice that even for a small number of recursions M that we obtain many converged states. One does not need $n_{\rm occ}$ levels of recursion to produce a substantial set of accurate eigenstates. At $M = n_{\rm occ}$, we have converged most (104) of the $n_{\rm occ}$ (128) states. In this example, it takes M=208 to accurately determine all eigenstates. Adding up all eigenvalues and comparing with the exact sum, $\sum_{i=1}^{128} f(\epsilon_i)\epsilon_i$, we find an error of $1.5 \times 10^{-5}\%$. Finally, increasing the level of recursion M beyond the point where all $n_{\rm occ}$ states have converged produces no more states (clones). We note in passing that a single calculation of the "a" and "b" recursion coefficients can be done with a single large M value (we chose 350 in this example) and the results can be analyzed for convergence for any M values less than this.

It must be pointed out that construction of the eigenfunctions is an order N^3 operation, or more precisely $M^2 \times N$. This is because the column vectors obtained from the recursion method (of length M) refer to the basis of states $|z_n\rangle$ which are of length N, and are in the space we actually want the eigenvectors. This is unfortunate, especially since clones force us to generate more vectors than needed. If the sequence were to accurately terminate after $n_{\rm occ}$ steps as it theoretically should (were it not for instabilities) then the states z_n would have been an orthogonal set of states in the occupied subspace. Since these are some unitary transformation of the occupied energy eigenstates, the charge density and band structure energy could be obtained without diagonalization of $H_{\rm rec}$ at all, and the state z_n could be used for charge densities just as eigenstates.

We now briefly describe a second approach. In some situations, it is undesirable to perform a recursion of $\hat{H}\hat{f}(\hat{H})$, particularly since multiplying by $\hat{f}(\hat{H})$ at each recursive step requires multiplication of the (sparse) matrix $\hat{H}$ and a vector K times, where K is the number of terms in the Tchebychev expansion and is on the order of one hundred. Instead, we begin with the random projected vector $|z_0\rangle$ $(= \hat{f}(\hat{H})|x\rangle)$, and allow ghost states to appear. The idea is that generally a ghost state does not appear until some occupied states have well converged in the recursion Hamiltonian $H_{\rm Rec}$. However, certain occupied states converge much more readily than others. By choosing several random vectors, and projecting out from the Hamiltonian these previously converged occupied states, we can obtain the full occupied spectrum of states.

As an example, we generate a projected random vector $|z_0\rangle$ and perform M levels of recursion using it as a starting vector. Here we arbitrarily choose M to be $n_{\rm occ}$ (the number of occupied states, i.e., 128 for this GaAs example). The number M is adjustable, and can actually be made much smaller than $n_{\rm occ}$. Most of the states generated will be ghost states, but a few will be genuine occupied states. These genuine states can be identified in a manner similar to the orbital peeling described earlier in the first approach. The dozen or so converged occupied states shown for a single random vector are shown in the lower panel of Fig. 4. Notice that the states that have converged are near the bottom and top of the occupied spectrum as is typical of a Lanczos-like recursion method.

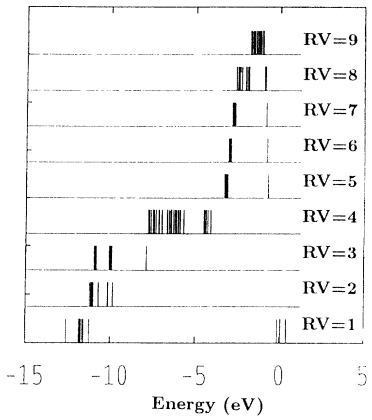

FIG. 4. The spectrum of converged eigenvalues generated by using a set of nine different random vectors (RV's). The eigenvalues are in eV units, with ~ -13 eV being the bottom of the valence band and $\sim +0.5$ eV being the top of the valence band. The first random vector converges to the "easy" states at the extremes of the occupied spectrum. Further random vectors fill in the spectrum sequentially, until all 128 occupied eigenstates are determined with nine (9) random vectors.

Next we choose a second random vector, then a third and so on, and project them onto the occupied subspace by application of the Fermi-Dirac operator. The Hamiltonian used in the recursion method for each subsequent random vector is replaced by a modified projected Hamiltonian

$$\hat{H} \longrightarrow \hat{H} - \sum_{i=1}^{n_c} \epsilon_i^c |\psi_i^c\rangle\langle\psi_i^c|. \tag{11}$$

Here ϵ_i^c and ψ_i^c are the n_c converged occupied eigenvalues and eigenvectors converged so far by all previous random vectors. This Hamiltonian forces the recursion method to generate only new occupied eigenvectors and not the "easy" eigenvectors previously found. The projection term of the Hamiltonian in Eq. (11) is very similar to a Kleinman-Bylander[20] separable pseudopotential, and will computationally have a similar cost in plane wave calculations.

In Fig. 4 we show the results of this procedure using nine (9) different random vectors. The second random vector, shown in the second panel from the bottom, converges to the extremal (lowest) energy eigenstates which were not found by the first random vector. As more random vectors are used, the spectrum fills in from the extremes, until finally for the last random vector (the ninth in this example) the last of the $n_{\rm occ}$ state is determined.

IV. CONCLUSION

In summary, we have further investigated the use of random vectors within the matrix formulation of the

electronic-structure problem. A random vector is appealing since in general it has components on all energy eigenvectors, and, in particular, the occupied states which we seek. We find that a Tchebychev approximation of the Fermi-Dirac projection operator is very accurate and a useful noniterative scheme for projecting the random vector onto the occupied subspace.

We have described two applications of the recursion method which satisfactorily determine the occupied eigenstates. In any application, care must be taken to account for the recursion method's tendency to produce multiple copies of occupied states (clones) or unoccupied eigenstates (ghosts) from a starting vector which is known to have its spectrum in the unoccupied subspace.

The methods we describe still scale with the cube of the number of atoms, but may be competitive with existing order N^3 techniques[13,14] when the occupied subspace is a small fraction of the size of the full space.

Note added in proof. We have recently implemented essentially the present method for *ab initio* molecular dynamics using a plane-wave basis and nonlocal pseudopotentials, and find excellent results compared to conventional plane-wave methods.

ACKNOWLEDGMENTS

We would like to acknowledge informative discussions with Alex Demkov, Jose Ortega, and Dmitri Tchikatilov. This work was supported by the Office of Naval Research under Contract Nos. ONR N00014-90-J-1304 and NSF DMR 93-22412. Pacific Northwest Laboratory is operated for the U.S. Department of Energy by Battelle Memorial Institute under Contract No. DE-AX06-76RLO 1830.

[1] D. A. Drabold and O. F. Sankey, Phys. Rev. Lett. **70**, 3631 (1993).
[2] J. Skilling, in *Maximum Entropy and Bayesian Methods*, edited by J. Skilling (Kluwer, Dordrecht, 1989), p. 455.
[3] R. N. Silver and H. Roder, Int. J. Mod. Phys. C. (to be published).
[4] A good collection of papers may be found in *Theory and Application of Moment Methods in Many Fermion Systems*, edited by B. J. Dalton, S. M. Grimes, J. P. Vary, and S. A. Williams (Plenum Press, New York, 1980).
[5] C. Lanczos, *Applied Analysis* (Dover, New York, 1988).
[6] B. Strohmaier, S. M. Grimes, and S. D. Bloom, Phys. Rev. B **32**, 1379 (1985).
[7] J. Hafner, Phys. Rev. B **27**, 678 (1983).
[8] V. Heine, R. Haydock, and M. Kelly, in *Solid State Physics*, edited by H. Ehrenrich, F. Seitz, and D. Turnball (Academic, New York, 1980), Vol. 35.
[9] A plane wave Hamiltonian is effectively made sparse by working in reciprocal space for the kinetic energy, and real space for the potential energy. The changing of spaces is accomplished by fast Fourier transforms.
[10] O. F. Sankey, A. Demkov, D. A. Drabold, A. Gibson, and D. Tchikatilov (unpublished).
[11] P. Vogl *et al.*, J. Phys. Chem. Solids **44**, 365 (1983).
[12] M. Daw, Phys. Rev. B **47**, 10 895 (1993).
[13] R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).
[14] M. P. Teter, M. C. Payne, and D. C. Allan, Phys. Rev. B **40**, 12 255 (1989).
[15] W. H. Press, Brian P. Flannery, Saul A. Teukolsky, and William T. Vetterling, *Numerical Recipes, The Art of Scientific Computing* (Cambridge University Press, Cambridge, 1986).
[16] We use the EISPACK routine "rst" for this purpose.
[17] G. H. Golub and C. F. VanLoan, *Matrix Computations*, 2nd ed. (John Hopkins University Press, Baltimore, 1989).
[18] J. Cullum and R. A. Willoughby, *Sparse Matrix Proceedings*, edited by I. S. Duff and G. W. Stewart (SIAM, Philadelphia, PA, 1978).
[19] N. R. Burke, Surf. Sci. **58**, 349 (1976), A. Gibson R. Haydock, and J. P. LaFemina, Phys. Rev. B **47**, 9229 (1993).
[20] L. Kleinman and D. M Bylander, Phys. Rev. Lett. **48**, 1425 (1982).