# Multicenter Integration Scheme for Electronic Structure Calculations of Periodic and Nonperiodic Polyatomic Systems

**Zijing Lin,† John E. Jaffe, and Anthony C. Hess***

*Condensed Matter Theory Group, Material Science Department, Pacific Northwest National Laboratories, Richland, Washington 99352*

*Received: September 10, 1998; In Final Form: December 28, 1998*

We present a numerical integration scheme designed to treat the type of multicenter integrals encountered in electronic structure calculations. By developing a notation that differentiates between those atomic centers where integrands have significant amplitudes and those where they do not, we find a way to decompose multicenter integrals (into sums over one-center integrals) such that the number of operations needed for a given matrix element does not increase with increasing system size. In addition, a new adaptive one-center grid is presented that accounts for the shell structures of core electrons while allowing for the vastly different behavior of integrands in the valence and tail regions. Through the use of model integrands the necessary grid points are automatically generated for a given system based on the accuracy requested. Our new multicenter decomposition scheme and one-center grid have been tested separately and in conjunction with each other. Results of such tests demonstrate that our decomposition scheme combined with our one-center grid provides significant improvements over existing multicenter integration schemes. In addition to demonstrating the efficiency of the method for any size system, we will show that the CPU cost of an integral remains constant for systems larger than some easily achievable threshold size. In general comparison shows that the larger the system, the higher is the percent gain in efficiency over previously published methods. In addition, the higher the accuracy targeted, the higher percentage the gain. Also, the higher the accuracy required for a given system, the higher is the gain in efficiency. The method is therefore of great use for large polyatomic molecules and periodic systems.

## I. Introduction

The Hohenberg−Kohn[1] theory states that given a suitable density, $\rho(r)$, the exchange and correlation contributions to the total energy of the system can be obtained from a three-dimensional integral of the general form

$$I = \int F(r)\, \mathrm{d}^3 r = \int F(\rho,(r),\nabla\rho(r),...)\, \mathrm{d}^3 r \qquad (1)$$

The integrand, $F(r)$, in the above expression extends over all space and is, in turn, a function of the density $\rho(r)$. Even if the exact form for $F$ were known, it would not be well approximated by simple mathematical forms or expansions in $\rho(r)$. In fact, all functional forms currently in use result in integrands that cannot be integrated analytically. To complicate matters further, the integrands associated with this class of physical problem all possess cusp singularities at the locations of the atomic nuclei that cause the entire problem to be rather poorly behaved, in a mathematical sense, and require specialized strategies to obtain high-quality results. Computational density functional schemes,[2] particularly those that employ the style of localized basis functions commonly used in quantum chemistry, require efficient and accurate numerical integration packages to be of practical value.

In what follows, we will describe an approach that is efficient, accurate, and scales very favorably with system size. We have done this by making improvements to three areas of numerical integration technology. First, we have formulated a new scheme to remove the cusp singularities in the integrands. We will show that this strategy greatly reduces the degree of angular variation in the product of the integrand and the integration weight compared to previously published methods. Second, our multicenter decomposition scheme is capable of exploiting the fact that not all regions of space contribute significantly to the total value of a given integral. The scheme draws upon economies inherent in the use of a local basis to identify and carefully integrate regions of space that make large contributions to the integral while spending less effort on regions that make very small contributions. This aspect of our approach becomes critically important when treating crystalline systems that employ periodic boundary conditions, since a straightforward sum over the "infinite" number of atoms is not only unnecessary but impractical. Finally, we have adopted a modified one-center integration scheme that better accounts for the atomic shell structure. We will demonstrate that most of the dramatic efficiency gains are achieved via the new multicenter decomposition scheme.

Section II presents the mathematical framework of our multicenter decomposition scheme. Section III discusses the details of our one-center integration scheme and automatic grid point generator. To compare our one-center integration strategy to those previously published and to allow us to clearly demonstrate where our efficiency gains are originating (since we have changed both the one-center integration and the multicenter decomposition strategies), we will adopt Becke's[3] decomposition scheme when testing our one-center integration

* To whom correspondence should be addressed. E-mail: ac_hess@pnl,-gov. Fax: (509) 375-2426.
† E-mail: zijinglin@usa.net. Fax: (604) 709-6162.

10.1021/jp983668x CCC: $18.00 
Published on Web 03/13/1999

grid. Section IV contains the results of using both our decomposition scheme and our one-center integration strategy on the set of test systems previously proposed by Perez, Becke, and San Fabian (PBS) in section IV. To demonstrate the practical feasibility of this scheme in self-consistent-field calculations, we also present results for bulk silicon and aluminum obtained using our periodic SCF Gaussian basis density functional program GAPSS. Concluding remarks can be found in section V.

## II. Theory

Numerical integration schemes for quantum mechanical[3−15] applications can be roughly divided into two groups. In the first group of methods, the cusps (located at nuclear positions) are treated by partitioning space into cells commensurate with the atomic organization of the system. Once the total space of integration has been divided into pieces (which can result in complicated geometric shapes), integration is then carried out on all of the various subcomponents. This strategy is known as the "cellular" approach and has been reviewed by Te Velde and Baerends.[5] An alternative to the cellular approach partitions the multicenter integrand into subcomponents by employing one-center nuclear weight functions.[3,6] Since our approach employs such weight functions and certain aspects of our strategy rely on ideas previously described by both Becke[3] and Delley,[6] we will begin this section with a brief discussion of the general characteristics of methods adopting this approach. The primary goal of this section is to develop the mathematical expressions that define our multicenter decomposition scheme.

**IIa. Popular Approach**. We begin by defining the general role of nuclear weighting functions[3] and establishing relationships between certain quantities that will appear in our equations that are either identical to or are closely related to quantities previously defined by other researchers.[3,6] First, consider a one-center partition function, $\omega_i(\vec{r})$, which satisfies the normalization condition

$$\sum_i \omega_i(\vec{r}) = 1 \tag{2}$$

It is important to note that the summation index, $i$, in eq 2 extends over all nuclear centers in the polyatomic system. By use of this partition function, a multicenter function, $F$, can be decomposed into a sum of one-center components,

$$F(\vec{r}) = \sum_i F_i(\vec{r}) \tag{3}$$

where

$$F_i(\vec{r}) = \omega_i(r_i)F(\vec{r}) \tag{4}$$

Thus, the three-dimensional multicenter integral,

$$I = \int F(\vec{r})\, \mathrm{d}^3 r \tag{5}$$

can be reduced to a sum of single-center integrations,

$$I = \sum_i I_i \tag{6}$$

where

$$I_i = F_i(\vec{r})\, \mathrm{d}^3 r \tag{7}$$

As will be seen below, the one-center partition function, $\omega_i(\vec{r})$, must be properly chosen if the one-center function, $F_i(\vec{r})$, is to be well behaved and high numerical accuracy and efficiency are to be achieved.

A common approach to constructing one-center partition functions, originating from Becke,[3] employs a cutoff profile, $S(\mu_{ij})$, for atoms $i$ and $j$, of the form

$$S(\mu_{ij}) = \frac{1}{2}\{1 - p\{p[p(\mu_{ij})]\}\} \tag{8}$$

where

$$p(\mu) = \frac{3}{2}\mu - \frac{1}{2}\mu^3 \tag{9}$$

and

$$\mu_{ij} = \frac{r_i - r_j}{R_{ij}} \tag{10}$$

Here, $r_i$ and $r_j$ represent distances to atom $i$ and $j$, respectively, from the point $\vec{r}$ and $R_{ij}$ is the internuclear separation. By use of these definitions, Becke's[3] one-center nuclear weight function is then defined by

$$\omega_i(r) = P_i(r)/\sum_i P_i(r) \tag{11}$$

where

$$P_i(r) = \prod_{j \neq i} S(\mu_{ij}) \tag{12}$$

Note that the one-center weight function is composed of a product of the cutoff functions, $S(\mu_{ij})$. The partition function $S(\mu_{ij})$ is selected in such a way that it takes on the value of unity (or very near unity) for $r$ at and near nucleus $i$ but vanishes (or nearly vanishes) at or near nucleus $j$. An additional property of the one-center weight functions is that $S(\mu_{ij}) + S(\mu_{ji}) = 1$. The net effect of multiplying $S(\mu_{ij})$ and a multicenter integrand is to remove the cusp centered at atom $j$ while leaving the cusp at center $i$ intact.

Our approach, to be described in the following subsection, also uses Becke's[3] cutoff profile function. However, our strategy only uses this function to perform "two-center" decompositions (where it is referred to as $S_{ij}$) while for those cases where $i = j$, a value of $S_{ij} = {}^1/_2$ is used. On the basis of the work of Hirshfeld,[16] Delley[6] has suggested that the spherical atomic density, $\rho_i(r)$ (or related functions such as $\rho_i^2(r)$ and $\rho_i(r)/r^2$), should be used to define $P_i(r)$ in eq 11. Our approach contains elements of both Becke's and Delley's schemes by using slightly modified cutoff profiles in conjunction with physically motivated one-center weight functions.

**IIb. New Multicenter Decomposition**. In quantum mechanical applications the dominant computational cost is the evaluation of integrals (matrix elements) of the form $A = \int \chi_{\alpha_i} V \chi_{\beta_j}\, \mathrm{d}\vec{x}$. For the purposes of this study we specifically wish to consider the basis functions $\chi_{\alpha_i}$ to be localized functions centered on atom "$i$" with the compound index "$\alpha$" denoting the angular and radial properties of the function. Since we are interested in electronic structure applications, $V$ can be assumed to be the exchange−correlation potential $V_{xc}$, the Hartree Coulomb potential is $V_H$ or their sum, and the Kohn−Sham effective potential is $V_{eff}$. In all cases the presence of density cusps or Coulomb singularities will introduce sharp features into the various integrands at the location of the nuclei. In fact, the largest contribution to these

integrals typically originates from regions of space that are in proximity to the nuclei. The total integrand, however, is the product of the potential and the basis functions used to form the matrix representation of the operator. The spatial behavior of the potential and that of the basis functions can be very different.

In this regard, we will explore the computational economies that can be derived from using "localized" basis functions to expand the ground-state wave function. The creation of suitable atom-centered local basis functions by forming linear combinations of Gaussian functions[17] is an example of an approach commonly used in computational chemistry[18] and physics.[19] With the exception of diffuse polarization functions, these basis functions $\chi_{\alpha_i}$ typically approximate atomic one-electron eigenfunctions and hence are strongly localized around the nucleus. When such functions are used, the cusp in $V$ due to the presence of some atom "$n$" that resides far from the basis functions on atoms "$i$" and "$j$" will be reduced owing to the smallness of the basis product $\chi_{\alpha_i}\chi_{\beta_j}$ in regions of space near atom "$n$". As a direct consequence, the contribution of atom "$n$" to the integral $A$ will be small. Stated equivalently, the dominant contributions to the integrand $\chi_{\alpha_i}V\chi_{\beta_j}$ arise from atoms "$n$" that are in the vicinity of the atoms "$i$" and "$j$".

To quantify these ideas, we first focus on understanding how to partition the potential (not the total integrand as in previous methods) into a sum over atomic contributions:

$$V = \sum_n V_n = \sum_n V\omega_n \tag{13}$$

Such a partitioning could, for example, be accomplished with a one-center partitioning function of the type proposed by Delley.[6] As usual, the sum is over all atoms in the system and $V_n$ represents the contribution of atom "$n$" to the total potential $V$. To develop a notation capable of distinguishing between the various contributions to a matrix element $A = \int \chi_{\alpha_i} V \chi_{\beta_j}\, d\vec{x}$, first consider the set $M$ composed of all atoms in the system except atoms $i$ and $j$. In addition, we define two additional sets, which are subsets of $M$, that consist of all atoms within some arbitrary (but finite) radius $R_o$ of atom $i$ or $j$ as $N_i$ and $N_j$, respectively. The intersection of the sets $N_i$ and $N_j$ (i.e., $I = N_i \cap N_j$) is therefore the set of atoms that are within $R_o$ of both atoms $i$ and $j$. This collection of atoms defines what we will refer to as the "inner zone" of atoms. The remainder of the polyatomic system, the "outer zone", is therefore defined as the set $O = M \backslash I$. By use of these definitions, eq 13 becomes

$$\begin{aligned} V &= V_i + V_j + \sum_{k\in I} V_k + \sum_{l\in O} V_l \\ &= V_{ij} + \sum_{k\in I} V_k \end{aligned} \tag{14}$$

where

$$V_{ij} = V_i + V_j + \sum_{l\in O} V_l \tag{15}$$

without any loss of generality or approximation being made.

By substituting eqs 14 and 15 into our previous definition for the general matrix element $A$, we thus obtain

$$\begin{aligned} A &= \int \chi_{\alpha_i} V \chi_{\beta_j}\, d\vec{x} \\ &= \int \chi_{\alpha_i} V_{ij} \chi_{\beta_j}\, d\vec{x} + \sum_{k\in I} \int \chi_{\alpha_i} V_k \chi_{\beta_j}\, d\vec{x} \end{aligned} \tag{16}$$

which is an expression that merits some discussion. The second term in eq 16 consists of a sum over all atoms within the "inner zone". As such, the three center integrals contained in this sum can be expected to make significant contributions to the integrated value of $A$. The terms in this sum require three-center grids in order to efficiently integrate their contribution to high levels of accuracy. By comparison, the first term has now taken on the appearance of a two-center integral even though it is clear, by direct substitution of eq 15, that it contains three-center components. These three-center components, however, are found within the "outer zone" of atoms $i$ or $j$ and thus make smaller contributions to the total value of $A$. It would be customary to integrate such terms using a three-center integration grid. The inner zone, however, can always be chosen in such a way that these three-center integrals will be within 2 orders of magnitude of the target accuracy for the matrix element (e.g., if the target accuracy is $1 \times 10^{-05}$, then these integrals will be smaller than $1 \times 10^{-03}$). These terms are therefore small enough (in a relative sense) when compared to the absolute accuracy target that the needed two to three significant figures of absolute accuracy in these terms can be obtained using two-center ($ij$) integration grids. Thus, the three-center integrals contained in the first term of eq 16 can be evaluated using two-center integration grids. For large polyatomic or periodic systems the number of such terms can be substantial; reducing the number of explicit integration centers can lead to dramatic cost savings without sacrificing numerical accuracy if done properly.

To derive explicit formula for the nuclear weights, we now employ the cutoff profile function $S_{ij}$ of eq 8 to decompose the three-center integrals ($ijk$) in eq 16 into two two-center ($ik$ and $jk$) integrals and finally decompose all two-center integrals into one-center terms by the same procedure. By direct substitution we thus obtain

$$\begin{aligned} A &= \int \chi_{\alpha_i} V_{ij} \chi_{\beta_j} (S_{ij} + S_{ji})\, d\vec{x} + \sum_{k\in I} \int \chi_{\alpha_i} V_k \chi_{\beta_j} (S_{ij} + S_{ji})\, d\vec{x} \\ &= \int \chi_{\alpha_i} V_{ij} \chi_{\beta_j} (S_{ij} + S_{ji})\, d\vec{x} + \sum_{k\in I} \int \chi_{\alpha_i} V_k \chi_{\beta_j} [S_{ij}(S_{ik} + S_{ki}) + S_{ji}(S_{jk} + S_{kj})]\, d\vec{x} \\ &= \int \{\chi_{\alpha_i} V_{ij} \chi_{\beta_j} S_{ij} + \sum_{k\in I} \chi_{\alpha_i} V_k \chi_{\beta_j} S_{ij} S_{ik}\}\, d\vec{x} + \\ &\qquad \int \{\chi_{\alpha_i} V_{ij} \chi_{\beta_j} S_{ji} + \sum_{k\in I_{ij}} \chi_{\alpha_i} V_k \chi_{\beta_j} S_{ji} S_{jk}\}\, d\vec{x} + \\ &\qquad \sum_{k\in I} \int \chi_{\alpha_i} V_k \chi_{\beta_j} \{S_{ij} S_{ki} + S_{ji} S_{kj}\}\, d\vec{x} \end{aligned} \tag{17}$$

Note that all integrals in the above expression contain a single integration center. For example, the last integral in the final term is centered at site "$k$", since the cusp at center $i$ has been removed by the function $S_{ji}$ and the cusp at $j$ is eliminated by $S_{kj}$. By substituting eqs 15 and 13 into eq 17 (and performing some algebra), we obtain the working expression for our integration scheme,

$$\int \chi_{\alpha_i} V \chi_{\beta_j}\, d\vec{x} = \int \chi_{\alpha_i} V \chi_{\beta} g_{ij}^i\, d\vec{x}_i + \int \chi_{\alpha_i} V \chi_{\beta} g_{ij}^j\, d\vec{x}_j + \sum_{k\in I} \int \chi_{\alpha_i} V \chi_{\beta} g_{ij}^k\, d\vec{x}_k \tag{18}$$

where

$$g_{ij}^i = S_{ij}\{1 + \sum_{k\in I} (S_{ik} - 1)\omega_k\} \tag{19}$$

and

$$g_{ij}^{k} = \omega_k\{S_{ij}S_{ki} + S_{ji}S_{kj}\} \tag{20}$$

The analogous quantity, $g_{ij}^{j}$, has the same form as that of $g_{ij}^{i}$ with the exception that the $i$ and $j$ indices are interchanged. We note that the final form represented by eq 18 is easy to use, since it involves only the multiplication of the original integrand by a nuclear weight function. When attempting to integrate the charge density (or when center $i = j$), eq 18 simplifies to

$$\int \chi_{\alpha_i} V \chi_{\beta i}\, \mathrm{d}\vec{x} = \int \chi_{\alpha_i} V \chi_{\beta i} G_{ii}^{i}\, \mathrm{d}\vec{x}_i + \sum_{k\in I} \int \chi_{\alpha_i} V \chi_{\beta i} G_{ii}^{k}\, \mathrm{d}\vec{x}_k \tag{21}$$

where

$$G_{ii}^{i} = 1 + \sum_{k\in I}(S_{ik} - 1)\omega_k \tag{22}$$

and

$$G_{ii}^{k} = \omega_k S_{ki} \tag{23}$$

In summary, we decompose the multicenter integral $\int \chi_{\alpha_i} V \chi_{\beta_j}\, \mathrm{d}\vec{x}$ into a sum over one-center contributions. Unlike previous methods, however, we do not place one-center grids on all atoms contained in the system and subsequently sum over each one-center contribution to obtain an estimate of the total integral. Instead we treat the total integral as if it were composed of two "zones". For an inner zone, where contributions to the total integral are large, all multicenter features must be retained and one-center integration must be done very accurately. The remaining portion of the material, composed of atoms not explicitly included in the inner zone, is composed of terms that are smaller in absolute magnitude and can therefore be computed with less relative accuracy using sufficiently large inner zone one-center grids. One of the motivations for adopting this approach is that it is computationally more efficient to chose slightly enlarged one-center grids in the inner zone (to sufficiently integrate the outer zone) than to increase the number of atomic centers where one-center integration grids will be explicitly constructed.

An added benefit associated with this approach is that the cost for each multicenter integral will not increase as the size of the system is increased. The inner zone can be defined independently of the system size assuming the system is not so small that all atoms are contained within this zone. For very large molecules or crystalline systems, the number of nonnegligible matrix elements increases linearly with the number of atoms in the systems. In such cases, the numerical integration costs in our scheme will also increase linearly. Of course, any scheme that exploits the local nature of the basis set can, in principle, achieve such scaling. The challenge is to design an approach that is robust and efficient that does not suffer from the use of excessively arbitrary cutoffs. In our scheme, the boundary between the inner and outer zones, of course, is not unique. It can, however, be selected on the basis of the accuracy requirements of the calculation and the integrands being evaluated and through a well-defined process. In the following sections we will demonstrate numerically that, in practice, a general definition of the inner zone size can be adopted that achieves the absolute accuracy objectives while exhibiting high overall efficiency.

## III. One-Center Integration and Automatic Grid Generator

Recently, there has been considerable emphasis on the development and comparison of one-center integration strategies in the literature.[8−15] Such studies typically adopt a multicenter decomposition scheme (frequently due to Becke[3]) and concentrate on understanding how various improvements to the single-center integration processes can be made. Here, we will describe our adaptive one-center integration scheme, discuss some characteristics associated with our automatic grid generator, and finally compare our results to those obtained by other researchers. To facilitate the comparison and to distinguish between the contributions made by our one-center integration scheme from those made by our multicenter decomposition strategy, we will use Becke's[3] nuclear weights in all results of this section unless otherwise stated.

**IIIa. One-Center Integration**. Once a multicenter decomposition scheme has been chosen, there are many ways to perform the necessary one-center integrations. In fact, there are several schemes[3,6−15] that all result in slightly different grids in use today. In this section we will focus on describing our approach. As will be seen below, an adaptive radial grid that can account for the shell structure of the center of integration is important.

*Radial Grids.* For its simplicity in application a unique analytical expression capable of determining the radial grid is desirable. Simplicity, however, is not necessarily the most important criterion in determining an effective quadrature grid. Optimal performance is typically achieved by tuning the radial grid to the cusp and shell structures of the atomic densities. This perspective drives us to empirically divide the radial integration grid into segments:

$$I = \int \mathrm{d}\Omega \int_0^{\infty} F(\vec{r}) r^2\, \mathrm{d}r \approx \int \mathrm{d}\Omega \int_0^{R_{\max}} F(\vec{r}) r^2\, \mathrm{d}r$$
$$= \sum_{i=1}^{n} \int \mathrm{d}\Omega \int_{R_i}^{R_{i+1}} F(\vec{r}) r^2\, \mathrm{d}r \tag{24}$$

For small $r$, we employ the geometric series $R_i$ (au) = (0, 1/(2$z$), 3/(2$z$), 9/(2$z$), ...), where $z$ is the nuclear charge of the atom of interest. This process is continued for values of $R_i < {}^3/_4$. For ${}^3/_4 < R_i < 1$, a value of $R_i = 1$ au is used followed by the sequence $R_n$ = (2, 3, 5, 7, 10, 15, $R_{\max}$) until the maximum value of $R$ has been reached. For all results reported here a value of $R_{\max} = 30$ au has been used, which is more than sufficient for most applications although $R_{\max}$ can be trivially enlarged to suit special needs when necessary. Gauss−Legendre quadrature is used within all radial subregions. Gill et al.[11] have also proposed a scheme that divides the radial grid into a few regions, although the divisions are not as extensive as what is being proposed here. The advantage of dividing the radial grid into several segments is that it provides the flexibility necessary to adapt to the changes in the integrand through the cusp and valence regions. With the type of grid just proposed, the addition of grid points to one radial segment is completely decoupled from activities in another segment. This allows grid points to be placed were they are most needed and provides a mechanism to quickly converge integrated quantities to high levels of accuracy.

*Angular Grids.* As described above, the radial grid is divided into subregions to improve efficiency. This strategy also allows the use different angular grids in each individual subregion and provides the needed flexibility to adapt to angular variations in the integrand. The angular grids are constructed from special point formulas that exactly integrate spherical harmonics over the surface of the unit sphere up to a maximum, $l_{\max}$.[20−24] The highest order special point formula implemented is $l = 59$, which generates a grid that contains 1202 integration points.[24] For

situations that require higher accuracy a product scheme that employs Gauss−Legendre quadrature for $\theta$ and a geometric trapezoidal quadrature for $\varphi$ are used. Regardless of the accuracy criterion, both $\theta$ and $\varphi$ are integrated to the same $l_{max}$. Although some efficiency is sacrificed by not allowing the maximum angular momentum to vary independently for $\theta$ and $\varphi$, we feel the numerical stability gained with respect to rotational transformations justifies integrating them to the same order. The smallest angular grid used in our approach for any radial subregion is capable of integrating a polynomial exactly to $l = 3$. No attempt is made by the numerical integration algorithms to reorient the polyatomic system from that input by a user in the geometry definition, though such reorientation might be beneficial.[25−27]

**IIIb. Automatic One-Center Grid Generator**. For ease of use and reliability an automatic grid generation algorithm that is capable of selecting the appropriate number and location of grid points necessary to integrate the integrand of interest to the target accuracy is highly desired. The basic procedure involved in our automatic grid generation algorithm is as follow.

1. A set of prototype integrands (e.g., $\rho_i V_{eff}$, $\rho_i V_{xc}$, $\rho$) that behave similarly to the actual integrands of interest are evaluated on a set of test grids. These prototype integrands share the attribute that they can be expediently evaluated on the test grids using a density that has been formed from the superposition of atomic densities. We also find that these test functions generate grids that reliably integrate actual matrix elements in our electronic structure program, GAPSS.[28,29]

2. Each radial subregion is searched in sequence to determine the number of radial and angular points needed to guarantee that the test integrands can be integrated to 1 order of magnitude greater accuracy than the target accuracy for the real integral. Within a radial subregion this is accomplished by first evaluating the test integrands on an unnecessarily large radial grid in combination with a moderately sized angular grid. The number of radial grid points ($N$) is then successively increased from small to large $N$ until the difference between the current value and that found for the very large radial grid is smaller than the absolute tolerance. When the smallest grid satisfying this criterion is determined, the procedure is halted and an analogous process is initiated for the angular grids combined with the newly determined radial grid. This procedure is repeated for all radial subregions for all atoms in the system. Absolute accuracy (as opposed to a relative accuracy) is always chosen as the basis for the acceptance of a given grid in our approach.

When the acceptance criterion for each subregion is set to be an order of magnitude more accurate than the target integral, the probability that the overall accuracy target will be met is greatly increased. The statistical issues involved can be understood by noting that the number of potential sources of error involved in selecting the grid is on the order of $2 \times (10-15) \times N_{center}$ (one each for the radial and angular components multiplied by the typical number of radial segments for a given atomic center times the total number of centers in the polyatomic system). If we assume that each error source contributes a random error greater or less than some value $\epsilon$, then the total statistical error would be of the order $[(20-30)N_{center}]^{1/2}$ larger than $\epsilon$. Thus, for a wide range of atomic centers, $N_{center}$, a value of $\epsilon$ that is approximately $^1/_{10}$ of the target accuracy would yield a final accuracy that is of the same order of magnitude as that targeted by the user. The signs of these errors can also be random. Therefore, some systems can experience fortuitous cancellation of errors where others may not be as fortunate. If, for example, all errors possessed the same sign and absolute magnitude, the total error could be as large as $(2-3) \times N_{center} \times \epsilon$. In practice, however, we find that the final errors do not deviate significantly from the "rule-of-thumb" given above.

**IIIc. Integration Efficiency of the One-Center Grid**. In a recent work, Perez-Jorda, Becke, and San-Fabian (PBS)[11] presented an adaptive one-center integration scheme and compared the performance of that scheme to several previously published methods. The results of that comparison indicated that the PBS strategy provided the most efficient approach for the set of molecules and integrands tested. In the following paragraphs we will present the results of our one-center integration scheme and compare the results to those obtained by PBS.[30] To facilitate the comparison, we have evaluated the same test integrand with the five accuracy targets ($10^{-3}-10^{-7}$ au) for the 21-molecule test suite of PBS at the molecular geometries reported in that work. The test integrand is the total molecular density that is synthesized from a superposition of atomic densities. The nuclear weight function proposed by Becke[3] has been used to decompose the multicenter integrals into sums over one-center integrals. Thus, any differences between the results of this study and that of PBS are due entirely to the two different one-center integration schemes being used.

The number of grid points required to integrate the total density for five absolute accuracy targets, $\epsilon$, and the resulting absolute errors for the 21-molecule data set are shown in Table 1. When the relative errors presented in Tables 1 and 2 of PBS are multiplied by the total number of electrons in the molecule and thus converted to an absolute error, a direct comparison of the results of this study and that of PBS can be made. By such a comparison, it can then be seen that our one-center integration scheme is generally more efficient, requiring fewer grid points than that of PBS for 20 out of 21 molecules. The only exception is for $C_{60}$, for which PBS's results are superior. The underlying reason for this is unclear to us. Possibly there is a significant advantage associated with decoupling the $\theta$ and $\varphi$ grids in C60 or there may be some unexpected benefit in choosing a special orientation as described in ref 11 because of the high symmetry of the molecule. Overall, however, our one-center grid is more efficient than that reported by PBS.

With the intent of providing an intuitive and unbiased method of comparing the integration efficiency of various one-center integration schemes, PBS suggested that the function $\chi = -\log_{10}(N_{grid}/N_{atoms})/\log_{10} \epsilon_r$ (where $N_{grid}$ is the number of grid points required to achieve a given relative error $\epsilon_r$) be used as a measure of efficiency. The properties of the function are such that the smaller the value of $\chi$ the greater the efficiency of the method (note that $\chi = 0$ or $10^{\chi} = 1$ represents infinitely high efficiency). Using this function and the data provided in Table 1 (converted to relative as opposed to absolute errors), we find that our grid is characterized by $10^{\chi} = 4.190$ compared to PBS's reported value of 4.484.

To further test the relationship between the efficiency of the one-center integration scheme and the multicenter decomposition strategy, we have performed tests using our one-center grids with the weight function

$$\omega_i(\rho_i/r_i^2) = \frac{\rho_i/r_i^2}{\sum_j (\rho_j/r_j^2)}$$

proposed by Delley.[6] For each accuracy target ($10^{-3}-10^{-7}$) we have again computed the average number of grid points needed to achieve a given target for all 21 molecules in the PBS test suite. Curves B and C of Figure 1a illustrates how the

**TABLE 1: Number of Grid Points of Our One-Center Grid Generated by Superposition of Atomic Densities (Upper Rows) and Absolute Errors (Lower Rows) for 21 Molecules at Five Absolute Tolerances $\epsilon$[a]**

| tolerance | E-3 | E-4 | E-5 | E-6 | E-7 |
|---|---|---|---|---|---|
| $H_2O$ | 2939 | 7487 | 11572 | 22611 | 43156 |
| | 7.0e-5 | 1.5e-5 | 5.1e-6 | 3.0e-7 | 3.3e-8 |
| $NH_3$ | 4252 | 16353 | 18799 | 45716 | 79321 |
| | 9.7e-4 | 4.4e-5 | 2.5e-6 | 4.9e-6 | 4.6e-8 |
| $CH_4$ | 4400 | 13013 | 28794 | 63949 | 99817 |
| | 2.7e-4 | 1.9e-5 | 5.7e-5 | 9.4e-8 | 5.2e-8 |
| $C_2H_6$ ($D_{3h}$) | 10172 | 30386 | 93142 | 190414 | 416628 |
| | 1.6e-3 | 5.5e-5 | 7.4e-6 | 4.3e-10 | 4.3e-8 |
| $C_2H_5OH$ | 14664 | 46243 | 90147 | 166425 | 329492 |
| | 1.4e-3 | 4.6e-5 | 6.3e-6 | 7.9e-7 | 6.6e-8 |
| ClOH | 3144 | 7056 | 12815 | 34432 | 79910 |
| | 2.2e-4 | 3.e-5 | 2.6e-6 | 7.7e-7 | 5.4e-8 |
| $BH_3$ | 3741 | 7787 | 15292 | 31669 | 63233 |
| | 9.8e-4 | 2.8e-4 | 1.5e-4 | 4.4e-6 | 4.0e-8 |
| CClFH−CClFH ($C_i$) | 7624 | 19910 | 40380 | 97694 | 169842 |
| | 1.3e-3 | 9.2e-5 | 1.7e-6 | 5.8e-7 | 7.3e-8 |
| $C_2H_6$ ($D_{3d}$) | 8804 | 27390 | 44044 | 127480 | 189104 |
| | 1.3e-3 | 2.1e-4 | 1.7e-4 | 2.7e-7 | 8.0e-7 |
| CFH−CFH | 5330 | 15560 | 38048 | 77776 | 151564 |
| | 5.6e-3 | 3.9e-5 | 1.1e-5 | 4.7e-8 | 1.6e-8 |
| $CCl_2-CH_2$ ($C_{2v}$) | 7624 | 19910 | 40380 | 97694 | 169842 |
| | 1.3e-3 | 9.2e-5 | 1.7e-6 | 5.8e-7 | 7.3e-8 |
| $C_6H_6$ | 18414 | 65124 | 159714 | 319024 | 489258 |
| | 7.5e-4 | 7.7e-5 | 2.0e-5 | 1.6e-6 | 7.3e-8 |
| $SF_6$ | 10953 | 26569 | 56969 | 91380 | 162287 |
| | 7.3e-4 | 3.9e-5 | 2.1e-6 | 5.9e-7 | 4.2e-8 |
| $C_5H_5N_5$ | 33985 | 71900 | 303730 | 629775 | 1204005 |
| | 6.7e-4 | 6.0e-4 | 7.3e-5 | 5.5e-7 | 4.0e-6 |
| $ZnCl_3(C_5H_6N_5)$ | 39667 | 169464 | 421827 | 846220 | 1808741 |
| | 7.2e-4 | 2.8e-4 | 5.5e-5 | 5.2e-6 | 2.5e-6 |
| $C_5H_6N_2O_2-C_6H_4O_2$ | 75255 | 460600 | 970533 | 1923858 | 3856271 |
| | 3.3e-3 | 2.2e-4 | 2.1e-5 | 1.1e-6 | 8.4e-9 |
| $C_{12}H_{14}N_2O_2$ | 53406 | 244760 | 774644 | 1454808 | 3424794 |
| | 2.9e-3 | 3.8e-4 | 1.6e-4 | 2.9e-5 | 2.3e-7 |
| $C_{15}H_{12}N_2O_2$ | 102969 | 294559 | 752082 | 1770858 | 3155695 |
| | 2.5e-4 | 6.1e-5 | 1.1e-4 | 8.2e-6 | 8.6e-6 |
| $C_{17}H_{21}NO_4$ | 79275 | 354970 | 992836 | 2095514 | 4037950 |
| | 4.5e-3 | 5.9e-4 | 2.8e-5 | 3.4e-6 | 1.7e-7 |
| $C_{20}H_{29}N_3O_2$ | 133701 | 482734 | 1268229 | 2883650 | 4995880 |
| | 4.9e-4 | 2.1e-3 | 3.6e-4 | 8.6e-6 | 7.9e-7 |
| $C_{60}$ | 310860 | 779220 | 2931420 | 6200340 | 11802468 |
| | 1.3e-3 | 1.3e-4 | 1.1e-4 | 1.2e-5 | 2.0e-6 |

[a] No advantage has been taken of molecular symmetry. The multicenter decomposition scheme described in ref 3 has been used.

**TABLE 2: Number of Grid Points Required to Integrate the Total Charge Density Using for Five Target Accuracies; Absolute Errors Are Also Illustrated[a]**

| tolerance | E-3 | E-4 | E-5 | E-6 | E-7 |
|---|---|---|---|---|---|
| $H_2O$ | 1065 | 2695 | 4948 | 7825 | 11436 |
| | 3.1e-4 | 1.8e-5 | 9.0e-7 | 2.8e-7 | 1.9e-8 |
| $NH_3$ | 1441 | 3439 | 7079 | 12205 | 16663 |
| | 4.9e-4 | 3.8e-5 | 1.1e-6 | 4.9e-8 | 6.6e-8 |
| $CH_4$ | 1683 | 3982 | 6538 | 14163 | 20158 |
| | 6.8e-4 | 4.7e-5 | 1.8e-7 | 4.4e-7 | 3.9e-8 |
| $C_2H_6$ ($D_{3h}$) | 3858 | 6262 | 14367 | 22200 | 34125 |
| | 3.3e-4 | 1.1e-4 | 3.5e-7 | 1.7e-7 | 1.5e-7 |
| $C_2H_5OH$ | 4484 | 8415 | 18185 | 30394 | 67572 |
| | 3.2e-4 | 3.6e-5 | 6.3e-6 | 3.8e-7 | 1.9e-7 |
| ClOH | 1942 | 3498 | 6942 | 10897 | 19082 |
| | 6.8e-4 | 2.9e-5 | 1.3e-7 | 4.7e-7 | 1.2e-8 |
| $BH_3$ | 1636 | 3543 | 6301 | 11126 | 18583 |
| | 1.3e-4 | 3.6e-5 | 4.0e-6 | 1.6e-7 | 6.3e-8 |
| CClFH−CClFH($C_i$) | 5002 | 9657 | 17452 | 42250 | 81220 |
| | 8.1e-4 | 1.3e-4 | 1.1e-6 | 3.3e-7 | 1.3e-7 |
| $C_2H_6$ ($D_{3d}$) | 4074 | 7168 | 14190 | 21643 | 47916 |
| | 1.9e-4 | 1.3e-4 | 8.1e-6 | 3.5e-7 | 1.3e-7 |
| CFH−CFH | 2932 | 6670 | 13555 | 31212 | 61815 |
| | 3.4e-4 | 2.5e-5 | 7.5e-6 | 3.8e-8 | 4.4e-8 |
| $CCl_2-CH_2$ ($C_{2v}$) | 2989 | 7411 | 12972 | 31934 | 69892 |
| | 5.1e-4 | 5.4e-5 | 4.5e-6 | 3.5e-7 | 9.3e-8 |
| $C_6H_6$ | 5936 | 12121 | 22349 | 37738 | 72509 |
| | 4.3e-5 | 1.8e-4 | 2.3e-5 | 5.1e-7 | 1.2e-7 |
| $SF_6$ | 4078 | 7240 | 12475 | 27937 | 44397 |
| | 1.9e-3 | 1.3e-4 | 8.5e-7 | 1.6e-6 | 8.7e-8 |
| $C_5H_5N_5$ | 6925 | 17314 | 28434 | 54911 | 97331 |
| | 2.2e-3 | 2.5e-4 | 7.7e-7 | 4.4e-7 | 9.1e-8 |
| $ZnCl_3(C_5H_6N_5)$ | 6225 | 11834 | 26909 | 56319 | 127191 |
| | 3.0e-3 | 1.0e-4 | 2.2e-5 | 1.4e-6 | 1.6e-7 |
| $C_5H_6N_2O_2-C_6H_4O_2$ | 11507 | 24135 | 39541 | 72351 | 157869 |
| | 5.3e-3 | 3.9e-4 | 4.1e-5 | 7.8e-7 | 5.1e-8 |
| $C_{12}H_{14}N_2O_2$ | 6041 | 15139 | 31249 | 67608 | 126171 |
| | 1.7e-3 | 2.2e-3 | 1.3e-4 | 7.0e-5 | 1.1e-6 |
| $C_{15}H_{12}N_2O_2$ | 12584 | 26387 | 44360 | 79836 | 180675 |
| | 1.3e-2 | 1.5e-3 | 1.0e-4 | 2.2e-5 | 4.0e-7 |
| $C_{17}H_{21}NO_4$ | 9573 | 21008 | 38657 | 77423 | 165136 |
| | 4.7e-3 | 4.2e-4 | 6.5e-5 | 1.2e-8 | 4.9e-7 |
| $C_{20}H_{29}N_3O_2$ | 23277 | 49647 | 75880 | 162002 | 272667 |
| | 3.3e-3 | 5.5e-4 | 5.8e-5 | 3.2e-6 | 8.6e-7 |
| $C_{60}$ | 28098 | 79296 | 133898 | 199916 | 332337 |
| | 2.7e-3 | 5.6e-4 | 2.2e-5 | 5.7e-7 | 1.8e-9 |
| $C_{240}$ | 100775 | 296244 | 473956 | 648840 | 1093610 |
| | 3.5e-2 | 4.5e-3 | 3.0e-4 | 9.8e-6 | 6.0e-8 |

[a] No advantage has been taken of molecular symmetry.

average number of grid points varies as a function of accuracy target using our one-center grids with Becke's[3] and Delley's[6] weights. For comparison, the figure also includes the published results of PBS (curve A) using Becke's[3] weight functions and the PBS one-center integration scheme. For each of the five target accuracies it can be seen that our one-center grid combined with either Becke's or Delley's weights is more efficient than the PBS result. Perhaps more interestingly, the integration efficiency index of Delley's weight function and our one-center grid is 3.806, which is substantially more efficient than that of Becke's weight with our grids. Figure 1b demonstrates how the number of grid points increases as a function of the number of atomic centers in the molecule for an absolute accuracy target of $1.0 \times 10^{-05}$. These data clearly show the anomaly mentioned above for $C_{60}$ and the fact that Delley's weight with our one-center grid possesses the best scaling as a function of system size.

Parts a and b of Figure 1 also demonstrate the pronounced effect that the multicenter decomposition scheme can have on the overall efficiency of a given method. For example, Becke's weight with our one-center grid yields an efficiency index of 4.190 while the same weights with PBS grids produces an index of 4.484. This level of difference represents a modest improvement. However, our one-center integration scheme in conjunction with Becke's and Delley's weights yields a PBS efficiency index of 4.190 and 3.806, respectively. These results suggest that when the multicenter decomposition scheme is properly selected, significant improvements in overall efficiency can be achieved. In the following sections we will show that by use of our new decomposition scheme, this is indeed the case.

Finally, we propose a slightly different merit function than that suggested by PBS. We propose a function that is more indicative of how the number of grid points increases as the absolute accuracy target is decreased. Thus, we suggest a function that is characterized by the prefactor $C$ and an order of convergence parameter $n$,

$$\lambda = C\left(\frac{N_{\text{grid}}}{N_{\text{atom}}}\right)^{-n} \tag{25}$$

where $\lambda$ is the absolute error resulting from the grid. A small prefactor means a method is efficient in achieving low accuracies, while a method with a large order of convergence can efficiently achieve high-accuracy targets. By a least-squares fit

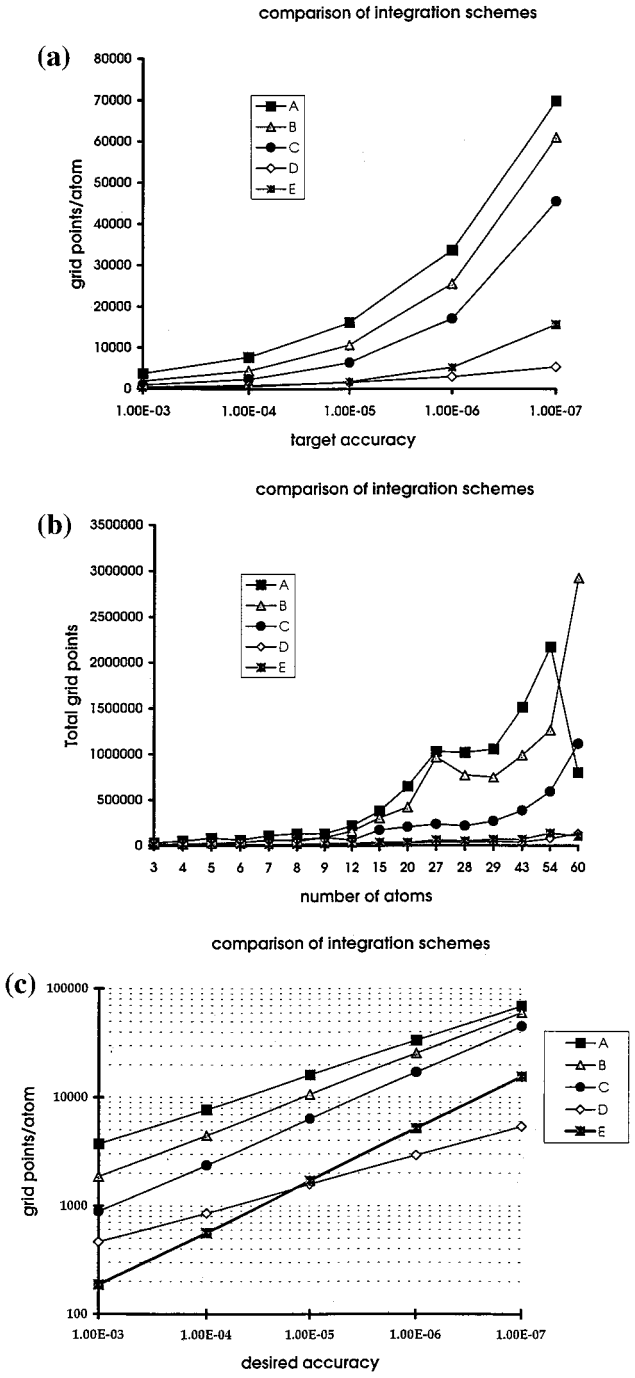


**Figure 1.** Comparison of several integration schemes' ability to integrate the total molecular charge. Curve A [squares] illustrates the result of using Becke's weighting scheme in combination with PBS's one-center (ref 11). Similarly, curve B [triangle] results from Becke's weight combined with our 1c mesh (section III). Curve C [circle] is due to a Delley weight and our 1c grid, while curve D [diamonds] demonstrates our multicenter integration scheme combined with our 1c grid. Curve E uses the same method as that for curve D but for integrating matrix element type integrals as described in section IV. (a) Average number of 1c grid points required to achieve the target accuracy. (b) Average number of 1c grid points as a function of the number of atoms in the systems for a target integration accuracy of $10^{-5}$. (c) Average number of grid points as a function target accuracy derived from the fitted prefactors and the rates of convergence expression found in section III.

to the data in Figure 1a, it is found that our grid with Becke's weight gives $C = 10^{5.6}$ and $n = 2.6$ while Delly's weight ($\omega_i$-($\rho_i/r_i^2$)) and our grid yield $C = 10^{3.9}$ and $n = 2.3$. Therefore, we might expect that Becke's weights would more readily converge to higher accuracy targets while Delley's weights are more efficient for typical accuracy requirements because of the much smaller prefactor. These conclusions are supported by log−log plots shown in Figure 1c.

## IV. Results for the New Multicenter Integration Scheme

This section will present results obtained using our one-center integration scheme in conjunction with our multicenter decomposition strategy outlined in section II. We will demonstrate the behavior of this approach by integrating the total charge densities of the 21 molecules in the PBS test suite for five different accuracy targets of $10^{-3}-10^{-7}$. In addition, we will discuss the application of our approach to quantities more relevant to the construction of the matrix representation of the Hamilitonian. It will be shown that our method is more efficient for all systems tested.

**IVa. Integration of Total Charge Densities**. The use of eq 21 to integrate the total charge density of a polyatomic system requires that one atom must be designated as atom "*i*". This action has the effect of collapsing any explicit or implied sums over this index. Thus, by selecting a specific "*i*" atom and replacing *V* with $\rho$, eq 21 reduces to

$$\int \rho \, d\vec{x} = \int \rho G_{ii}^{i} \, d\vec{x} + \sum_{k \neq i} \int \rho G_{ii}^{k} \, d\vec{x} \qquad (26)$$

where $G_{ii}^{i}$ and $G_{ii}^{k}$ are given by eqs 22 and 23 with the understanding that the sum over *k* extends to all atoms in the system except atom "*i*". From eq 26 it is clear, however, that atoms "*i*" and "*k*" are treated differently. Although the total charge obtained is independent of which atom is chosen as atom "*i*", the number of grid points per atom is dependent on this selection. By convention we select the atom with the largest nuclear charge as the "*i*" atom. If more than one such atom exists, the atom that is closest to the geometric center of the molecule is chosen as the "*i*" center.

The result of using our one-center grid and decomposition scheme to obtain the total charge density for the PBS test suite and C240 is contained in Table 2. A comparison of the data in Tables 1 and 2 quickly reveal that a significant reduction in the number of grid points for all 21 molecules for all accuracy targets has been achieved using the new decomposition scheme. Curve D of parts a and b of Figure 1 further illustrate this point by demonstrating how the number of grid points varies as a function of accuracy target and number of atoms, respectively.

We can further demonstrate the importance of the multicenter decomposition scheme on the integration efficiency by investigation how the average grid saving factor

$$S_a = \frac{1}{21} \sum_{m=1}^{21} \frac{N_{BLH}{}^{m}(\epsilon_i)}{N_{LH}{}^{m}(\epsilon_i)} \qquad (27)$$

varies as a function of the accuracy target $\epsilon$. In this expression, $N_{BLH}{}^{m}(\epsilon)$ represents the number of grid points needed using Becke's weights and our one-center grid to achieve a given accuracy target $\epsilon$ (the analogous quantity indexed by LH corresponds to the use of our weights and grids). Note that the sum is over all 21 molecules in the test suite. The denominator of the expression represents the number of grid points needed using our decomposition scheme and one-center grid, and *m* is an index that represents the specific molecules in the test suite. Figure 2a plots the behavior of the computed ratios as a function of the five target accuracies. From this figure it can be seen

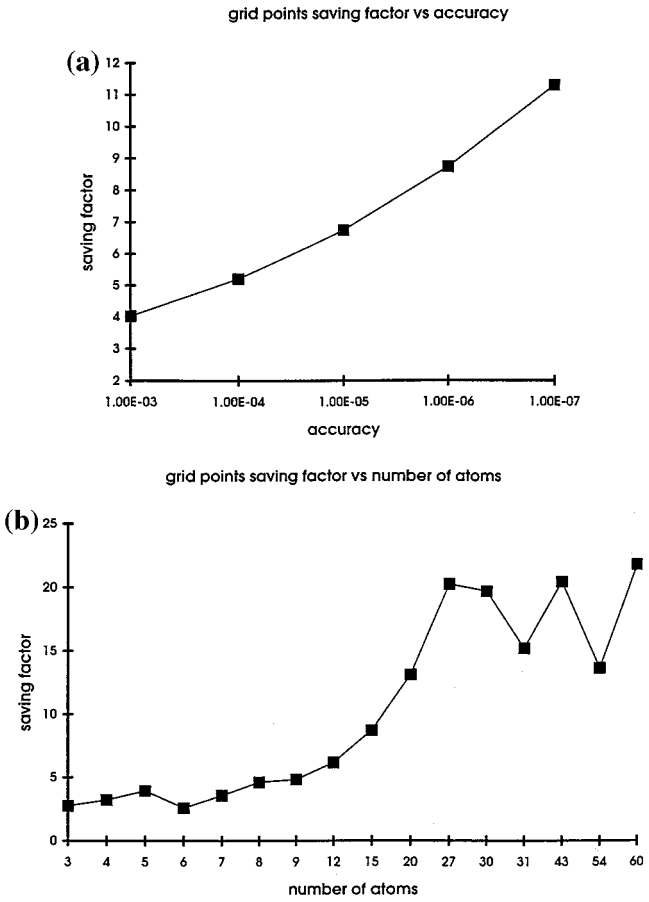


**Figure 2.** Improvement by the present multicenter decomposition scheme over that of ref 3 to integrate the total charge. The same one-center 1c grid is used in both calculations. The graph illustrates effects due solely to multicenter decomposition scheme. (a) Ratio of the number of grid points needed vs target accuracy. (b) Ratio of the number of grid points needed vs number of atoms in the test molecule.

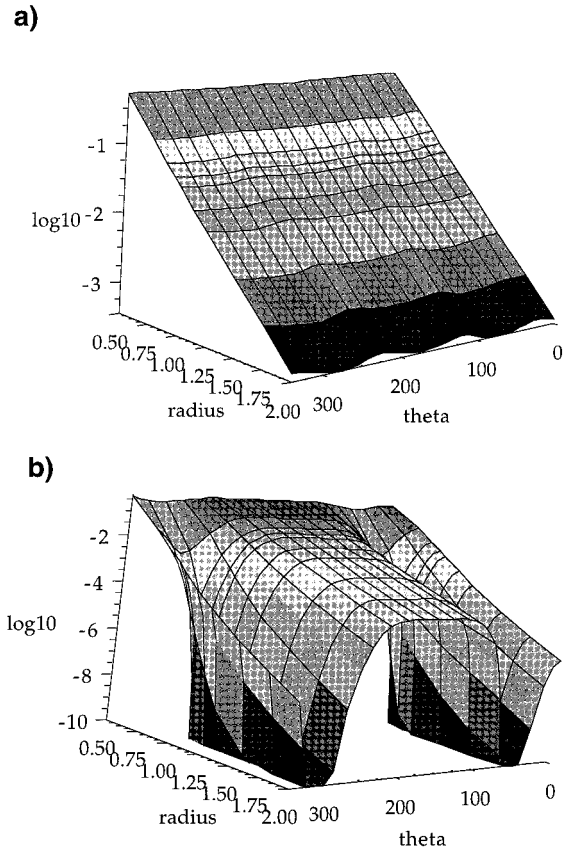


**Figure 3.** Comparison of one-center integrands (total charge density times one-center weights) for $H_2O$ using our method and that of ref 3. The coordinate system is as follows. The integration origin is located at the oxygen center, the radius is the distance between a point P and the O nucleus in units of OH bond length. All points P are contained within the molecular plan. $\theta$ is the angle between the $C_{2v}$ axis and that of OP in degrees. Part a is a semilogrithmic plot of the integrand variation using the approach described in this work, whereas part b demonstrates the behavior using the approach described in ref 3. Only values larger than $1 \times 10^{-10}$ are shown in part b.

that our decomposition scheme and one-center grid are 4−11 times more efficient than that reported by PBS. It is also important to note that this "saving factor" increases as the absolute error in the calculation decreases.

By performing a least-squares fit of the data in Table 2 to eq 25, we find that our method is characterized by a prefactor $C = 10^{7.0}$ and a convergence order $n = 3.8$. Note that although our prefactor is larger than Becke's ($C = 10^{5.6}$) and Delley's ($C = 10^{3.9}$), our order of convergence is higher than Becke's ($n = 2.6$) and Delley's ($n = 2.3$). Therefore, our approach both is more efficient at practical accuracy levels and is significantly more efficient when high-accuracy calculations are required. This can clearly be seen by comparing curves B−D in Figure 1c, which use our one-center grid with three different multicenter decomposition schemes.

By inspection of the data in Table 2, it is also clear that as the number of atoms in a molecule increases, so does the percentage reduction in the number of grid points needed to achieve the accuracy targets using our approach. Simply stated, the "larger" the polyatomic system the more efficient our approach becomes compared to previously publish schemes. This observation can be quantified by forming the saving factor,

$$S_m = \frac{1}{5}\sum_{i=1}^{5} \frac{N_{\mathrm{BLH}}{}^{m}(\epsilon_i)}{N_{\mathrm{LH}}{}^{m}(\epsilon_i)} \tag{28}$$

where the variables are defined as in eq 27. Note that in this expression the averages are being formed over the five target accuracies (as opposed to the 21 molecules in eq 27). A plot of the resulting ratios, as a function of the number of atoms in the system, is shown in Figure 2b. From these data it can be seen that the "saving factor" ranges from 3 to 20 and increases as the number of atoms in the system increases.

Although it is difficult to ascribe the increases in efficiency due to the new decomposition scheme to any particular aspect of the theory, a qualitative explanation can be obtained by investigating the spatial variation of integrand weight product. Figure 3 shows the log of the variation in the one-center integrand resulting from Becke's (Figure 3a) and our (Figure 3b) decomposition schemes for the total density times the partition weight for the oxygen atom in a water molecule. With respect to the radial direction, it can be seen that either decomposition scheme exhibits a smooth monotonic decay and is therefore easily integrated. In fact, the radial integration of Becke's integrand may be slightly easier than ours, since the function decays more rapidly to negligible values. Along the angular direction, however, Becke's integrand exhibits pronounced oscillations. Since our integrand is smoother, it is not surprising to expect that our approach would require fewer grid points to achieve a particular accuracy. Similar results can be obtained for other molecules in the test suite. The oscillations in either decomposition scheme are introduced by the use of sharply varying cutoff functions (see eq 12).[3,13] Becke's scheme,

however, uses a product formula, which has a tendency to create nodes in the angular variation of the integrand to combine these functions, whereas our approach uses a summation formula to obtain the nuclear weights. In systems that contain many nuclei the product form appears to enhance the oscillatory behavior of the integrand in the angular directions. To integrate the resulting function accurately can therefore require large angular grids. For this reason our approach becomes progressively more efficient as the number of atoms in the system is increased.

**IVb. Tests Relevant to Fock Matrix Construction**. In this subsection we concentrate on testing our approach with integrands that more accurately reflect the challenges typical of electronic structure calculations than those presented by the total charge density. Specifically, we wish to apply the method to the evaluation of matrix elements of the form $\langle\chi_{\alpha_i}|V|\chi_{\beta_j}\rangle$ more relevant to the construction of the Hamiltonian. To avoid an intractable investigation of the complicated dependence of $\langle\chi_{\alpha_i}|V|\chi_{\beta_j}\rangle$ on the basis set (and the possible loss of generality of our results), we chose instead to test the integration of the quantity $\rho_i V$, where $\rho_i$ is the atomic density of atom "*i*". As a consequence, the required formulas are given by eqs 21−23 with the exception that the product $\chi_{\alpha_i}\chi_{\beta i}$ should be replaced by $\rho_i$, i.e.,

$$\int \rho_i V \, d\vec{x} = \int \rho_i V G^i_{ii} \, d\vec{x}_i + \sum_{k\in I} \int \rho_i V G^k_{ii} \, d\vec{x}_k \qquad (29)$$

With respect to the potential *V*, we will report grids capable of simultaneously integrating both the LDA effective potential, $V_{\text{eff}}^{\text{LDA}}$, and the exchange−correlation potential only, $V_{\text{xc}}^{\text{LDA}}$. These potentials have been selected because of their common use in density functional theory. For these tests (and in our electronic structure program GAPSS[28,29]) we use the specific partition function $\omega_k = V_{\text{H}}^{k_{\text{atom}}}/\sum_{i_{\text{atom(allatoms)}}} V_{\text{H}}^{i_{\text{atom}}}$ in eq 13 where $V_{\text{H}}$ is the atomic Hartree potential obtained from an atomic Hartree−Fock calculation.

The number of grid points needed to integrate the integrands just described for the five target accuracies for the entire PBS test suite is found in Table 3. If a polyatomic system contains multiple atomic centers with the same chemical identity and those atoms are not symmetry-related, then the largest grid determined by the grid generator for that set of atoms will be assigned to all atoms of that chemical type. Since no analytical result exists for these test cases, the errors reported in Table 3 have been obtained by comparison to results derived with grids whose absolute accuracy target was $10^{-8}$ au. From a least-squares fit to the data using eq 25 we find that $C = 10^{1.7}$ and the order of convergence *n* is 2.1. By comparison of curve E in parts a−c of Figure 1 and other curves in the same figure, its is clear that the method being proposed is highly competitive.

**IVc. Testing the Inner and Outer Zone Partitioning Scheme**. In the previous section the inner zone radius, $R_o$, was chosen such that all atoms were contained in the inner zone. This condition can always be achieved in finite systems by choosing $R_o$ to be larger than any nuclear−nuclear distance in the polyatomic system. In this subsection, we demonstrate how the zone partitioning scheme can be used to take advantage of the rapid decay of operators when represented in local bases. By using systems large enough to profit from this partitioning scheme, we will demonstrate how the selection of the zone radius affects the accuracy and efficiency associated with computing the prototypical matrix elements defined in eq 29. The effect of using three inner zone radii (7,10, and 15 au) for two large molecules ($C_{60}$ and $C_{240}$) and three crystals (MgO, eight atoms in a simple cubic cell, $a = 4.21$ Å; $TiS_2$, trigonal, three atoms per cell, $a = 3.4073$ Å, $c = 5.6953$ Å; silica_zms5, orthorhomobic, 288 atoms per unit cell, $a = 20.07$ Å, $b = 19.92$ Å, $c = 13.42$ Å) is shown in Table 4.

**TABLE 3: Number of Grid Points Generated by $\int \rho_i V \, d\vec{r}$ ($V = V_{\text{eff}}^{\text{LDA}}$ and $V_{\text{xc}}^{\text{LDA}}$) at Five Accuracy Targets and Resulting Maximum Absolute Errors (Lower Rows) for 21 Molecules[a]**

| tolerence | E-3 | E-4 | E-5 | E-6 | E-7 |
|---|---|---|---|---|---|
| $H_2O$ | 1392 | 4031 | 16476 | 33609 | 110892 |
| | 1.4e-4 | 2.0e-5 | 1.6e-6 | 6.6e-7 | 1.1e-8 |
| $NH_3$ | 1665 | 5132 | 14337 | 38730 | 142149 |
| | 8.0e-4 | 1.9e-5 | 8.5e-7 | 3.5e-7 | 5.1e-8 |
| $CH_4$ | 1534 | 3830 | 20382 | 37865 | 87075 |
| | 5.4e-4 | 3.0e-5 | 8.9e-6 | 3.3e-7 | 1.6e-7 |
| $C_2H_6$ ($D_{3h}$) | 2263 | 5526 | 12242 | 37281 | 135200 |
| | 2.3e-4 | 2.5e-5 | 2.8e-6 | 7.2e-7 | 9.2e-8 |
| $C_2H_5OH$ | 4149 | 9357 | 23099 | 82798 | 346674 |
| | 8.4e-4 | 9.6e-5 | 1.9e-6 | 1.7e-7 | 1.0e-7 |
| ClOH | 1741 | 3715 | 6158 | 19822 | 83779 |
| | 3.5e-5 | 3.4e-5 | 7.8e-6 | 4.3e-7 | 3.5e-8 |
| $BH_3$ | 1161 | 4101 | 9529 | 48492 | 196244 |
| | 1.7e-4 | 3.2e-5 | 3.1e-6 | 3.8e-7 | 7.9e-8 |
| CClFH−CClFH(Ci) | 3992 | 10705 | 27575 | 73968 | 246193 |
| | 3.6e-4 | 9.7e-6 | 2.7e-6 | 3.3e-7 | 2.2e-8 |
| $C_2H_6$ ($D_{3d}$) | 2579 | 5742 | 10890 | 32203 | 138805 |
| | 3.5e-4 | 5.0e-5 | 3.0e-6 | 4.1e-7 | 1.4e-8 |
| CFH−CFH | 2300 | 5282 | 16907 | 72406 | 146578 |
| | 3.8e-4 | 5.7e-5 | 3.0e-6 | 2.1e-7 | 5.9e-8 |
| $CCl_2-CH_2$($C_{2v}$) | 2234 | 5655 | 14129 | 39726 | 190967 |
| | 6.3e-5 | 5.3e-5 | 3.9e-6 | 5.3e-7 | 5.4e-8 |
| $C_6H_6$ | 4147 | 9884 | 24497 | 86139 | 421821 |
| | 6.4e-4 | 2.5e-5 | 1.2e-5 | 7.2e-7 | 1.4e-8 |
| $SF_6$ | 2828 | 6133 | 13328 | 26265 | 65114 |
| | 3.0e-4 | 3.2e-5 | 7.8e-6 | 7.1e-8 | 2.9e-8 |
| $C_5H_5N_5$ | 5647 | 13728 | 36684 | 104195 | 431752 |
| | 4.1e-4 | 1.3e-5 | 1.8e-5 | 2.9e-6 | 3.4e-7 |
| $ZnCl_3(C_5H_6N_5)$ | 6373 | 14085 | 33244 | 91886 | 360060 |
| | 7.3e-4 | 3.2e-4 | 6.8e-6 | 2.0e-6 | 4.1e-7 |
| $C_5H_6N_2O_2-C_6H_4O_2$ | 11447 | 24996 | 62402 | 198280 | 593262 |
| | 1.8e-4 | 3.3e-5 | 9.8e-6 | 1.1e-6 | 2.0e-8 |
| $C_{12}H_{14}N_2O_2$ | 8969 | 16882 | 50088 | 113325 | 445389 |
| | 7.4e-5 | 2.9e-5 | 7.6e-6 | 8.9e-7 | 1.2e-7 |
| $C_{15}H_{12}N_2O_2$ | 8868 | 21015 | 66982 | 108391 | 282218 |
| | 1.9e-3 | 2.7e-4 | 3.8e-6 | 1.5e-6 | 5.3e-7 |
| $C_{17}H_{21}NO_4$ | 13886 | 38621 | 68149 | 200098 | 606887 |
| | 9.8e-4 | 1.4e-4 | 2.1e-5 | 5.6e-7 | 1.7e-7 |
| $C_{20}H_{29}N_3O_2$ | 28297 | 60466 | 137569 | 351557 | 888362 |
| | 5.3e-4 | 6.4e-6 | 3.8e-5 | 2.0e-7 | 5.3e-7 |
| $C_{60}$ | 22375 | 49225 | 102957 | 242240 | 690930 |
| | 5.3e-4 | 1.5e-4 | 1.3e-5 | 2.7e-6 | 7.7e-8 |

[a] No advantage has been taken of molecular symmetry.

Several issues associated with making a proper selection of $R_o$ can be understood by inspecting the data in Table 4 associated with $C_{60}$. For example, the results reported in that table for $R_o = 15$ au can be seen to be identical to that in Table 3. This is because all atoms in this molecule are within 15 au of each other and are therefore contained in the inner zone. The lack of an entry for the accuracy target $10^{-7}$ for $R_o = 7$ au, however, is an example of selecting a radius that is too small to economically achieve the accuracy target. In this case, the integrand has not been properly decomposed and even very large one-center grids centered on inner zone atoms (300 000 points/atom) are unable to integrate the integrand to the desired levels. Under such circumstances, it is less expensive and more reliable to choose a larger inner zone radius. Using a slightly larger radius will more effectively decompose the integrand and reduce the contribution to the integral from the outer zone. This increase in inner zone radius allows smaller grids to be used in the integration of the outer zone at the expense of increasing the number of atoms (and therefore one-center grids) in the inner zone. Actual physical systems, however, are not very sensitive

**TABLE 4: Number of Grid Points, as a Function of Inner Zone Radius, Required to Integrate the Same Quantity Evaluated in Table 3 for Several Large Molecules and Crystals[a]**

| tolerance | E-3 | E-4 | E-5 | E-6 | E-7 |
|---|---|---|---|---|---|
| | | | $C_{60}$ | | |
| $R_o = 7$ au | 6997 | 15616 | 53007 | 162458 | |
| | 6.0e-4 | 1.4e-4 | 1.2e-5 | 1.7e-6 | |
| $R_o = 10$ au | 12079 | 26582 | 57904 | 148052 | 505333 |
| | 5.0e-4 | 1.5e-4 | 1.2e-5 | 2.1e-6 | 1.4e-8 |
| $R_o = 15$ au | 22375 | 49225 | 102957 | 242240 | 690930 |
| | 5.3e-4 | 1.5e-4 | 1.3e-5 | 2.7e-6 | 7.7e-8 |
| | | | $C_{240}$ | | |
| $R_o = 7$ au | 4321 | 14343 | 81522 | 226418 | |
| | 1.2e-4 | 3.1e-5 | 8.7e-6 | 1.0e-6 | |
| $R_o = 10$ au | 11386 | 38291 | 57793 | 165434 | 495464 |
| | 9.5e-5 | 2.4e-5 | 1.2e-5 | 9.6e-7 | 8.0e-7 |
| $R_o = 1$ au | 23632 | 79748 | 119711 | 342664 | 996145 |
| | 9.1e-5 | 2.2e-5 | 1.2e-5 | 6.9e-7 | 2.3e-7 |
| | | | MgO | | |
| $R_o = 7$ au | 8348 | 20344 | 76695 | 225967 | 580917 |
| | 7.4e-4 | 8.3e-5 | 4.0e-6 | 4.3e-6 | 1.3e-7 |
| $R_o = 10$ au | 22898 | 55462 | 147281 | 359677 | 954336 |
| | 7.7e-4 | 1.7e-5 | 3.4e-6 | 6.0e-7 | 1.4e-7 |
| $R_o = 15$ au | 68746 | 168513 | 453378 | 1103836 | 2897164 |
| | 7.7e-4 | 7.8e-5 | 8.1e-6 | 5.5e-6 | 2.4e-7 |
| | | | $TiS_2$ | | |
| $R_o = 7$ au | 4907 | 12986 | 53405 | 105222 | 392591 |
| | 3.8e-4 | 4.7e-4 | 6.4e-4 | 8.2e-7 | 1.4e-7 |
| $R_o = 10$ au | 8555 | 21572 | 49725 | 161252 | 241751 |
| | 4.9e-4 | 3.6e-4 | 6.3e-4 | 1.1e-6 | 3.7e-7 |
| $R_o = 15$ au | 33188 | 81456 | 193083 | 602710 | 854030 |
| | 5.1e-4 | 4.0e-4 | 6.3e-4 | 1.5e-6 | 9.5e-7 |
| | | | silicious ZSM-5 | | |
| Ro = 7au | 4848 | 24576 | 80990 | 269816 | |
| | 2.3e-4 | 2.1e-4 | 2.4e-4 | 3.5e-6 | |
| $R_o = 10$ au | 11518 | 25669 | 86997 | 237137 | 1188304 |
| | 5.5e-4 | 1.1e-4 | 2.2e-4 | 2.3e-6 | 2.9e-8 |
| $R_o = 15$ au | 32788 | 73733 | 243115 | 563525 | 2897445 |
| | 5.5e-4 | 1.3e-4 | 2.2e-4 | 8.0e-7 | 5.9e-8 |

[a] No symmetry has been used in these calculations; maximum absolute errors are also reported.

to the specific selection of $R_o$. For all systems reported in Table 4, for example, an inner zone radius of 10 or 15 au has virtually no effect on the ability to achieve any of the five accuracy targets. According to our experience and the data in Table 4 a radius of 7−8 au will cost-effectively produce the accuracy target typical in electronic structure theory (in the range of $10^{-3}$−$10^{-4}$ au). For higher accuracy calculations an inner zone radius of 10 au is used in GAPSS[28,29] and recommended for general use.

With an appropriately chosen inner zone radius the number of grid points needed to evaluate an integral depends primarily on the number of atoms in the inner zone sphere (as opposed to the total number of atoms in the system). For this reason our approach exhibits excellent scaling as a function of system size. For example, the number of atoms inside the $R_o = 10$ au inner zone spheres of $C_{240}$ and $C_{60}$ are 36 and 32, respectively. The results in Table 4 show that the number of grid points for the two molecules at a given accuracy are approximately equal. Similar correlation can be found for the other systems in Table 4. In general, the cost of an integral with our method will not increase with increasing system size once some "critical" size has been exceeded. As a consequence, the larger the system the greater the savings with respect to conventional approaches. The application of our method to periodic systems is particularly valuable, since it provides a natural mechanism to distinguish between regions of the crystal that contribute heavily to a given integral (and therefore must be treated very carefully) and those regions where a more approximate treatment is sufficient. The data for the bulk crystalline systems, MgO, TiS2, and ZSM-5, demonstrate that the method behaves very similarly for infinite periodic systems.

**TABLE 5**

(a) Comparison of Equilibrium Lattice Constant, Bulk Modulus, Direct and Indirect Band Gaps, and Valence Bandwidth of Crystalline Si with Other Typical LDA Calculations, and Experiment

| | present | other LDA | experiment[c] |
|---|---|---|---|
| $a$ (Å) | 5.4165 | 5.4241,[a] 5.4098[b] | 5.4294 |
| $B$ (GPa) | 101.2 | 99,[a] 100[b] | 100.8 |
| $E_{g,direct}$ (eV) | 2.80 | 2.6[b] | 4.185 |
| $E_{g,indirect}$ (eV) | 0.69 | 0.51,[a] 0.7[b] | 1.17 |
| VB width (eV) | 11.86 | 11.8[b] | 12.5 |

(b) Equilibrium Lattice Constant, Bulk Modulus, and Cohesive Energy of Face-Centered Cubic Al, Calculated by the Method of the Present Paper with the LDA and PBE-96 GGA Functionals, Compared with Another LDA Calculation and Experiment

| | present, LDA | present, GGA | other LDA | expt |
|---|---|---|---|---|
| $a$ (Å) | 3.9875 | 4.0418 | 4.0199,[d] 4.0112[e] | 4.022[e] |
| $B$ (GPa) | 82.9 | 74.1 | 79.7,[d] 71.5[e] | 72.7[f] |
| $E_{coh}$ (eV) | 3.98 | 3.50 | 4.07,[d] 3.65[e] | 3.37[g] |

[a] Reference 35. [b] Reference 36. [c] Reference 37. [d] Reference 19a. [e] Reference 38. [f] Reference 39. [g] Reference 40.

The scheme just discussed is fully implemented in our periodic Gaussian basis DFT code GAPSS and is routinely used to efficiently obtain physically correct results. As brief examples of applications of our method, we show calculations of bulk properties of silicon and aluminum in Table 5. These results were obtained with our code GAPSS that finds exchange−correlation energies and Hamiltonian matrix elements by the integration techniques described above. Other aspects of GAPSS have been described elsewhere.[28] Orbital basis sets were of approximately TZVP quality and were constructed by reoptimizing molecular DFT basis sets[32] for the crystalline environment. Density basis sets used in the Coulomb part of the calculation were taken from the molecular sets[32] with essentially no modifications. We used the VWN parametrization[33] of the local density approximation for Si and Al and the PBE96 form[34] of the generalized gradient approximation for Al. Structural parameters were obtained from calculated energy versus volume curves by fitting to Murnaghan's equation. Our results are in close agreement with the typical high-quality LDA calculations[19a,35,36,38] shown in Table 5. As usual, the LDA results all slightly underestimate the lattice constants and overestimate the cohesive energies relative to experiment[37−40] but give good structural properties overall. The GGA yields improved values for the bulk modulus and cohesive energy of Al relative to experiment but slightly overcorrects the lattice constant, as is typical for that approximation. Also as expected, the DFT eigenvalue differences from all calculations (Table 5a) are significantly smaller than the corresponding experimental excitation energies, but here again, our results agree closely with other calculations. Thus, our method gives accurate results for prototypical examples of both covalent (Si) and metallic (Al) bonding in extended systems.

## V. Concluding Remarks

In this work, we have formulated a multicenter integration scheme where the cost for an integral does not increase with increasing system size. This method can be efficiently applied to periodic systems where the application of economical

algorithms is critical. We have demonstrated that the largest gains in efficiency originate from our new decomposition scheme. We have also proposed and discussed an adaptive one-center integration scheme and automatic grid generator and compared the capabilities of these formulations to those that have been previously published. Finally, we have demonstrated that when combined, our one-center integration scheme and decomposition strategy are more efficient, to the best of our knowledge, than previously published methods.

**Acknowledgment.** We acknowledge the U.S. DOE, Office of Basic Energy Sciences, Division of Chemical Sciences and Geosciences and Engineering for support of this work.

## References and Notes

(1) Hohenberg, P.; Kohn, W. *Phys. Rev. B* **1964**, *136*, 864.
(2) Kohn, W.; Sham, L. *J. Phys. Rev. A* **1965**, *140*, 1133.
(3) Becke, A. D. *J. Chem. Phys.* **1988**, *88*, 2547.
(4) Boys, S. F.; Rajagopal, P. *Adv. Quantum Chem.* **1965**, *2*, 1.
(5) te Velde, G.; Baerends, E. J. *J. of Comput. Phys.* **1992**, *99*, 84.
(6) Delley, B. *J. Chem. Phys.* **1990**, *92*, 508.
(7) Pederson, M. R.; Jackson, K. A. *Phys. Rev. B* **1990**, *41*, 7453.
(8) Andzelm, J.; Wimmer, E. *J. Chem. Phys.* **1992**, *96*, 1280.
(9) Murray, C.; Handy, N.; Laming, G. *Mol. Phy.* **1993**, *78*, 997.
(10) Gill, P.; Johnson, B.; Pople, J. *Chem Phys. Lett.* **1993**, *209*, 506.
(11) Perez-Jorda, J.; Becke, A.; San-Fabian, E. *J. Chem. Phys.* **1994**, *100*, 6520.
(12) Treutler, O.; Ahlrich, R. *J. Chem. Phys.* **1995**, *102*, 346.
(13) Perez-Jorda, J. *J. Chem. Phys.* **1994**, *101*, 1738.
(14) Mura, M. E.; Knowles, P. J. *J. Chem. Phys.* **1996**, *104*, 9848.
(15) Stratmann, R. E.; Scuseria, G. E.; Frisch, M. J. *Chem. Phys. Lett.* **1996**, *257*, 213.
(16) Hirshfeld, F. L. *Theor. Chim. Acta B* **1977**, *44*, 129.
(17) Szabo, Attila; Ostlund, Neil S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; Macmillan Publishing Co., Inc.: New York, 1982.
(18) Labanowski, J. K., Andzelm, J. W., Eds. *Density Functional Methods in Chemistry*; Springer-Verlag: New York, 1991.
(19) See, for example, the following. (a) Boettger, J. C.; Trickey, S. B. *Phys. Rev. B* **1996**, *53*, 3007. (b) Schroer, P.; Krueger, P.; Pollmann, J. *Phys. Rev. B* **1993**, *47*, 6971.
(20) Stroud, A. H. *Approximate Calculation of Multiple Integrals*; Prentice-Hall: Englewood Cliffs, NJ, 1971.
(21) Lebedev, V. I. *Zh. Vyechisl. Mat. Mat. Fiz.* **1975**, *15*, 48.
(22) Lebedev, V. I. *Zh. Vyechisl. Mat. Mat. Fiz.*, **1976**, *16*, 293.
(23) Lebedev, V. I. *Sib. Mat. Zh.* **1977**, *18*, 132.
(24) Delley, B. *J. Comput. Chem.* **1996**, *17*, 1152.
(25) Johnson, B. G.; Frisch, M. J. *Chem. Phys. Lett.* **1993**, *216*, 133.
(26) Johnson, B. G.; Gill, P. M. W.; Pople, J. A. *Chem. Phys. Lett.* **1994**, *220*, 377.
(27) Johnson, B. G.; Frisch, M. J. *J. Chem. Phys.* **1994**, *100*, 7429.
(28) Jaffe, J. E.; Hess, A. C. *J. Chem. Phys.* **1996**, *105*, 10983.
(29) Lin, Zijing; Hess, A. C. In preparation.
(30) By so-called backward trimming, ref 13 gives significantly better results than ref 11. Nevertheless, we focus on comparing our results to those of ref 11 mainly because (1) it is easier to make detailed comparison, since there are more data in ref 11, (2) the comparison is primarily for separating the effects of 1C grid points from those of multicenter decomposition schemes, and (3) higher efficiency could be achieved by various techniques, such as "pruning" and searching the $\theta$ and $\varphi$ grids separately. However, this is not interesting here, as discussed in the text.
(31) It is important to note that the specific form of the partitioning function used in this work (or the work of others for that matter) is not unique. In the latest versions of Dmol, for example, a functional form different from that cited in ref 6 is currently being used.
(32) Godbout, N.; Salahub, D. R.; Andzelm, J.; Wimmer, E. *Can. J. Chem.* **1992**, *70*, 560.
(33) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200.
(34) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865.
(35) Boettger, J. C. *Int. J. Comput. Chem. Symp.* **1996**, *30*, 133.
(36) Lambrecht, W. R. L.; Andersen, O. K. *Phys. Rev. B* **1986**, *34*, 2439.
(37) *Semiconductors: Group IV Elements and III−V Compounds*; Madelung, O., Ed.; Springer-Verlag: Berlin, 1991; Section 1.2.
(38) Lam, P. K.; Cohen, M. L. *Phys. Rev. B* **1981**, *24*, 4224.
(39) Greene, R. G.; Luo, H.; Ruoff, A. L. *Phys. Rev. Lett.* **1994**, *73*, 2075.
(40) Boettger, J. C.; Trickey, S. B. *Phys. Rev. B* **1984**, *29*, 6434.