Lawrence Berkeley National Laboratory

Title
A brief comparison between grid based real space algorithms and spectrum algorithms for electronic structure calculations

Permalink
https://escholarship.org/uc/item/9bg6t4tg

Author
Wang, Lin-Wang

Publication Date
2008-06-13

A brief comparison between grid based real space algorithms and spectrum algorithms for electronic structure calculations

Lin-Wang Wang
NERSC System Architecture Team
Lawrence Berkeley National Laboratory

Quantum mechanical ab initio calculation constitutes the biggest portion of the computer time in material science and chemical science simulations. As a computer center like NERSC, to better serve these communities, it will be very useful to have a prediction for the future trends of ab initio calculations in these areas. Such prediction can help us to decide what future computer architecture can be most useful for these communities, and what should be emphasized on in future supercomputer procurement. As the size of the computer and the size of the simulated physical systems increase, there is a renewed interest in using the real space grid method in electronic structure calculations. This is fueled by two factors. First, it is generally assumed that the real space grid method is more suitable for parallel computation for its limited communication requirement, compared with spectrum method where a global FFT is required. Second, as the size N of the calculated system increases together with the computer power, O(N) scaling approaches become more favorable than the traditional direct O(N^3) scaling methods [1]. These O(N) methods are usually based on localized orbital in real space [2], which can be described more naturally by the real space basis. In this report, we will compare the real space methods versus the traditional plane wave (PW) spectrum methods, for their technical pros and cons, and the possible of future trends. For the real space method, we will focus on the regular grid finite different (FD) method and the finite element (FE) method. These are the methods used mostly in material science simulation. As for chemical science, the predominant methods are still Guassian basis method, and sometime the atomic orbital basis method. These two basis sets are localized in real space, and there is no indication that their roles in quantum chemical simulation will change anytime soon. We will focus on the density functional theory (DFT) [3], which is the most used method for quantum mechanical material science simulation.

To calculate a system under density functional theory, especially under its local density approximation (LDA) [3, 4], one needs to solve the minima of the total energy as a function of the single particle wavefunction $\{\psi_i\}$:

$$E[\psi] = \int \left[-\frac{1}{2} \sum_i \psi_i^*(r) \nabla^2 \psi_i(r) + V_{ion}(r)\rho(r) \right] d^3r + \frac{1}{2} \int \frac{\rho(r)\rho(r')}{|r-r'|} d^3r d^3r'$$
$$+ \int \rho(r)\varepsilon_{XC}(\rho(r)) d^3r + \sum_{R,R'} \frac{Z_R Z_{R'}}{|R-R'|}$$

(1)

here the occupied charge density is

$$\rho(r) = \sum_i |\psi_i(r)|^2 \qquad (2)$$

and $V_{ion}(r)$ is the ionic potential, and Z_R is the nuclei charge at R, and the function $\varepsilon_{XC}(x)$ is the LDA exchange correlation function. The wavefunctions ψ_i satisfies the following orthonormal condition:

$$\int \psi_i^*(r)\psi_j(r)d^3r = \delta_{i,j} \qquad (3)$$

Finding the minimum of the energy in Eq.(1) is equivalent to find the solution for the following Kohn-Sham equation [4] (sometime it is also called Schrodinger's equation):

$$[-\frac{1}{2}\nabla^2 + V(r) + \hat{V}_{NL}]\psi_i(r) = \varepsilon_i\psi_i(r) \qquad (4)$$

here the total potential V(r) has the following expression:

$$V(r) = V_{ion}(r) + \int \frac{\rho(r')}{|r-r'|}d^3r + \mu_{XC}(\rho(r)) \qquad (5)$$

here the second term is the Coulomb potential due to charge $\rho(r)$, and the third term is the LDA exchange correlation potential coming from the derivative of $\rho\varepsilon_{XC}(\rho)$. In Eq.(4), we have also introduced a nonlocal potential $\hat{V}_{NL}$, which is a nonlocal operator acting on the wavefunction. This nonlocal potential is needed for pseudopotential calculations [1].

Given the above formalism, the different numerical methods distinguish themselves by their ways to represent the wavefunction ψ_i. In the plane wave (PW) method, the wavefunctions are expanded by plane wave basis set as:

$$\psi_i(r) = \sum_q C_i(q)e^{iq\bullet r} \qquad (6)$$

Usually a periodic box (supercell) is chosen. Then the reciprocal lattice of the supercell defines a grid of q in the Fourier space. As a convention, all the q points within a sphere defined by a kinetic energy cutoff $E_c = \frac{1}{2}q_c^2$ is chosen in the summation of Eq.(6). In the PW method, the wavefunctions are kept in reciprocal space (q-space), represented by the coefficients $C_i(q)$. Major operations, like the enforcement of the orthonormal conditions of Eq.(3) are carried out in this reciprocal space representation. However, to carry out operations like the $V(r)\Psi_i(r)$, the wavefunction is transformed via FFT into the real space on a regular real space grid. After the V(r) and $\Psi_i(r)$ multiplication, the result is then

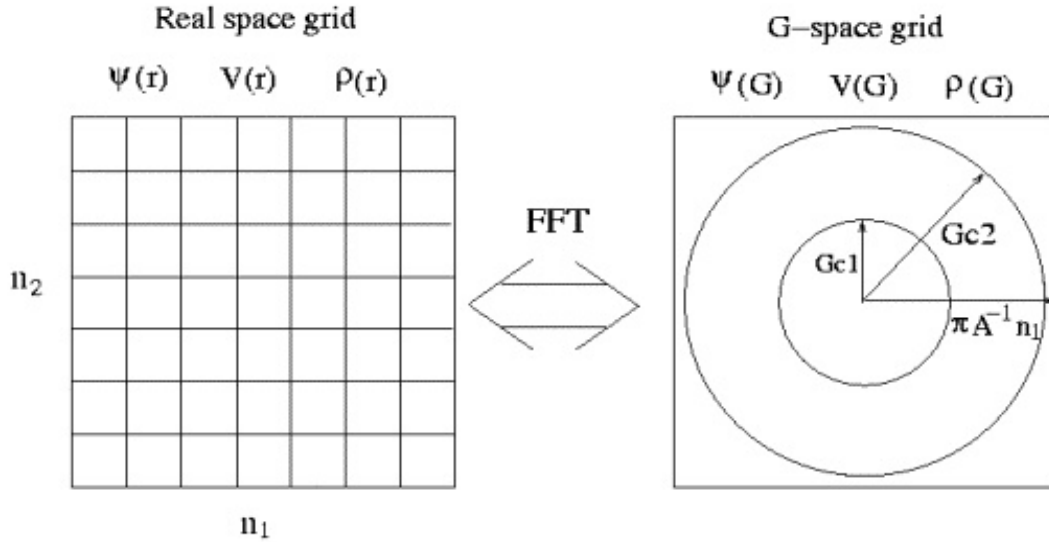Fourier transformed back to the reciprocal space. This due space representation is illustrated in Fig.1.

Fig.1, the due space representation of the PW method. The reciprocal space (right box) and real space (left box). The wavefunction q vectors (G=q in this figure) inside the cutoff Gc1 are chosen in the summation of Eq.(6), while the potential V and charge density ρ are represented by the planewave q inside a larger cutoff Gc2. Gc1 is half of Gc2. FFT is used to transform the wavefunction from reciprocal space representation to a real space grid.

The PW formalism has the following advantage. Given the wavefunction expression in Eq.(6), the local minimum of Eq.(1) can be evaluated almost numerically exactly without further approximations. The only numerical approximation for the evaluation of Eq.(1) is the calculation of the LDA exchange correlation term (the fourth term) in Eq(1), but the error is found to be very small [1]. Thus, the solution is numerically exact, and it is truly variational. The only approximation is the use of limited planewave basis function in Eq.(6). This numerical exactness is only rivaled by the Gaussian method in quantum chemistry calculation, where the numerical integrals have analytical expressions.

The numerically variational feature provide the following two properties: (1) the total energy found from Eq.(1) is always an upper bound of the exact solution; (2) the error in the energy ΔE is proportional to the square of the error in the wavefunction $\Delta \Psi$. Besides, a variational solution allows one to calculate the atomic forces using the Hellman-Feynman theory [1]. This is very important for atomic relaxation and molecular dynamics simulations.

The computational disadvantage of the PW method for large scale parallel computation is the requirement for the FFT. While the number of floating point operation is moderate, the global communication for the FFT can be a serious bottleneck. Efficient FFT have been demonstrated using a few thousand processors in electronic structure calculations

[5]. But for even larger number of processors, the communication message size can be very fragmented, and the communication latency can significantly slow down the FFT. However, a critical difference between the electronic structure calculations and many other calculations (e.g., fluid dynamics and classical electric-magnetic field simulation) is that there are many wavefunctions ψ_i (the index i) to be calculated in electronic structure calculations. For a few thousand atom system, the number of the wavefunctions is also a few thousands. This provides another way of parallelization, that is a parallelization among the wavefunction index i. Using this double parallelization scheme (one on the wavefunction index, another on the planewave vector q), Q-box has successfully run on >100,000 processors on the IBM BG/L computer [6].

To avoid the possible impediment of a large global FFT, and to implement the localized orbital O(N) method, many researchers have started to explore the real space grid methods. There are two most widely explored real space grid methods, one is the finite element (FE) method, another is the finite difference method (FD). Let's first discuss the FE method.

There are many possible FE basis set [7]. One commonly used one is the so called C^0 FE basis set, within which each basis function $\Phi_m(r)$ is a continuous function within a finite local region crossing a few grid points in each direction. The basis function $\Phi_m(r)$ is one at one grid point (r_m), and zero at all the other grid point. As a result, the wavefunction ψ_i is expanded as:

$$\psi_i(r) = \sum_m C_i(m)\Phi_m(r) \qquad (7)$$

and $C_i(m) = \psi_i(r_m)$, and here r_m is the grid point.

C^0 FE basis set does not require the basis function to be smooth, i.e., its derivative could be not continuous, and its Laplacian (first term in Eq.(1)) could be ill defined. However, this problem can be avoided by rewriting the first term in Eq.(1) as: $\int \nabla \psi_i^*(r) \bullet \nabla \psi_i(r) d^3r$. There could be different orders of C^0 FE basis functions. The linear order $\Phi_m(r)$ is a piece wise straight line between grid points (or a linear function in a 3D region). A quadratic $\Phi_m(r)$ is a second order function of r, while a cubic $\Phi_m(r)$ is a third order polynomial. Regardless of its order, one $\Phi_m(r)$ is only nonzero within a few grid points interval in any direction. This makes the basis set spatially local. With the FE basis set, the Eq.(1) and the corresponding Eq.(4) can be solved variationally [8, 9]. But not everything has been worked out, and there are many complications. Imagine a Nth order polynomial is used for the basis function $\Phi_m(r)$ to describe the wavefunction ψ_i, then because the charge density $\rho(r)$ calculated from Eq(2) is a quadratic function of the wavefunction, a 2Nth order FE basis function is needed to describe the charge density. Further more, it is not clear what order is needed to describe the potential in Eq.(5). If the potential V(r) is not described accurately in a given order of FE basis, then the corresponding matrix element $\int \Phi_m(r)V(r)\Phi_n(r)d^3r$ for the Kohn-Sham equation cannot

be calculated accurately. This can introduce serious errors. Currently, the most discussion in the literature is centered around how to calculate the Laplacian term: $\int \psi_i^*(r)\nabla^2\psi_i(r)d^3r$, but not at the potential term. The potential term is further complicated by the need to solve the Poisson equation. Currently, the Poisson equation is solved in real space by multilevel method, either use FE basis or FD method [10]. But whether the solved potential described in a FE basis or on a grid contains enough information to evaluate $\int \Phi_m(r)V(r)\Phi_n(r)d^3r$ accurately is not fully investigated. The central question is: how big a numerical approximation is involved when we evaluate the total energy in Eq.(1) based on the wavefunction expression in Eq.(7). For the planewave approach [1], there are techniques at every step of the total energy evaluation to make sure that the final total energy is calculated exactly given the original wavefunction expansion of Eq.(6). In the current FE formalism, sometime higher density grid is used to calculate ρ(r) and V(r), e.g., a double grid had been used [9]. Although in planewave formalism, a double grid can ensure that the discrete numerical summation result will be the same as the exact analytical result [1], there is no such guarantee in the FE method.

Despite the above uncertainties in the formalism, it has been shown that the FE method can result is a variational solution of the LDA total energy. When the grid size h decreases, the total energy also decreases, approaching the exact value (calculated by the PW method) from above. Thus, the solutions are lower bounded. Using 3D 32-node "serendipity" finite elements, which satisfies the cubic completeness [7], it has been shown that the total energy converges to the exact result as $O(h^{-6})$ [8]. Unfortunately, in that work of Pask et.al. [8], it has not been shown whether the ΔE scales as the square of ΔΨ, an more stringent requirement for a variational solution.

Computationally, the FE is more local and compact than the FD method. Only short range neighboring grid interaction (hence communication) is needed. It produces very sparse matrix, and fits well with large parallel computation. Besides, mature computer packages exist for parallel FE calculation. FE has also been tested based on nonuniform grid and adaptive grid, especially for all electron calculations. Unfortunately, it is found that huge number (tens of thousands) of finite element basis per atom is required for such all electron calculations [11,12]. Thus, currently, pseudopotential is used to reduce the number of basis set, and regular equal space grid is used to generate the finite elements. Overall, despite the maturity of the FE method itself in other fields, like the civil engineers and fluid dynamic, and the existence of FE computer packages, its usage in electronic structure calculations is still very limited. Partly this is due to the relative complication of its formalism, and partly it is because there are still technical uncertainties (e.g., the solution and representation of the potential function to guarantee the variational solution of the total energy). Further algorithmic research is still needed to make the FE method more widely accepted by the electronic structure community. At this stage, it is difficult to say how large a role the FE method will play in the future of electronic structure calculations.

We now discuss the FD method. FD method benefits from the fact that it is very similar to the planewave method, which is widely known in material science simulation

community. In the FD method [13], the wavefunction $\psi_i(r)$ is represented by its values at the grid points. This is a discretization approach of the exact equations (1) and (4). There is no basis set for the wavefunction expansion, and the subsequent numerical evaluation of the total energy in Eq.(1) is also discretized. In the FD approach, the Laplacian is evaluated by a finite difference formula:

$$\frac{\partial^2 \psi}{\partial^2 x}(x_i, y_j, z_k) = \frac{1}{h_x^2} \sum_{n=-M/2}^{M/2} A_n^M \psi(x_i + nh_x, y_j, z_k) \tag{8}$$

This is an Mth order formula. The error could scale as $O(h^{M-1})$ for a smooth function $\psi(r)$. Note that, in a plane wave method, we have

$$\nabla_x^2 \psi(x_i, y_j, z_k) = \sum_q q_x^2 C(q) e^{iq \bullet r_i} = \sum_{i'} A(x_i - x_{i'}) \psi(x_{i'}, y_j, z_k) \tag{9}$$

This is essentially the same as the Eq.(8). The only difference is that, in Eq.(9), the convolution summation is carried out over all the grid point i', while in Eq.(8), the summation is done up to the M/2_th neighbor. But for large enough M, the A_n^M in Eq.(8) can be chosen to approach $A(x_i - x_{i'})$ in Eq.(9).

In practice, one major difference between the FD method and the PW method is that, in FD method the wavefunction is kept in real space grid, while in the PW method, it is kept in reciprocal space as in Eq.(6). If the same real space numerical grids are used for these two methods, then due to the use of half sized sphere in the PW method as shown in Fig.1, the number of coefficients in the PW method is more than 8 times smaller than the real space FD method.

In terms of formalism, the FD method is almost exactly the same as the PW method, except the way in which the Laplacian is carried out, and the book keeping of the wavefunction (one in real space, one in reciprocal space). But due to the lack of basis set in FD, it loses the variational feature of the PW and the FE calculations. For one thing, the total energy can approach the final energy from below as the grid point increases. This is shown in Fig.2. Furthermore, the ΔE does not scale as the square of $\Delta \Psi$. Instead, it just scales as $\Delta \Psi$ itself. This is shown in Fig.3, taken from Ref.14. For the same ΔE, and the atomic forces, it is found that [15] the number of grid point (n_1, n_2, n_3) of FD method should be between 0.5 to 1 of the number of grid points used in the PW method. As a result, the number of coefficient (thus the memory) of the real space FD method could be about 4 times that of the PW method for the same accuracy calculation. This could be a serious draw back for the FD calculation. Not only it increases its memory requirement, it increases the computation to enforce the orthonormal condition of Eq.(3). Because for large system calculations, the orthogonalization enforcement takes much more time (it is the origin of the $O(N^3)$ scaling) than the Schrodinger's equation application of the wavefunction (Eq.(4), the FFT), thus the increase of the wavefunction

coefficient can slow down the overall calculation, and makes the FD method slower than the PW method.
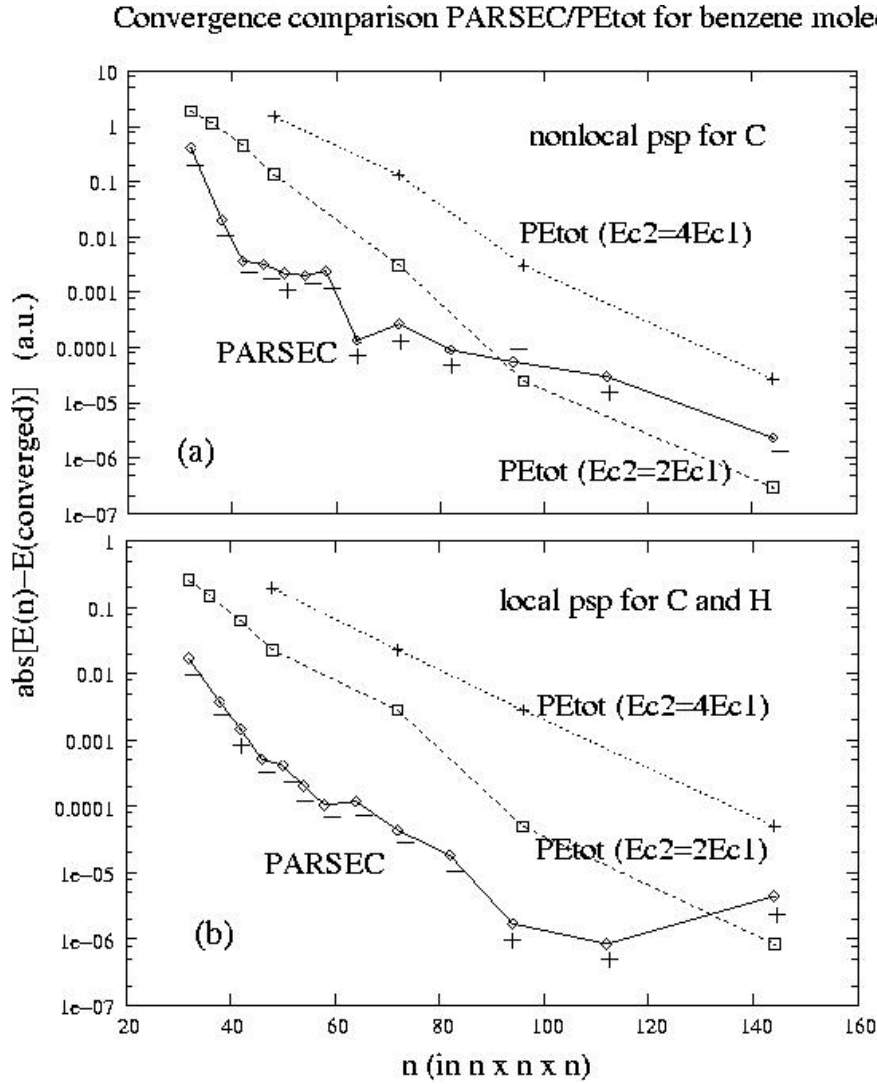

Convergence comparison PARSEC/PEtot for benzene molecu

(a) nonlocal psp for C
PEtot (Ec2=4Ec1)
PARSEC
PEtot (Ec2=2Ec1)

(b) local psp for C and H
PEtot (Ec2=4Ec1)
PARSEC
PEtot (Ec2=2Ec1)

abs[E(n)−E(converged)] (a.u.)

n (in n x n x n)

Fig.2, the convergence comparison between the FD code PARSEC and the PW code PEtot. A C_6H_{12} benzene molecule is calculated inside a supercell. The cubic supercell has an n^3 real space grid. The errors of the total energies are plotted against the size of the grid n. In (a), nonlocal pseudopotential is used for the C atom, while in (b) only local pseudopotentials are used. In the PW PEtot calculation, the size of the grid n is directly related to the planewave cut off Ec2 ($= 0.5Gc_2^2$), as shown in Fig.1. In a normal calculation, as shown in Fig.1, Gc1=0.5Gc2, thus Ec2 = 4 Ec1 ($= 0.5Gc_1^2$). However, for a fixed n, one can always make the Ec1 (Gc1) larger in the program (although it might lose some of the exactness in the total energy evaluation of Eq.(1)). Thus, we have also shown the PEtot results with Ec2=2Ec1. In terms of energy, it approaches the final result faster. The PEtot results approach the final result from above, while the PARSEC result for a given n can be either larger (denoted by +) or smaller (denoted by -) than the final result.

Another draw back of the FD method is the difficulty to use preconditioner in the iterative solution of the Schrodinger's equation. Although multigrid method can be used as the preconditioner, in practice, due to its coding complication, and due to the intrinsic atomic length scale (multigrid cannot be grosser than the characteristic atomic features), the usage of the multigrid technique is limited, and in many cases not used at all. However, due to the lack of efficient preconditioner, different methods than the conventional preconditioned conjugate gradient method [1] has been developed to solve the Schrodinger's equation. These methods (restart Lanczos, and Davison methods) do not use the preconditioner [16]. Recently, a Chebyshev polynomial filter method is developed. What is really helpful for this filter method is its reduction of the required orthogonalization steps. Using this method, a 10,000 atom Si quantum dot has been calculated (although it took more than a month to converge the result on a cluster machine).
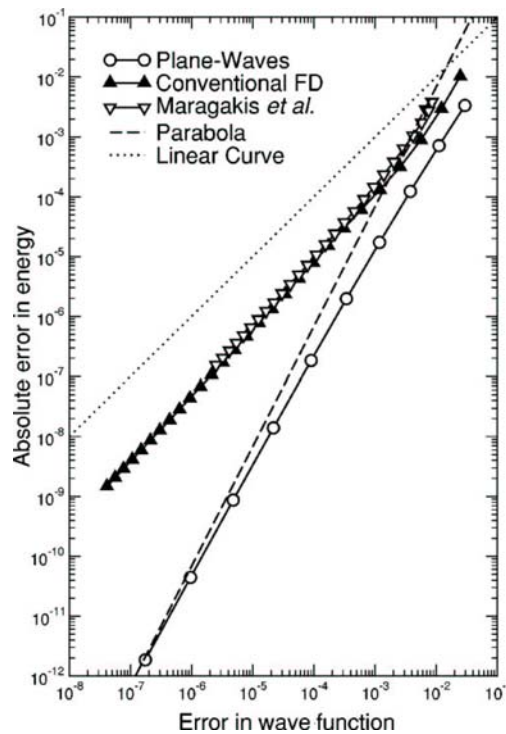

Fig.3, the error in the total energy versus the error in the wavefunction (taken from Ref.14). For the conventional FD and the modified FD method by Maragakis et al. [15], the absolute energy error scales as the first power of the wavefunction error. However, for the plane wave method, the energy error scales as the square of the wavefunction error, a property of variational solution.

As shown in Fig.2, one consequence of losing the variational principle is that the calculated total energy for a given grid size can be smaller than the exact result. This is mostly due to the finite difference kinetic energy expression. Because the kinetic energy in Eq.(8) is a convolution in real space, it can be expressed as a multiplication in

reciprocal space: $\sum_q f(q)C(q)\exp(iq \bullet r)$. Here f(q) is the Fourier transformation of A_n^M used in Eq.(8). It turns out, for the A_n^M used in conventional FD formulism, f(q) is smaller than $0.5q^2$ for large q. Thus, for large q (small number of grid points), the FD kinetic energy is a under estimation. This under estimation leads to a smaller kinetic energy, hence a total energy smaller than the exact solution. To correct this, Maragakis et.al [17] have introduced a new A_n^M which ensures that the f(q) is always larger than $0.5q^2$ for all q. This makes the FD result always approaches the exact solution from the above. Unfortunately, this also makes the method converges slower. More importantly, this still does not make ΔE scale as the square of ΔΨ. Like the original FD method, the ΔE scales in the same order as ΔΨ, as shown in Fig.3.

A recent significant advance of the FD method in electronic structure calculation is the development and release of the PARSEC package by Chelikowsky and Saad's group [16, 18, 19]. This is a parallel code with several new algorithms implemented. As reported in Ref.18, it can be faster than the traditional PW code. For a fair comparison, however, one has to use a grid which provides the same accuracy as in a standard PW code. Unfortunately, we have so far not seen a comprehensive study comparing the FD results from PARSEC with the PW results. We have encountered cases where the published PARSEC results significantly differ from our own PW calculations. But fairly large systems (a few thousand atoms) have been calculated using the PARSEC package, especially with their Chebyshev filter algorithm. Sometime to calculate such large systems (e.g., quantum dots), group theory is used to take the advantage of the symmetric of the system [16], and block diagonalize the Hamiltonian matrix. This is a technique heavily used in the early days (70's-80's) of ab initio calculations. It is currently not used in most ab initio calculation package (due to the use of iterative method). Overall, there is no doubt that with the release of this code, the FD method will play a significant part in large scale electronic structure calculations.

The parallel implementation of PARSEC is not as straight forward as one might think. The high order FD method (e.g, M=12) makes the communication not so local, especially if the number of processors is large for a given grid size. Fortunately, there are many well established routines to partition the real space grid data to achieve a load balance and reduce the communication cost. This is an area heavily studied in the applied math community. In PARSEC, the package Metis [20] is used to partition the grid in a distributed memory parallel machine. Thus, in short, the grid is not distributed in regular cubic as one might imagine. Instead, special partitioning routines are used to help with this task. So far, PARSEC can scale to about hundred processors. But it is yet to be proved that it can scale to thousands of processors.

One of the biggest drive for real space method is the O(N) algorithms which usually employ the localized orbitals. PARSEC is still a $O(N^3)$ code. As far as the traditional $O(N^3)$ algorithm is concerned, the PW approach with its double parallelization in both q space and wavefunction index will probably still be adequate for future large computers. However, when the computer gets larger, and the systems to be simulated also get larger,

some kind of O(N) approach is unavoidable. Does this mean the future of quantum mechanical material science simulation will be dominated by real space codes? This is not necessarily true. First, there are many ab initio methods, especially the ones beyond LDA, are based on planewave formalism (for example, the GW method). These methods are mostly naturally implemented by planewave codes. Second, as for O(N) methods, PW method can still play an important role. For example, Payne's group has developed a "FFT box" technique to describe the localized orbitals [21]. The FFT involved is not the full FFT of the whole system, rather it is carried out in a fixed size box independent of the overall size of the whole system. Similarly, we have developed a linear scaling three dimensional fragment (LS3DF) method [22], which solves the small fragments of a large system using PW method. Again, the FFT involved does not cover the whole space. It is also a local FFT. Thus, some kind of local FFT can still play a role in the future O(N) methods.

In summary, the real space method will certainly play an important role in future electronic structure calculations. This is especially true for the FD method, mostly due to its conceptual simplicity and similarity to the PW method. However, more studies are needed to establish the accuracy and the true computational costs of these methods. Due to the possible increase of the number of expansion coefficients, some part of the calculations (e.g, the orthogonalization part) can be more expensive than the PW method. It is also needed to establish their scalability in large parallel computation. Some of the real space methods (e.g,, the high order FD method) might not be so straight forward to be implemented. Very often, their implementations rely on existing FD and FE software. The real space method might play a major role in future O(N) algorithms. But the parallel scalabilities of such codes are still need to be demonstrated. PW code will not disappear in future electronic structure calculations. It might still play a dominant role. However, it is not likely that the algorithm will require a full machine and whole physical system FFT for larger than 10,000 processor computers. Instead, a double parallelization over the PW vector q and wavefunction index i might be used. As a result, the FFT will only be run on part of the machine (e.g, on $N_{proc}^{1/2}$ number of processors, where N_{proc} is the total number of processors used in the calculation). The PW method can also be used in future O(N) methods, where the FFT will be carried out on a small number of processors (or even on a single processor) over a fixed physical size (not the whole physical system). Thus, when choosing future machines, the speed of FFT will continue to be a major factor for electronic structure calculations. However, one might not need a full FFT using all the computer processors.

[1] M.C. Payne, M.P. Teter, D.C. Allan, T.A. Arias, and J.D. Joannopoulos, Rev. Mod. Phys. **64**, 1045 (1992).
[2] S. Goedecker, Rev. Mod. Phys. **71**, 1085 (1999).
[3] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
[4] W. Kohn and L.J. Sham, Phys. Rev. **140**, A1133 (1965).
[5] A. Canning, L.W. Wang, A. Williamson, A. Zunger, J. Comp. Phys. **160**, 29 (2000).

[6] F. Gygi, et.al., Proceeding of Supercomp. 2005 (ACM, 2005).

[7] H. Kardestuncer, et.al., *Finite Element Handbook*, (McGraw-Hill, New York, 1987).

[8] J.E. Pask, B.M. Klein, C.Y. Fong and P.A. Sterne, Phys. Rev. B **59**, 12352 (1999).

[9] J.E. Pask, P.A. Sterne, Modelling Simul. Mater. Sci. Eng. **13**, R71 (2005).

[10] J.E. Pask, P.A. Sterne, Phys. Rev. B **71**, 113101 (2005).

[11] S.R. White, J.W. Wilkins, M.P. Teter, Phys. Rev. B **39**, 5819 (1989).

[12] E. Tsuchida and M. Tsukada, Phys. Rev. B **52**, 5573 (1995).

[13] T.L. Beck, Rev. Mod. Phys. **72**, 1041 (2000).

[14] C.K. Skylaris, O. Dieguez, P. D. Haynes, and M.C. Payne, Phys. Rev. B **66**, 73103 (2002).

[15] J.-L. Fattebert and F. Gygi, Phys. Rev. B **73**, 115124 (2006).

[16] L. Kronik, A. Makmal, M.L. Tiago, M.M.G. Alemany, M. Jain, X. Huan, Y. Saad, J.R. Chelikowsky, Phys. Stat. Sol. (b), **243**, 1063 (2006).

[17] P. Maragakis, J. Soler, and E. Kaxiras, Phys. Rev. B **64**, 193101 (2001).

[18] A. Stathopoulos, S. Ogut, Y. Saad, J. Chelikowsky, and H. Kim, Comput. in Sci. Eng. 19, July/Aug. (2000).

[19] http://www.ices.utexas.edu/parsec/index.html

[20] G. Karypis and V. Kumar, Proc. 10th Int'l Parallel Processing Symp., IEEE Computer Soc. Press, Los Alamitos, Calif., 314 (1996).

[21] C.K. Skylaris, A.A. Mostofi, P.D. Haynes, O. Dieguez, M.C. Payne, Phys. Rev. B **66**, 35119 (2002).

[22] L.W. Wang, Z. Zhao, J. Meza, Tech. Rep. LBNL-61691, Lawrence Berkeley National Laboratory (2006).