

LAWRENCE LIVERMORE NATIONAL LABORATORY

Partition-of-unity finite-element method for large scale quantum molecular dynamics on massively parallel computational platforms

J. E. Pask, N. Sukumar, M. Guney, W. Hu

March 1, 2011

Disclaimer

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

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

Partition-of-unity finite-element method for large scale quantum molecular dynamics on massively parallel computational platforms

LDRD 08-ERD-052 Final Report

John E. Pask (PI)

Lawrence Livermore National Laboratory, Livermore, California 94550

N. Sukumar (Co-PI), M. Guney, and W. Hu

Dept. of Civil and Environmental Eng., University of California, Davis, California 95616

Abstract

Over the course of the past two decades, quantum mechanical calculations have emerged as a key component of modern materials research. However, the solution of the required quantum mechanical equations is a formidable task and this has severely limited the range of materials systems which can be investigated by such accurate, quantum mechanical means. The current state of the art for large-scale quantum simulations is the planewave (PW) method, as implemented in now ubiquitous VASP, ABINIT, and QBox codes, among many others. However, since the PW method uses a global Fourier basis, with strictly uniform resolution at all points in space, and in which every basis function overlaps every other at every point, it suffers from substantial inefficiencies in calculations involving atoms with localized states, such as first-row and transition-metal atoms, and requires substantial nonlocal communications in parallel implementations, placing critical limits on scalability. In recent years, *real-space* methods such as finite-differences (FD) and finite-elements (FE) have been developed to address these deficiencies by reformulating the required quantum mechanical equations in a strictly local representation. However, while addressing both resolution and parallel-communications problems, such local real-space approaches have been plagued by one key disadvantage relative to planewaves: excessive degrees of freedom (grid points, basis functions) needed to achieve the required accuracies. And so, despite critical limitations, the PW method remains the standard today.

In this work, we show for the first time that this key remaining disadvantage of real-space methods can in fact be overcome: by building known atomic physics into the solution process using modern *partition-of-unity* (PU) techniques in finite element analysis. Indeed, our results show order-of-magnitude reductions in basis size relative to state-of-the-art planewave based methods. The method developed here is completely general, applicable to any crystal symmetry and to both metals and insulators alike. We have developed and implemented a full self-consistent Kohn-Sham method, including both total energies and forces for molecular dynamics, and developed a full MPI parallel implementation for large-scale calculations. We have applied the method to the gamut of physical systems, from simple insulating systems with light atoms to complex d- and f- electron systems, requiring large numbers of atomic-orbital enrichments. In every case, the new PU FE method attained the required accuracies with substantially fewer degrees of freedom, typically by an order of magnitude or more, than the current state-of-the-art PW method. Finally, our initial MPI implementation has shown excellent parallel scaling of the most time-critical parts of the code up to 1728 processors, with clear indications of what will be required to achieve comparable scaling for the rest.

Having shown that the key remaining disadvantage of real-space methods can in fact be overcome, the work has attracted significant attention: with sixteen invited talks, both domestic and international, so far; two papers published and another in preparation; and three new university and/or national laboratory collaborations, securing external funding to pursue a number of related research directions. Having demonstrated the proof of principle, work now centers on the necessary extensions and optimizations required to bring the prototype method and code delivered here to production applications.

1 Introduction

Of paramount importance in the investigation of new and unknown materials, and of more common materials under extreme conditions, is the generality and accuracy of the theoretical methods employed in their study. First and foremost, a first-principles (*ab initio*) quantum mechanical description is required whenever departures from isolated-atomic or known condensed-matter configurations may be significant. However, a merely *ab initio* approach is not enough in the investigation of such unfamiliar materials systems: the approach must be general, equally applicable to all atomic species and configurations, and systematically improvable so that errors can be clearly known and strictly controlled.

The PW pseudopotential method [1] is among the most widely used *ab initio* methods which afford this level of generality and systematic improvability. The accuracy and generality of the PW method arises from its nature and basis: a variational expansion approach in which solutions are represented in a Fourier basis. By virtue of the completeness of the basis, any condensed matter system can be modeled with arbitrary accuracy, in principle, by simply adding sufficient wavenumbers to the basis. In practice, however, the PW method has a number of significant limitations with respect to the solution of large, complex problems, and this has inspired extensive research in the past decade into alternative approaches better suited for such problems [2–4]. Among the more mature of these to date are the real-space finite-difference (FD), finite-element (FE), and wavelet based approaches.

Like the standard PW method [1], the FE method [2–5] is a variational expansion approach. However, whereas the PW method employs a Fourier basis, with every basis function overlapping every other at every point in the domain, the FE method employs a basis of strictly local piecewise polynomials, each overlapping only its immediate neighbors. Because the FE basis consists of polynomials, the method is completely general and systematically improvable, like the PW method. Because the FE basis is strictly local, however, the method offers some significant advantages with respect to large-scale calculations. First, because the FE basis functions are localized, they can be concentrated where needed in real space to increase the efficiency of the representation. The PW basis has the same resolution at all points in space and this leads to substantial inefficiencies in the solution of highly inhomogeneous problems such as those involving first-row and transition-metal atoms, and in finite or semi-infinite system calculations such as those involving atoms, molecules, or surfaces. Recent progress on this issue in the context of PW bases includes ultrasoft pseudopotentials [6], optimized pseudopotentials [7,8], adaptive coordinate transformations [9–11], and, more recently, the projector augmented wave (PAW) method [12]. However, in working to soften the problem (allowing less strongly localized solutions), these developments apply equally to real-space approaches. Second, the FE method can accommodate a wide variety of boundary conditions: Dirichlet boundary conditions for molecules or clusters, Bloch boundary conditions for crystals, a mixture of these for surfaces, etc. The PW method is limited to periodic boundary conditions. Third, and most significantly for large-scale calculations, the strict locality of the FE basis facilitates massively parallel computations by minimizing the need for nonlocal communications. PW methods rely on Fourier transforms which can lead to substantial inefficiencies in such distributed computations due to the need for large numbers of computational nodes to communicate with one another. Recent progress on this issue in the context of PW bases includes new fast Fourier transform formulations [13, 14], localized planewave formulations [15], and efficient distribution schemes for electronic states and planewave coefficients on large numbers of processors [16].

The advantages of a strictly local, real-space approach in large-scale calculations have been amply demonstrated in the context of finite difference (FD) methods [2,3,17–27]. These methods allow for some variable resolution in real space, can accommodate a variety of boundary conditions, and require no computation- or communication-intensive transforms. However, FD methods achieve these advantages by giving up the use of a basis altogether, instead discretizing all terms directly on a real-space grid, which leads to disadvantages such as limited accuracy in integrations [3,23,28] and nonvariational convergence [2,3]. Perhaps most significantly, however, because FD methods lack a basis, it is difficult to build known physics into the methods in order to increase the efficiency of the representation; whereas basis oriented approaches such as the FE method allow for the possibility of incorporating such known physics by, for example, adding functions to the basis which contain that physics. Indeed, as we discuss below, we exploit this property in the present work to substantially reduce the required basis size. By retaining the use of a basis while remaining strictly local in real space, FE methods combine significant advantages of both PW and FD approaches.

The main disadvantages of FE methods relative to FD methods are that they can be substantially more complex to implement, produce generalized rather than standard eigenproblems (as produced by most FD methods), and can require more storage. The greater complexity of FE based approaches has limited their maturation relative to FD based approaches but the gap has continued to narrow as extensive efforts have continued over the years and more high-quality standard FE libraries have become freely available [3]. The additional storage required for system matrices is not an issue for large problems since it grows only linearly with system size while other storage requirements common to both methods, such as for eigenvectors, grow quadratically. Finally, recent advances in iterative eigensolvers [25, 29, 30] and efficient parallel implementations, have greatly reduced the relative difficulty of solving the generalized eigenproblem. Most importantly for the present work, however, both increased storage and generalized eigenproblem may be more than compensated by exploiting the much greater flexibility of the FE method to concentrate degrees of freedom where needed, by incorporation of known physics into the basis, thus substantially reducing the degrees of freedom, and so eigenproblem size, required to achieve a given accuracy.

The finite-element method [5] has had a long history of success in diverse applications ranging from civil engineering to quantum mechanics. Applications in engineering go back to the 1950s. Applications to the electronic structure of isolated atomic and molecular systems began to appear in the 1970s [31,32]. White et al. [33] applied the method to full 3D atomic and molecular calculations in 1989, finding substantial advantages from the generality and strict locality of the basis but also inefficiencies in all-electron calculations of small molecules with uniform meshes due to the rapid variations of the wavefunctions near the nuclei. Applications to solids have appeared more recently. Hermansson and Yevick [34] applied the FE method to non-self-consistent solid-state electronicstructure calculations in 1986, finding that it was less efficient than the PW method for small problems with relatively smooth potentials, when uniform meshes and cubic- or lower-order bases were employed. Subsequent work has focused on the application of the method to large-scale calculations, where the strict locality of the basis can provide significant advantages in the context of massively parallel computations. Tsuchida and coworkers [35–39] have applied the method to full self-consistent molecular and solid-state electronic-structure calculations. They have implemented nonuniform meshes [35], adaptive coordinate transformations [36], pseudopotential and all-electron calculations [35], molecular dynamics [37], and an O(N) formulation [37]; and have demonstrated the favorable efficiency of the method relative to FD approaches [36]. Their calculations so far have been confined to orthorhombic unit cells, Γ -point Brillouin zone sampling, and insulating systems. The PI and coworkers [4,40–45] have formulated a completely general FE based electronic-structure method, allowing arbitrary Bravais lattices, full Brillouin zone sampling, and the treatment of metallic as well as insulating systems. Initial non-self-consistent positron calculations in solids [41] demonstrated the capacity of the FE based approach for large-scale calculations, accomplishing the largest such calculations reported to that time. The PI and coworkers have addressed, in particular,



Figure 1: Self-consistent FE and exact energy bands for Cu for $4 \times 4 \times 4$ and $6 \times 6 \times 6$ uniform meshes of cubic elements. The FE bands are essentially indistinguishable from the exact self-consistent solution over the whole of the k-space path already for the $6 \times 6 \times 6$ mesh. (The "exact" solution is taken from a well converged PW calculation.)

the enforcement of the required Bloch-periodic boundary conditions in the context of a general C^0 basis [40], the use of nonlocal operators in the context of self-consistent calculations [4], and the computation of the crystal potential and total energy in real space [44]. They have recently reported the first FE based *ab initio* electronic-structure calculations of metals [45] (Figure 1). Torsti *et al.* [3] have recently reported initial non-self-consistent calculations of bulk Si and have demonstrated the remarkable efficiency of FE methods with advanced mesh refinement in side by side comparisons with FD methods. Work continues [46–48] along the classical mesh refinement path to treat the rapid oscillations in the wavefunctions in the vicinity of the atomic nuclei; however, in exploiting only the *scale* of oscillations and not their known orbital *nature*, such methods must invariably overcome a quite substantial degree-of-freedom disadvantage relative to mature orbital-based methods [49–52] in order to be competitive.

The great flexibility of the finite-element method to concentrate degrees of freedom in real space where needed and omit them where not, via standard mesh refinement techniques, allows a much more efficient representation of highly inhomogeneous problems, as occur in *ab initio* electronic structure, than is possible by either FD or PW approaches. However, it has recently been recognized in the context of molecular calculations that one can do much better still by adding functions to the standard FE basis which contain the known physics in the most difficult parts of the domain i.e., in the vicinity of the atomic nuclei, in the present case, where the wavefunctions vary most rapidly. Dusterhoft et al. [53] added atomic orbitals to the standard FE basis in 2D axisymmetric all-electron calculations of C_2 . They found an order of magnitude increase in efficiency relative to standard mesh-refinement techniques. More recently, Yamakawa and Hyodo [54,55] demonstrated the substantial gains in efficiency which can be accomplished by adding well-chosen Gaussian basis functions to the standard FE basis in full 3D all-electron molecular calculations. A significant disadvantage of the above approaches, however, noted in Ref. [55], is that by adding such extended functions to the otherwise strictly local FE basis, sparseness and locality are compromised, leading to inefficiencies in parallel implementation. Modern partition-of-unity (PU) techniques [56–59] in finite-element analysis provide an elegant and highly efficient solution to this problem. Recently, the PI and collaborators have addressed the enforcement of Bloch-periodic boundary conditions and the use of orbital enrichment in a *partition-of-unity finite-element* (PUFE) formulation for quantummechanical eigenproblems [60], and the application of *enriched finite-element* (EFE) methods to the all-electron Coulomb problem [61]; demonstrating in both contexts substantial reductions in basis size relative to classical FE. Most recently, order-of-magnitude reductions in basis size relative to current state-of-the-art PW methods in full self-consistent Kohn-Sham calculations have been demonstrated [62] (see Section 3.1 below).

Clearly, in the presence of such precise a priori information about the nature of the solution, as in the present context, it can be of great advantage to use that information to arrive at the solution most efficiently. This has been appreciated from the earliest days of multi-atom calculations in quantum chemistry, employing all manner of Slater, Gaussian, and/or numerical orbital basis sets [51, 52, 63]. In the present work, we show for the first time that this can in fact be done while retaining both strict locality and systematic improvability of the basis as a whole: by building known atomic physics into the solution process using modern partition-of-unity techniques in finite element analysis. Indeed, our initial results show order-of-magnitude reductions in basis size relative to state-of-the-art PWs for a broad range of problems, especially those involving localized states. The method developed here is completely general, applicable to any crystal symmetry and to both metals and insulators alike. We have developed and implemented a full self-consistent Kohn-Sham method, including both total energies and forces for molecular dynamics, and developed a full MPI parallel implementation for large-scale calculations. We have applied the method to the gamut of physical systems, from simple insulating systems with light atoms to complex d- and f- electron systems, requiring large numbers of atomic-orbital enrichments. In every case, the new PUFE method attained the required accuracies with substantially fewer degrees of freedom, typically by an order of magnitude or more, than the current state-of-the-art PW method. Finally, our initial MPI implementation has shown excellent parallel scaling of the most time-critical parts of the code up to 1728 processors, with clear indications of what will be required to achieve comparable scaling for the rest.

2 Partition-of-unity finite elements

The partition of unity (PU) method [56–59] generalized the classical finite-element (FE) method by providing a means to incorporate local asymptotic solutions of boundary-value problems into the FE approximation. Any set of functions $\{\phi_i(\mathbf{x})\}_{i=1}^n$ that sum to unity form a partition of unity. The generalized finite-element method [64–66] and the extended finite-element method [67–71] are both particular instances of the partition-of-unity finite-element (PUFE) method. The PU paradigm permits the modeling of singularities, discontinuities, or sharp gradients on a fixed mesh. The striking advantages are that regular meshes suffice, system matrices are sparse, and high-accuracy can be obtained on relatively coarse meshes. In addition, in contrast to classical FE, remeshing is not needed for moving boundary problems. Early applications were in the modeling of holes and cracks in 2D [70] and 3D [71], and in coupling with level sets and fast marching methods to model material interfaces [72] and three-dimensional crack growth [73–75]. There is growing interest in PUFE methods and new applications continue to emerge—for example, wave propagation [76–78], phase transformations [79], multiscale modeling [80], and reduced-order modeling [81] in recent years.

In the PUFE method, in addition to the classical FE basis functions, basis functions that can represent discontinuities in the physical quantities [68] or their derivatives [72, 82] are included so that these geometric features can be represented without the mesh needing to conform to them. For example, a jump discontinuity in the pressure across a surface would indicate that the usual basis



Figure 2: Modeling a discontinuity at $x = x_0$. (a) Classical FE bases; and (b) Discontinuous enriched bases. The function H(x) = 0 if $x < x_0$ and unity otherwise. Nodes 2 and 3 are enriched.

functions should be enriched with a Heaviside function, which is zero on one side of the surface and unity on the other. The additional Heaviside function is known as an *enrichment function*. The classical and enriched basis functions on a 1D grid are shown in Figure 2. The PU method goes further than the mere representation of jump discontinuities; enrichment functions can also include, for example, the form of an asymptotic solution near a singularity or essentially any known property of the solution.

In classical FE, a basis function $\phi_j(\mathbf{x})$ is associated with node j in the mesh. In Figure 2a, classical FE basis functions in 1D are illustrated. Let $\omega_j = {\mathbf{x} : \phi_j(\mathbf{x}) > 0}$ be the region of support for $\phi_j(\mathbf{x})$. The nodes belonging to an element are given by the connectivity of the element, whereas the dual ω_j , is the collection of elements that are associated with a specific node j. The PUFE approximation for a scalar-valued function u is of the general form [56]:

$$u^{h}(\mathbf{x}) = \sum_{j=1}^{n} \phi_{j}(\mathbf{x}) \left(u_{j} + \sum_{\alpha=1}^{m} \varphi_{\alpha}(\mathbf{x}) a_{j\alpha} \right),$$
(1)

where $\varphi_{\alpha}(\mathbf{x})$ are enrichment functions, u_j are classical degrees of freedom, and $a_{j\alpha}$ are additional enriched degrees of freedom. The support of the enriched basis $\varphi_{\alpha}\phi_j$ is identical to that of ϕ_j . If the span of the exact solution for a boundary-value problem is contained in the PUFE basis, then the PUFE solution will be identical to the exact solution. Note that a single node can be enriched using many different enrichment functions. The caveat, however, is that the enriched basis must constitute a linearly independent set, which is more often the case with non-polynomial enrichment functions than with higher-order polynomials, which can introduce linear dependencies [83].

3 Application of PUFE method to the Kohn-Sham equations

In *ab initio* calculations, the required solutions (wavefunctions and potentials) vary most rapidly in the vicinity of the atomic nuclei and become more and more atomic-like the closer one gets to the atomic centers. Hence, we employ atomic wavefunctions as enrichments in the solution of the crystal eigenproblem and atomic potentials as enrichments in the solution of the associated Coulomb problem.

3.1 Self-consistent calculations

Because the Fourier basis is global, the convergence of the energies and eigenvalues with respect to number of basis functions (DOFs) in the PW method is spectral, i.e., faster than any polynomial. Whereas, being polynomial based, the convergence of FE and other such real-space methods is determined by polynomial completeness. For a polynomial basis complete to order p, the errors



Figure 3: Error in total energy of LiH versus number of degrees of freedom using standard PW, real-space FE, and new real-space PUFE methods.

in energies and eigenvalues of self-adjoint operators are $O(h^{2p})$ while the errors in the associated eigenfunctions are $O(h^{p+1})$ [5]. Hence, for sufficiently high accuracies, the spectral convergence of PW based methods must dominate and require fewer DOFs than fixed-degree polynomial based methods such as FE. However, at lower accuracies, with PU enrichment in particular, the FE basis can require fewer DOFs, substantially fewer in fact, as we show below.

To assess the efficiency of the PUFE method relative to current state-of-the-art electronicstructure methods on real, 3D, self-consistent problems, we developed a prototype self-consistent PUFE code by extending our existing conventional FE electronic structure code according to the formulation of Ref. [60]. We then compared the prototype PUFE code head-to-head against a standard PW code and modern conventional FE code in total energy calculations of LiH, using hard, transferable HGH pseudopotentials [84]. The results were striking. Figure 3 shows the error in total energy for LiH versus number of degrees of freedom (basis functions) using standard PW, modern real-space FE, and new prototype real-space PUFE codes. The horizontal line indicates an error of 10^{-3} Ha, typical in *ab initio* calculations. First, note that PW calculations required a factor of 8 fewer DOFs to reduce the error to the required level than the real-space cubic FE method: which demonstrates the extent of the DOF disadvantage real-space methods must overcome to be competitive with planewaves. Now consider the cubic PUFE result. This achieved the required level of accuracy with a *factor of 16 fewer DOFs* than the PW method.

Moreover, since real-space codes use local bases, they tend to be less computationally expensive per degree of freedom than planewave codes. Our experience comparing well optimized parallel FD code MGmol [85] and planewave code Qbox [16] on moderate numbers of processors indicates about a factor of 3 advantage for the real-space code with respect to computational expense per degree of freedom. Others [86] have found even larger advantages. Hence, the factor of 16 fewer degrees of freedom for the real-space PUFE code relative to planewaves suggests *well over an order magnitude speedup* relative to planewaves for a mature optimized implementation. And, due to the strict locality of the basis, that advantage should only increase with increasing numbers of processors.

To assess the performance of the PUFE method in the worst case, we applied it to a difficult



Figure 4: KS-DFT calculations for CeAl. (a) Enrichment functions for Ce; (b) Enrichment functions for Al; and (c) Error in total energy of CeAl versus number of degrees of freedom using standard PW, real-space FE, and new real-space PUFE methods.



Figure 5: Schematic charge density in unit cell. Total density $\rho = \rho^+ + \rho^-$.

f-electron system: triclinic CeAl, with atoms displaced from ideal positions. Because Ce has a full complement of s, p, d, and f states in valence, it requires 17 enrichment functions to span the occupied space (whereas Li requires only two), making this a particularly severe test for the efficiency of PUFE relative to planewaves. The enrichment functions for Ce and Al are shown in Figs. 4a and 4b, respectively. Figure 4c shows the error in the total energy for CeAl versus number of degrees of freedom (basis functions) using standard PW, modern real-space FE, and new real-space PUFE codes. First, we note that PW calculations required a factor of 16 fewer degrees of freedom to reduce the error to the required level than the real-space cubic FE method, which again demonstrates the extent of the degree-of-freedom disadvantage real-space methods must overcome to be competitive with planewaves. Now consider the cubic PUFE result. Remarkably, even with the large number of enrichment functions for Ce, the PUFE method still achieved the required level of accuracy with a *factor of 5 fewer DOFs* than the current state-of-the-art PW method.

3.2 All-electron Coulomb problem

All-electron calculations in a FE basis require the calculation of the Coulomb potential V_C due to a net neutral charge density (nuclear and electronic) $\rho = \rho_n + \rho_e \equiv \rho^+ + \rho^-$. For periodic solids, we are confronted with three distinct divergences which must be addressed: (1) the 1/r divergence of the electrostatic potential at the nuclei, (2) the divergence of both potential and energy lattice sums due to the long-range 1/r nature of the Coulomb interaction, and (3) the infinite self energies of the nuclei. For finite systems (e.g., molecules), there are no long-range divergences, but the others persist. A common approach to handle such lattice summations is the Ewald method [87], which relies on Fourier transforms. The limitations of the reciprocal space approach have inspired much research on the Coulomb problem using real-space and local-orbital based approaches [2–4, 88, 89]. Other approaches include smearing the nuclear charge (distributed nucleus approximation) [90] and using nonsingular, though rapidly varying, electronic density as source [2, 3, 33, 46–48, 55, 91, 92]. In Pask et al. [61], we introduced a real-space approach to treat the all-electron Coulomb problem without distributed nucleus approximation, sphere-interstitial matching, need for FFTs, or Ewald summation.

To resolve the divergence of the Coulomb potential, we introduce a smooth neutralizing density $\tilde{\rho}$ and write the total charge ρ in the unit cell as

$$\rho(\mathbf{x}) = \rho^{+}(\mathbf{x}) + \rho^{-}(\mathbf{x}) = \underbrace{\rho^{+}(\mathbf{x}) - \tilde{\rho}(\mathbf{x})}_{\text{neutralized nuclear density}} + \underbrace{\rho^{-}(\mathbf{x}) + \tilde{\rho}(\mathbf{x})}_{\text{neutralized electronic density}} \equiv \tilde{\rho}^{+}(\mathbf{x}) + \tilde{\rho}^{-}(\mathbf{x}), \quad (2)$$

where $\rho^+(\mathbf{x}) = \sum_i q_i \delta(\mathbf{x} - \tau_i)$ is the total nuclear charge density and $\rho^-(\mathbf{x})$ is the electronic charge density (Figure 5). We form $\tilde{\rho}$ in the unit cell as a sum of smooth, strictly local densities $\tilde{\rho}_I$ centered



Figure 6: Finite element (FE) and enriched finite element (EFE) solutions for crystalline diamond. (a) Charge densities; (b) Enrichment functions; and (c) Error in Coulomb energy per atom.

at atomic positions τ_I with integrals $\int \tilde{\rho}_I = q_I$: $\tilde{\rho} = \sum_I \tilde{\rho}_I$, where the sum is over all sites I in the crystal such that $\tilde{\rho}_I \neq 0$ (i.e. $\tilde{\rho}_I$ nonvanishing) in the unit cell.

The total potential $V(\mathbf{x})$ in the unit cell may now be written as

$$V(\mathbf{x}) = V^{+}(\mathbf{x}) + V^{-}(\mathbf{x}) = \tilde{V}^{+}(\mathbf{x}) + \tilde{V}^{-}(\mathbf{x}) = \underbrace{\sum_{I} \frac{q_{I}}{|\mathbf{x} - \boldsymbol{\tau}_{I}|} - \tilde{V}_{I}(\mathbf{x})}_{\text{known analytically}} + \tilde{V}^{-}(\mathbf{x}), \qquad (3)$$

and therefore the potential associated with $\tilde{\rho}^-$ can be obtained from a solution of Poisson's equation

$$\nabla^2 \tilde{V}^-(\mathbf{x}) = -4\pi \tilde{\rho}^-(\mathbf{x}),\tag{4}$$

subject to periodic boundary conditions. Since the total Coulomb energy per unit cell in the allelectron case is divergent, we extract the divergent nuclear self-energy analytically. The weak form, discrete equations, and expressions for the total Coulomb energy are provided in Ref [61].

We now apply the method to the computation of the all-electron Coulomb potential in crystalline diamond. The unit cell has lattice constant a = 6.75, with fcc lattice vectors $\mathbf{a}_1 = a/2(0, 1, 1)$, $\mathbf{a}_2 = a/2(1, 0, 1)$, $\mathbf{a}_3 = a/2(1, 1, 0)$, and carbon atoms at positions $\tau_1 = (0, 0, 0)$ and $\tau_2 = (1/4, 1/4, 1/4)$ in lattice coordinates. We consider nuclear point charges $q_i = 6$ at positions τ_i and take the electronic density ρ^- from overlapping all-electron atomic densities. We take cutoff radius $r_c = 1.4$ a.u. so that the neutralizing densities $\tilde{\rho}_I(\mathbf{x}) = 6g(|\mathbf{x} - \tau_I|)$ are nonoverlapping, where g(r) is a smooth C^2 cutoff function which vanishes identically for $r > r_c$. We solve this problem using cubic FE and enriched FE (PUFE with all nodes enriched and constrained to have the same coefficient). The enrichment function is constructed using a one-dimensional radial spectral element solver to obtain the atomic potential solutions. The numerical results are shown in Figure 6: charge densities are plotted in Figure 6a, the two enrichment functions, one for each atom, are shown in Figure 6b, and the convergence for FE and enriched FE is illustrated in Figure 6c. The enriched finite element solution has an accuracy of order 10^{-5} Ha in the Coulomb energy for a $16 \times 16 \times 16$ mesh (28,674 degrees of freedom), whereas the the best uniform-mesh FE result provides an accuracy of only 10^{-2} Ha on a $32 \times 32 \times 32$ mesh (229,376 degrees of freedom).

3.3 Adaptive quadrature

In the preceding section, we presented a new formulation for the Coulomb problem, which was solved using enriched FE. An open and critical issue in enriched/PUFE techniques is efficient numerical



Figure 8: Computing $\int_{\Omega} f(r) d\mathbf{x}$. (a) Plot of f(r); (b),(c) Convergence with respect to number of integration points in each direction and minimum distance between Gauss points, respectively.

integration. Since the enrichment functions are sharp and localized, we resorted to high-order tensor-product quadrature to obtain the results shown in Figure 6 [61]; however, for the method to be viable, efficiency in computing the stiffness matrix and force vector entries ($\mathbf{Kd} = \mathbf{f}$) is paramount. For the integration of functions with peaks at or near a boundary over the interval, the Möbius transformation has been adopted in quantum chemistry [93–95]; for electronic-structure calculation with a 1/r term in the integrand of the Hamiltonian, Batcho [46] used the Duffy transformation [96, 97]. In general, since the atom can be located anywhere inside the element, the above approaches do not meet our needs. Instead, we have developed an adaptive quadrature scheme that yields remarkable accuracy at minimal cost—this provides the required efficiency to solve the all-electron equations of DFT using the PUFE approach.

The integrand of the element stiffness matrix contains terms such as ψ , ψ^2 , $\nabla \psi$, $\nabla \psi \cdot \nabla \psi$, and $\psi \nabla \psi$ (ψ is the enrichment function; see Figure 7). The numerical integration algorithm that we have developed constructs a quadrature rule over each finite element that satisfies a given error tolerance for all the above integrands. Adaptive integration schemes are normally recursive in nature and have a few common ingredients: a quadrature rule that can be applied to the integration domain to provide a local estimate of the integral; a procedure to estimate the local integration error; a strategy to partition the integration domain into smaller divi-



sions of the same shape; and a stopping criterion [98–100]. Our quadrature construction algorithm is specifically designed for high accuracy on parallelepipeds. We use a tensor-product of one-dimensional Gauss quadrature rules with five points in each direction to evaluate the local integrals. A tensor-product quadrature with eight points in each direction is used as the reference integral to estimate the local error. If the absolute error of integration is larger than the prescribed tolerance, the integration domain is uniformly divided into eight cells and the adaptive integration is performed over each cell recursively. This process is started with the six integrands of Figure 7, and at each step only those functions whose integration error is higher than the tolerance are passed to the next level; until all functions are accurately integrated.

Mesh	Error tolerance	Number of integration points		
		Pure FE	Tensor-product	Adaptive quadrature
$2 \times 2 \times 2$	7×10^{-3}	1000	512000	128750
$4 \times 4 \times 4$	2×10^{-3}	8000	169000	72750
$8 \times 8 \times 8$	2×10^{-4}	64000	624000	137500
$12\times12\times12$	2×10^{-5}	216000	1840000	297375
$16\times16\times16$	6×10^{-6}	512000	14020000	540000
$20\times 20\times 20$	4×10^{-6}	1000000	27196000	1171500
$24 \times 24 \times 24$	7×10^{-7}	1728000	46852000	1896000

Table 1: Comparing tensor-product and adaptive quadrature schemes for the diamond problem.

3.3.1 Analysis of adaptive scheme

In all-electron calculations, the charge density has a cusp at the atomic sites (nuclei). However, this does not have an adverse effect on the efficiency of our integration method, and a high rate of convergence is realized. The following example demonstrates the efficacy of the approach. Consider the integration of the function $f(\mathbf{x}) = 1 - r$, where r is the distance of the point \mathbf{x} to the origin. The domain is $\Omega = [-1, 1]^d \subset \mathbb{R}^d$. A plot of the function $f \in C^0(\Omega)$ in 2D is shown in Figure 8a. We use Gauss quadrature to integrate f, independent of the location of the cusp. Figures 8b and 8c depict the convergence of the integration scheme as the number of integration points is increased: a rate of convergence of d + 1 is observed in d-dimensions. If $f \in C^m(\Omega)$ ($m \ge 0$) and $\Omega \subset \mathbb{R}^d$, then the convergence rate is m + d + 1. This "blessing of dimensionality" makes our approach particularly efficient in three dimensions for integrands with strongly localized inhomogeneities (such as cusps).

3.3.2 All-electron diamond

We now revisit the all-electron problem for diamond, which was considered in Section 3.2. We apply the adaptive integration algorithm and results are compared to the tensor-product rule in Table 1. The error tolerance (the input of the quadrature construction algorithm) is the maximum allowable error that produces a stable result. Since cubic FE shape functions are used, a 5-point Gauss rule is employed in each direction. Adaptive integration proves to be markedly superior to tensor-product quadrature. The improvement is most evident for finer meshes where higher accuracies are required: the integration demand of the PUFE solution is only marginally greater than the pure FE solution on fine meshes. And the adaptive scheme achieves the required accuracy with an *order of magnitude* fewer total quadrature points than the previously implemented tensor-product scheme.

3.4 Parallel implementation

For an efficient parallel implementation, we assume the meshes are uniform and that the number of elements in each direction of the Poisson mesh is an integer multiple of the number in the Schrödinger mesh. Because the orbital/potential enrichments take up the rapid variations in the desired crystal wavefunctions/potentials, a uniform FE mesh is sufficient, as demonstrated in the preceding sections. Also, because the Poisson solution is such a small, and linearly scaling, part of the total self-consistent solution, increasing its number of elements as an integer multiple of the wavefunction mesh presents no problem. Finally, in our initial implementation, each partition has the same number of elements and nodes and, therefore, the number of elements in a given lattice direction will be an integer multiple of partition count in that direction.



Figure 9: Example 6x8 mesh.

A straightforward static scheduling is employed, where there is a one-to-one mapping between partitions and processes. This scheduling works well if all partitions have similar workloads. This may not always hold, however, since the execution time for applying the nonlocal projectors may vary among partitions. The main reason for this variation is the sparsity of the nonlocal projector vectors. The partitions with no atoms will tend to have sparser nonlocal projector vectors due to the locality of the projector functions (on the order of half nearest neighbor distance). Consequently, partitions with no atoms will have smaller execution times for nonlocal projector computations. Whereas this is not a significant problem for scaling, it does point to opportunities for further optimizations.

MPI is employed for the parallelization throughout. Synchronous (blocking) MPI subroutines are used. Subsequent implementations will exploit nonblocking MPI calls to overlap computations with communications.

3.4.1 Domain decomposition

The Schrödinger and Poisson meshes are partitioned so that the result vector found by matrixvector products has an equal number of elements in each partition. The partitioning of the mesh is illustrated on an example 6x8 mesh (Figure 9). Here, the gray nodes are image nodes. There are no additional degrees of freedom (DOFs) for image nodes due to the Bloch-periodic boundary conditions of problem.

The results from matrix-vector multiplications are stored based on the partitioning shown in Figure 10. As shown in the Figure, the vectors are partitioned in a nonoverlapping fashion. An MPI_Gather can be used to gather all results on a single processor. However, currently, the results are collected on the main processor (processor 0) by MPI_Send and MPI_Recv calls. The result vectors belonging to partitions (w1, w2, w3, and w4) and the global result vector w are also shown.

3.4.2 Parallel matrix-vector multiplication for $H_{\rm loc}$

The element-by-element (EBE) matrix-vector multiplication for the local part of the Hamiltonian $H_{\rm loc}$ is written as

$$w = \sum_{e} H^{e}_{\text{loc}} v^{e}, \tag{5}$$



Figure 10: 2x2 partitioning of the nodes and the corresponding result vectors w.



Figure 11: The elements and nodes used for the EBE computations on partition-1.

where H_{loc}^e are the local element matrices, and v^e are the local element vectors corresponding to element DOFs. In finite element computations, the entries of the global vectors v and w correspond to unique DOFs. To compute a result vector entry w_i , entries of v corresponding to the adjacency set of DOF i (iUAdj(i)) are required. Therefore, for a given partition, the size of vector v is larger than that of vector w since the vector v also includes the entries corresponding to the adjacency set of the DOFs within the partition. Similarly, for EBE computations, all element matrices connected to the DOFs within the partitions are used. Figure 11 shows the elements and nodes used for partition-1. Here, element matrices corresponding to yellow colored elements are used in EBE computations. The red colored nodes correspond to the components of the w_1 vector. The components of the v_1 vector correspond to red colored nodes plus green colored nodes. As shown in Figure 11, elements on opposing boundaries are also required for the EBE computations due to the Bloch-periodic boundary conditions of the problem.

Once the elements are assigned to processes (e.g., yellow elements for partition-1 in Figure 11), each process computes the element Overlap and Laplacian matrices corresponding to its elements. The element connectivity information must be modified according to the vectors v and w distributed among the processes. The element connectivity information is updated at the partitioning stage based on the distributed result w and operand v vectors.

The parallel matrix-vector multiplication is performed in three main steps:

- 1. Main process (rank 0) distributes the global operand vector v to all processes. The entries of distributed vectors v_d are duplicated for the nodes at the boundary of the partitions (green nodes in Figure 11). The rank 0 process sends v_d by calling MPI_Send and other processes make a corresponding MPI_Recv call to get v_d .
- 2. Each process computes the EBE matrix-vector product. The element DOF connectivities are modified according to the indices of w_d and v_d . The results are stored in distributed result vectors w_d . The entries of w are stored in a nonoverlapping fashion (red nodes in Figure 11).
- 3. The distributed result vectors w_d are gathered at the main process (rank 0). The global result vector w is composed of the w_d vectors. The results are gathered using MPI_Send and MPI_Recv calls. MPI_Gather might be used in future implementations. (See Figure 10 for gathering of w vectors.)

3.4.3 Parallel matrix-vector multiplication for H_{nonloc}

The matrix-vector computations for the nonlocal part of the potential (H_{nonloc}) are performed as the summation of dot-products:

$$w = \sum_{l} f_l h_l f_l^H \cdot v, \tag{6}$$

where f_l are nonlocal projectors and h_l are weights. There can be a large number of nonlocal projectors and the operation given in Eq. 6 consumes the largest share of execution time for sufficiently large problems. For parallel computations, the entries of f_l and v are distributed in a nonoverlapping fashion (Figure 10). Unlike the local operator computations, communication between processors is required during the nonlocal operator computations. The nonlocal computations are divided into two parts:

$$c_l = f_l^H \cdot v \tag{7}$$

$$w = \sum_{l} f_l h_l c_l \tag{8}$$



Figure 12: Partitioning of the elements for density computations on example 6x8 mesh.

The operation shown in Eq. 7 is a dot-product and requires communication between processors. First, each process computes the scalars c_l corresponding to the nodes within its partition (red nodes in Figure 11). To obtain a global c_l value, first, c_l values computed on individual partitions are summed. Then, the result is distributed to all processors. These two operations can be performed with a single MPI_AllReduce. After the global c_l values are distributed to processors, the remaining computations shown in Eq. 8 can be performed without further data transfers.

After the result vector w is found at each partition, the global result is gathered at the main processor. This gathering operation proceeds in the same way as for the H_{loc} computations (Figure 10).

3.4.4 Parallel density computations

The density computations are performed on the nonoverlapping subdomains of the Poisson mesh. Figure 12 shows the nonoverlapping subdomains for the example 6x8 mesh. Once the eigenfunctions and eigenvalues are calculated, the charge densities are computed at each Gauss point without any communication between processors. However, some communication between the processors is required for the numerical integration of the charge densities. The integration is required for normalization and it can be performed independently in each partition using the Gauss point charge density values within the elements of each partition. The integration can be written as the summation of integrals over each partition:

$$\int \rho \, dV = \sum_{p} \int \rho \, dV_p \tag{9}$$

$$I_p = \int \rho \, dV_p,\tag{10}$$

where ρ is the charge density at Gauss points, V is the unit cell volume, V_p is the partition volume, and I_p is the integral computed on partition p.

The I_p values computed on each partition are summed and the summation is sent to all processors. The two operations (summation + broadcast) can be performed with a single MPI_AllReduce. Once the integral is received by all processors, the charge densities are normalized at each processor independently.



Figure 13: Parallel speedup of 108 atom Al₃Cu calculation on up to 1728 processors of LLNL sierra.

The parallel computation of the charge density on the Schrödinger mesh is carried out in the same way as for the Poisson mesh.

3.4.5 Initial results

Figure 13 shows the parallel speedup of a 108 atom Al₃Cu calculation on up to 1728 processors of LLNL sierra. In this initial parallel implementation, only the most computationally intensive parts of the code were parallelized. As is clear from the Figure, the most computationally intensive part of the PUFE calculation, the construction of the electronic density, scales almost perfectly up to the full 1728 processors allocated, requiring just 8 sec to complete on 1728 processors. The construction of local matrices scales well also as shown just below the density, with some degradation due to additional gather-scatter operations. Less attention was given to this portion due to its smaller computation time: as it is, requiring less than 1 sec on 1728 processors. The sparse matrix-vector multiplies, however, did not scale well in the initial implementation. Subsequent analysis showed that this was a result of excessive gather-scatter operations in the implementation. The vast majority of these, however, can be removed by keeping the relevant data structures distributed rather than gathering on every processor before every operation. These optimizations are underway presently.

4 Outlook

In this work, we have shown for the first time that a local, systematically improvable, real-space basis can attain the accuracies required in quantum mechanical materials calculations with not only fewer but *substantially fewer* degrees of freedom than current state-of-the-art planewave based methods, as implemented in VASP, ABINIT, QBox, and a host of other codes the world over. In demonstrating this, the work has attracted increasing attention as it has progressed. Over the course of the project, sixteen external invited talks were presented, including three international; two papers [60, 61] have been published and another [62] is in preparation; and multiple external collaborations have been forged, resulting in multiple external proposals funded. Prof. N. Sukumar at UC Davis was awarded an NSF-DMS grant to work on quadrature and polynomial order/continuity issues for the PUFE Kohn-Sham method. Profs. Zhaojun Bai and N. Sukumar at UC Davis were awarded a UC-Lab Fee Research grant to develop efficient parallel eigensolvers for the PUFE method. Dr. Kristopher Andersen at the Naval Research Laboratory was awarded an HPTi PETTT grant to develop a high-order spectral-element projector-augmented-wave alternative to the PUFE method, leveraging much of the expertise, code, and collaborations developed in the present project. Most recently, we have begun a collaboration with the quantum chemistry group of Prof. T. J. Martinez at Stanford on the application of the PUFE methodology developed here to molecular, cluster, and nanostructural materials; and an associated NSF proposal has been submitted. In addition to attracting three postdocs and two graduate students based at UC Davis, the work has attracted students to the Laboratory to work with the PI, Dr. Pask. Graduate student, Ondrej Certik, came to the Lab to work on high-order radial solvers, used to construct enrichment functions for the larger PUFE electronic structure method, as a participant in the 2010 Computational Chemistry and Materials Science Summer Institute. Mr. Certik has now returned to the Lab to work with Dr. Pask on his Thesis. Most recently, two more graduate students, Zhenfei Liu at UC Irvine and Cindy Wang at Cal Tech, have applied to the 2011 Computational Chemistry and Materials Science Summer Institute to work with Dr. Pask on related research.

Given the substantial external interest, multiple collaborations developed, and associated funding now secured or submitted, the work begun here continues actively on several fronts. Having demonstrated the proof of principle, work now centers on the necessary extensions and optimizations required to bring the prototype method and code delivered here to production applications, with an eye toward Stockpile Stewardship applications, in particular.

Acknowledgments This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

References

- W. E. Pickett. Pseudopotential methods in condensed matter applications. Computer Physics Reports, 9(3):115–197, 1989.
- [2] T. L. Beck. Real-space mesh techniques in density-functional theory. *Rev. Mod. Phys.*, 72(4):1041–1080, 2000.
- [3] T. Torsti, T. Eirola, J. Enkovaara, T. Hakala, P. Havu, V. Havu, T. Höynälänmaa, J. Ignatius, M. Lyly, I. Makkonen, T. T. Rantala, J. Ruokolainen, K. Ruotsalainen, E. Räsänen, H. Saarikoski, and M. J. Puska. Three real-space discretization techniques in electronic structure calculations. *phys. stat. sol.* (b), 243(5):1016–1053, 2006.
- [4] J. E. Pask and P. A. Sterne. Finite element methods in ab initio electronic structure calculations. Model. Simul. Mater. Sci. Eng., 13(3):R71–R96, 2005.
- [5] G. Strang and G. J. Fix. An Analysis of the Finite Element Method. Prentice-Hall, Englewood Cliffs, 1973.
- [6] D. Vanderbilt. Soft self-consistent pseudopotentials in a generalized eigenvalue formalism. *Phys. Rev. B*, 41(11):7892–7895, 1990.
- [7] A. M. Rappe, K. M. Rabe, E. Kaxiras, and J. D. Joannopoulos. Optimized pseudopotentials. *Phys. Rev. B*, 41(2):1227–1230, 1990.

- [8] J. S. Lin, A. Qteish, M. C. Payne, and V. Heine. Optimized and transferable nonlocal separable abinitio pseudopotentials. *Phys. Rev. B*, 47(8):4174–4180, 1993.
- [9] F. Gygi. Adaptive Riemannian metric for plane-wave electronic-structure calculations. *Europhys. Lett.*, 19(7):617–622, 1992.
- [10] A. Devenyi, K. Cho, T. A. Arias, and J. D. Joannopoulos. Adaptive Riemannian metric for all-electron calculations. *Phys. Rev. B*, 49(19):13373–13376, 1994.
- [11] D. R. Hamann. Band-structure in adaptive curvilinear coordinates. Phys. Rev. B, 51(15):9508-9514, 1995.
- [12] P. E. Blöchl. Projector augmented-wave method. Phys. Rev. B, 50(24):17953–17979, 1994.
- [13] S. Goedecker, M. Boulet, and T. Deutsch. An efficient 3-dim FFT for plane wave electronic structure calculations on massively parallel machines composed of multiprocessor nodes. *Comput. Phys. Commun.*, 154(2):105–110, 2003.
- [14] A. Canning and D. Raczkowski. Scaling first-principles plane-wave codes to thousands of processors. Comput. Phys. Commun., 169(1-3):449–453, 2005.
- [15] C. K. Skylaris, P. D. Haynes, A. A. Mostofi, and M. C. Payne. Introducing ONETEP: Linearscaling density functional simulations on parallel computers. J. Chem. Phys., 122(8):084119, 2005.
- [16] F. Gygi, E. W. Draeger, B. R. De Supinski, R. K. Yates, F. Franchetti, S. Kral, J. Lorenz, C. W. Ueberhuber, J. A. Gunnels, and J. C. Sexton. Large-scale first-principles molecular dynamics simulations on the BlueGene/L platform using the Qbox code. In *Proceedings of Supercomputing 2005*, page 24, 2005.
- [17] J. R. Chelikowsky, N. Troullier, and Y. Saad. Finite-difference-pseudopotential method: Electronic-structure calculations without a basis. *Phys. Rev. Lett.*, 72(8):1240–1243, 1994.
- [18] M. M. G. Alemany, M. Jain, L. Kronik, and J. R. Chelikowsky. Real-space pseudopotential method for computing the electronic properties of periodic systems. *Phys. Rev. B*, 69(7):075101, 2004.
- [19] A. P. Seitsonen, M. J. Puska, and R. M. Nieminen. Real-space electronic-structure calculations: Combination of the finite-difference and conjugate-gradient methods. *Phys. Rev. B*, 51(20):14057–14061, 1995.
- [20] F. Gygi and G. Galli. Real-space adaptive-coordinate electronic-structure calculations. Phys. Rev. B, 52(4):R2229–R2232, 1995.
- [21] K. A. Iyer, M. P. Merrick, and T. L. Beck. Application of a distributed nucleus approximation in grid based minimization of the Kohn-Sham energy functional. J. Chem. Phys., 103(1):227– 233, 1995.
- [22] T. Hoshi, M. Arai, and T. Fujiwara. Density-functional molecular-dynamics with real-space finite-difference. *Phys. Rev. B*, 52(8):R5459–R5462, 1995.
- [23] E. L. Briggs, D. J. Sullivan, and J. Bernholc. Real-space multigrid-based approach to largescale electronic structure calculations. *Phys. Rev. B*, 54(20):14362–14375, 1996.

- [24] N. A. Modine, G. Zumbach, and E. Kaxiras. Adaptive-coordinate real-space electronicstructure calculations for atoms, molecules, and solids. *Phys. Rev. B*, 55(16):10289–10301, 1997.
- [25] J. L. Fattebert. Finite difference schemes and block Rayleigh quotient iteration for electronic structure calculations on composite grids. J. Comput. Phys., 149(1):75–94, 1999.
- [26] J.-L. Fattebert and J. Bernholc. Towards grid-based O(N) density-functional theory methods: optimized non-orthogonal orbitals and multigrid acceleration. *Phys. Rev. B*, 62(3):1713–1722, 2000.
- [27] J.-L. Fattebert and F. Gygi. Linear scaling first-principles molecular dynamics with controlled accuracy. Comput. Phys. Comm., 162:24–36, 2004.
- [28] T. Ono and K. Hirose. Timesaving double-grid method for real-space electronic-structure calculations. *Phys. Rev. Lett.*, 82(25):5016–5019, 1999.
- [29] A. V. Knyazev. Toward the optimal preconditioned eigensolver: Locally optimal block preconditioned conjugate gradient method. Siam J. Sci. Comput., 23(2):517–541, 2001.
- [30] P. Arbenz, U. L. Hetmaniuk, R. B. Lehoucq, and R. S. Tuminaro. A comparison of eigensolvers for large-scale 3D modal analysis using AMG-preconditioned iterative methods. *Int. J. Numer. Meth. Eng.*, 64(2):204–236, 2005.
- [31] A. Askar. Finite-element method for bound-state calculations in quantum-mechanics. J. Chem. Phys., 62(2):732–734, 1975.
- [32] L. R. Ram-Mohan. Finite Element and Boundary Element Applications in Quantum Mechanics. Oxford University Press, New York, 2002.
- [33] S. R. White, J. W. Wilkins, and M. P. Teter. Finite-element method for electronic-structure. *Phys. Rev. B*, 39(9):5819–5833, 1989.
- [34] B. Hermansson and D. Yevick. Finite-element approach to band-structure analysis. *Phys. Rev. B*, 33(10):7241–7242, 1986.
- [35] E. Tsuchida and M. Tsukada. Electronic-structure calculations based on the finite-element method. Phys. Rev. B, 52(8):5573–5578, 1995.
- [36] E. Tsuchida and M. Tsukada. Adaptive finite-element method for electronic-structure calculations. Phys. Rev. B, 54(11):7602–7605, 1996.
- [37] E. Tsuchida and M. Tsukada. Large-scale electronic-structure calculations based on the adaptive finite-element method. J. Phys. Soc. Jpn., 67(11):3844–3858, 1998.
- [38] E. Tsuchida, Y. Kanada, and M. Tsukada. Density-functional study of liquid methanol. Chem. Phys. Lett., 311(3-4):236-240, 1999.
- [39] E. Tsuchida. Ab initio molecular-dynamics study of liquid formamide. J. Chem. Phys., 121(10):4740–4746, 2004.
- [40] J. E. Pask, B. M. Klein, C. Y. Fong, and P. A. Sterne. Real-space local polynomial basis for solid-state electronic-structure calculations: A finite-element approach. *Phys. Rev. B*, 59(19):12352–12358, 1999.

- [41] P. A. Sterne, J. E. Pask, and B. M. Klein. Calculation of positron observables using a finite element-based approach. Appl. Surf. Sci., 149(1-4):238-243, 1999.
- [42] J. E. Pask, B. M. Klein, P. A. Sterne, and C. Y. Fong. Finite-element methods in electronicstructure theory. *Comput. Phys. Commun.*, 135(1):1–34, 2001.
- [43] J. E. Pask and P. A. Sterne. Finite elements in ab initio electronic-structure calculations. In S. Yip, editor, *Handbook of Materials Modeling*, volume 1, pages 423–437. Springer, Dordrecht, 2005.
- [44] J. E. Pask and P. A. Sterne. Real-space formulation of the electrostatic potential and total energy of solids. *Phys. Rev. B*, 71(11):113101, 2005.
- [45] J. E. Pask and P. A. Sterne. Ab initio electronic structure calculations of metals by the finite element method. Bull. Am. Phys. Soc., 50:1107, 2005.
- [46] PF Batcho. Computational method for general multicenter electronic structure calculations. *Phys. Rev. E*, 61(6, Part B):7169–7183, 2000.
- [47] Eric J. Bylaska, Michael Holst, and John H. Weare. Adaptive Finite Element Method for Solving the Exact Kohn-Sham Equation of Density Functional Theory. J. Chem. Theory Comput., 5(4):937–948, 2009.
- [48] Lauri Lehtovaara, Ville Havu, and Martti Puska. All-electron density functional theory and time-dependent density functional theory with high-order finite elements. J. Chem. Phys., 131(5):054103, 2009.
- [49] H. L. Skriver. The LMTO Method. Springer, Berlin, 1984.
- [50] D. J. Singh and L. Nordstrom. Planewaves, Pseudopotentials, and the LAPW Method. Springer, New York, 2nd edition, 2006.
- [51] S. F. Boys. Electronic wave functions. I. A general method of calculation for the stationary states of any molecular system. Proc. R. Soc. Lon. A, 200(1063):542–554, 1950.
- [52] W. J. Hehre, R. F. Stewart, and J. A. Pople. Self-consistent molecular-orbital methods. 1. Use of Gaussian expansion of Slater-type atomic orbitals. J. Chem. Phys., 51(6):2657, 1989.
- [53] C. Dusterhoft, D. Heinemann, and D. Kolb. Dirac-Fock-Slater calculations for diatomic molecules with a finite element defect correction method (FEM-DKM). *Chem. Phys. Lett.*, 296(1-2):77–83, 1998.
- [54] S. Yamakawa and S. Hyodo. Electronic state calculation of hydrogen in metal clusters based on Gaussian-FEM mixed basis function. J. Alloy. Compd., 356(2):231–235, 2003.
- [55] S. Yamakawa and S. Hyodo. Gaussian finite-element mixed-basis method for electronic structure calculations. *Phys. Rev. B*, 71(3):035113, 2005.
- [56] J. M. Melenk and I. Babuška. The partition of unity finite element method: Basic theory and applications. Computer Methods in Applied Mechanics and Engineering, 139:289–314, 1996.
- [57] I. Babuška and J. M. Melenk. The partition of unity method. International Journal for Numerical Methods in Engineering, 40:727–758, 1997.

- [58] C. A. Duarte and J. T. Oden. An h-p adaptive method using clouds. Computer Methods in Applied Mechanics and Engineering, 139:237–262, 1996.
- [59] I. Babuška, U. Banerjee, and J. E. Osborn. Survey of meshless and generalized finite element methods: A unified approach. Acta Numerica, 12:1–125, 2003.
- [60] N. Sukumar and J. E. Pask. Classical and enriched finite element formulations for blochperiodic boundary conditions. Int. J. Numer. Meth. Eng., 77(8):1121–1138, 2009.
- [61] J. E. Pask, N. Sukumar, and S. E. Mousavi. Linear scaling solution of the all-electron Coulomb problem in solids. Int. J. Multiscale Computational Engineering, 2011. in press. Available at http://arXiv.:1004.1765.
- [62] J. E. Pask, N. Sukumar, M. E. Guney, and W. Hu. Partition of unity finite elements for electronic structure calculations. *Phys. Rev. Lett.*, 2011. in preparation.
- [63] T. H. Dunning Jr. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. J. Chem. Phys., 90:1007, 1989.
- [64] T. Strouboulis, I. Babuška, and K. Copps. The design and analysis of the generalized finite element method. *Computer Methods in Applied Mechanics and Engineering*, 181(1–3):43–69, 2000.
- [65] T. Strouboulis, K. Copps, and I. Babuška. The generalized finite element method : An example of its implementation and illustration of its performance. *International Journal for Numerical Methods in Engineering*, 47(8):1401–1417, 2000.
- [66] T. Strouboulis, K. Copps, and I. Babuška. The generalized finite element method. Computer Methods in Applied Mechanics and Engineering, 190(32–33):4081–4193, 2001.
- [67] T. Belytschko and T. Black. Elastic crack growth in finite elements with minimal remeshing. International Journal for Numerical Methods in Engineering, 45(5):601–620, 1999.
- [68] N. Moës, J. Dolbow, and T. Belytschko. A finite element method for crack growth without remeshing. International Journal for Numerical Methods in Engineering, 46(1):131–150, 1999.
- [69] J. Dolbow, N. Moës, and T. Belytschko. Discontinuous enrichment in finite elements with a partition of unity method. *Finite Elements in Analysis and Design*, 36(3–4):235–260, 2000.
- [70] C. Daux, N. Moës, J. Dolbow, N. Sukumar, and T. Belytschko. Arbitrary cracks and holes with the extended finite element method. *International Journal for Numerical Methods in Engineering*, 48(12):1741–1760, 2000.
- [71] N. Sukumar, N. Moës, B. Moran, and T. Belytschko. Extended finite element method for three-dimensional crack modelling. *International Journal for Numerical Methods in Engineering*, 48(11):1549–1570, 2000.
- [72] N. Sukumar, D. L. Chopp, N. Moës, and T. Belytschko. Modeling holes and inclusions by level sets in the extended finite-element method. *Computer Methods in Applied Mechanics* and Engineering, 190(46–47):6183–6200, 2001.
- [73] N. Sukumar, D. L. Chopp, and B. Moran. Extended finite element method and fast marching method for three dimensional fatigue crack propagation. *Engineering Fracture Mechanics*, 70(1):29–48, 2003.

- [74] D. L. Chopp and N. Sukumar. Fatigue crack propagation of multiple coplanar cracks with the coupled extended finite element/fast marching method. *International Journal of Engineering Science*, 41(8):845–869, 2003.
- [75] N. Sukumar, D. L. Chopp, E. Béchet, and N. Moës. Three-dimensional non-planar crack growth by a coupled extended finite element and fast marching method. *International Journal* for Numerical Methods in Engineering, 76(5):727–748, 2008.
- [76] P. Ortiz and E. Sanchez. An improved partition of unity finite element model for diffraction problems. International Journal for Numerical Methods in Engineering, 50(12):2727–2740, 2001.
- [77] X. K. Li and H. Y. Zhou. Partition of unity finite element method for short wave propagation in solids. Applied Mathematics and Mechanics, 26(8):1056–1063, 2005.
- [78] P. Gamallo and R. J. Astley. The partition of unity finite element method for short wave acoustic propagation on non-uniform potential flows. *International Journal for Numerical Methods in Engineering*, 65(3):425–444, 2006.
- [79] H. Ji, D. Chopp, and J. E. Dolbow. A hybrid extended finite element / level set method for modeling phase transformations. *International Journal for Numerical Methods in Engineer*ing, 54(8):1209–1233, 2002.
- [80] J. Fish and Z. Yuan. Multiscale enrichment based on partition of unity. International Journal for Numerical Methods in Engineering, 62(10):1341–1359, 2005.
- [81] W. Aquino, J. C. Brigham, C. J. Earls, and N. Sukumar. Generalized finite element method using proper orthogonal decomposition. *International Journal for Numerical Methods in Engineering*, 2009. in press.
- [82] T. Belytschko, C. Parimi, N. Möes, N. Sukumar, and S. Usui. Structured extended finite element methods for solids defined by implicit surfaces. *International Journal for Numerical Methods in Engineering*, 56(4):609–635, 2003.
- [83] M. Griebel and M. A. Schweitzer. A particle-partition of unity method for the solution of elliptic, parabolic, and hyperbolic PDEs. SIAM Journal of Scientific Computing, 22(3):853– 890, 2000.
- [84] C. Hartwigsen, S. Goedecker, and J. Hutter. Relativistic separable dual-space Gaussian pseudopotentials from H to Rn. Phys. Rev. B, 58(7):3641–3662, 1998.
- [85] J.-L. Fattebert. *MGmol, a real-space multigrid first-principles molecular dynamics code.* CASC, Lawrence Livermore National Laboratory, 2001-2006.
- [86] J. J. Mortensen, L. B. Hansen, and K. W. Jacobsen. Real-space grid implementation of the projector augmented wave method. *Phys. Rev. B*, 71(3):035109, 2005.
- [87] P. P. Ewald. Die Berechnung optischer und electrostatischer Gitterpotentiale. Ann. der Physik, 64:253, 1921.
- [88] T. A. Arias. Multiresolution analysis of electronic structure: Semicardinal and wavelet bases. *Rev. Mod. Phys.*, 71(1):267–311, 1999.

- [89] J. M. Soler, E. Artacho, J. D. Gale, A. Garcia, J. Junquera, P. Ordejon, and D. Sanchez-Portal. The SIESTA method for ab initio order-N materials simulation. J. Phys.-Condes. Matter, 14(11):2745–2779, 2002.
- [90] S. Goedecker and O. V. Ivanov. Linear scaling solution of the Coulomb problem using wavelets. Solid State Commun., 105(11):665–669, 1998.
- [91] H. Murakami, V. Sonnad, and E. Clementi. A 3-dimensional finite-element approach towards molecular SCF computations. Int. J. Quantum Chem., 42(4):785–817, 1992.
- [92] E. Tsuchida and M. Tsukada. Real-space approach to electronic-structure calculations. Solid State Commun., 94(1):5–8, 1995.
- [93] H. H. H. Homeier and E. O. Steinborn. Numerical integration of functions with a sharp peak at or near one boundary using Möbius transformation. *Journal of Computational Physics*, 87:61–72, 1990.
- [94] H. H. H. Homeier and E. O. Steinborn. On the evaluation of overlap integrals with exponential-type basis functions. *International Journal Of Quantum Chemistry*, 42:761–778, 1992.
- [95] R. López and G. Ramírez. Calculation of two-center exchange integrals with STOs using Möbius transformations. International Journal Of Quantum Chemistry, 49:11–19, 1994.
- [96] M. G. Duffy. Quadrature over a pyramid or cube of integrands with a singularity at a vertex. SIAM Journal on Numerical Analysis, 19(6):1260–1262, 1982.
- [97] S. E. Mousavi and N. Sukumar. Generalized Duffy transformation for integrating vertex singularities. *Computational Mechanics*, 45(2–3):127–140, 2010.
- [98] J. Berntsen and T. O. Espelid. Algorithm-706 DCUTRI: An algorithm for adaptive cubature over a collection of triangles. ACM Transactions on Mathematical Software, 18(3):329–342, 1992.
- [99] J. Berntsen, R. Cools, and T. O. Espelid. Algorithm 720 An algorithm for adaptive cubature over a collection of 3-dimensional simplices. ACM Transactions on Mathematical Software, 19(3):320–332, 1993.
- [100] W. Gander and W. Gautschi. Adaptive quadrature Revisited. BIT Numerical Mathematics, 40(1):84–101, 2000.