Preprint ANL/MCS-P1384-1206

QUASICONTINUUM-LIKE REDUCTION OF DFT CALCULATIONS OF NANOSTRUCTURES

Dan Negrut*
Materials Science Program
Department of Mechanical Engineering
The University of Wisconsin
Madison, WI 53706
negrut@wisc.edu

Mihai Anitescu
Mathematics and Computer Science Division
Argonne National Laboratory
Argonne, IL 60439
anitescu@mcs.anl.gov

Anter El-Azab
Materials Theory Group
College of Engineering
Florida State University
Tallahassee, FL 32310
anter@eng.fsu.edu

Peter Zapol
Materials Science and Chemistry Divisions
Argonne National Laboratory
Argonne, IL 60439
zapol@anl.gov

December 21, 2006

Abstract

Density functional theory (DFT) can accurately predict chemical and mechanical properties of nanostructures, although at a high computational cost. A quasicontinuum-like framework is proposed to substantially increase the size of the nanostructures that can be simulated ab initio. This increase stems from two facts. First, in order to find approximate ground state electron density (the electronic problem*), expensive DFT calculations are limited to a small number of subdomains, and the solution is interpolated everywhere else. The* electronic problem *embeds interpolation and coupled cross-domain optimization techniques through a process called electronic reconstruction. Second, for the optimization of nuclei positions (the* ionic problem*), computational gains result from explicit consideration of a reduced number of representative nuclei interpolating the positions of the rest of nuclei following the quasicontinuum paradigm.*

In the proposed approach, the new ionic configuration is the solution of a nonlinear system obtained as the first-order optimality condition for the minimization of the total energy associated with the nanostructure. This is an optimization problem with equilibrium constraints because the electronic density is itself the solution of a minimization problem. Numerical tests using the Thomas-Fermi-Dirac functional demonstrate the validity of the proposed framework within the orbital-free density functional theory.

INTRODUCTION

Nanostructures have dimensions in the range of $1 \sim 100$ nm and typically contain $10^2 \sim 10^8$ atoms. Density functional methods within the Kohn-Sham approach[1] are typically applied to systems with fewer than 100 atoms. Contemporary implementations of order-N methods[2] such as SIESTA[3], ONETEP[4], and CONQUEST[5] that exhibit linear scaling of computation time with system size enable an increase in the number of atoms by one to two orders

*Address all correspondence to this author.

of magnitude on massively parallel computers. Larger system sizes are accessible to classical interatomic potential methods but these methods cannot account for spin and charge relaxation and conjugation effects, which are important in modeling reactions, electronic excitation, and bond breaking processes. Therefore, new computational paradigms are needed that enable larger scale electronic structure calculations.

A combination of methods with different fidelity is often used to reduce computational effort if only local information is needed with high accuracy. An example of such an approach is the ONIOM method[6] for computations of chemical properties. However, such schemes have an inherent problem with conditions at the boundaries of different fidelity regions. Another approach to reduce computational effort, called the quasicontinuum method, is based on explicit treatment of only representative atoms and on interpolations for the rest. It has been successfully used in atomistic studies of mechanical properties with classical potentials and has been under development to include electronic-level calculations. This type of approach is particularly suitable for many nanostructures because large regions of the structures are perturbed relatively little as compared to periodic structures and, therefore, can be treated by using interpolation schemes.

The present work proposes a quasicontinuum-like technique that, by closing the spatial scale gap, renders electronic structure information at the nanoscale. The proposed methodology follows in the steps of the quasicontinuum approach discussed in[7;8;9] for mechanical analysis at the nanoscale. Specifically, this is an extension of the work in[8;9], because, rather than considering a potential-based interatomic interaction that has a limited range of validity and is difficult to generalize to inhomogeneous materials, the methodology proposed uses *ab initio* methods to provide for particle interaction. At the same time it is a generalization of the method proposed in[7] because, rather than considering electron density within each mesh discretization element separately, the proposed method treats the electronic density distribution in all elements in a generic optimization framework.

Our approach does not rely on a strict periodicity assumption; it merely assumes that the material displays a nearly periodic structure in certain regions of the nanostructure. However, in order to bridge the gap between subatomic scale associated with the electron density and the nanoscale associated with the structures investigated, we have assumed that almost everywhere in the nanostructure the optimized structure results in only small deformations of periodic structure. This assumption is referred to as *near-periodicity*, because the nonperiodic part of the state variables is approximated as a macroscopic smoothly varying field. As explained later, the *near-periodicity* assumption enables the use of interpolation for electronic structure reconstruction.

Here we define the *small deformations* that are present in most of the material. The nanostructure is considered to have an initial reference configuration $D^0 \subset \mathbb{R}^3$. The structure undergoes a deformation described by a deformation mapping $\Phi(\mathbf{r}^0, t) \in \mathbb{R}^3$, which gives the location $\mathbf{r}$ in the global Cartesian reference frame of each point $\mathbf{r}^0$ represented in the undeformed material frame. As indicated, the mapping might depend on time t. The variable t does not necessarily represent the time contemporary with the structure under consideration. In a static simulation framework this variable might be an iteration index of an optimization algorithm that solves for the system ground state.

The components of the deformation gradient are introduced as

$$F_{iJ} = \frac{\partial \Phi_i}{\partial r_J^0} \quad ,$$

where upper-case indices refer to the material frame, and lower-case indices to the Cartesian global frame. Thus, $\mathbf{F} = \nabla_0 \ \Phi$, where ∇_0 represents the material gradient operator. Using the repeated index summation rule, we express the deformation of an infinitesimal material neighborhood $\mathrm{d}\mathbf{r}^0$ about a point $\mathbf{r}^0$ of D^0 as

$$\mathrm{d}r_i = F_{iJ} \ \mathrm{d}r_J^0 \quad .$$

If $\mathbf{u} = \mathbf{r} - \mathbf{r}_0$, the concept of small distortion is equivalent to requiring that the spectral radius of $\bar{\mathbf{F}} = \nabla \ \mathbf{u}$ be sufficiently small; that is,

$$||\nabla_0 \ \mathbf{u}||_2 < \mathcal{K}$$

is expected to hold almost everywhere in the domain D^0, for a suitable chosen value of $\mathcal{K}$.

With the two assumptions introduced, we anticipate computational savings due to a two-tier interpolation-based approach that will reduce the dimension of the problem. First, the electronic structure will be evaluated in some domains by interpolation using adjacent regions in which a DFT-based approach is used to accurately solve the electronic structure problem; this procedure is called *electronic density reconstruction*. Second, the positions of the nanostructure nuclei will be expressed in terms of the positions of a reduced set of so-called representative nuclei, *repnuclei*. The proposed approach solves only for the positions of these *repnuclei*; the entire deformation field (mapping) is defined based on an appropriate representation for $\Phi(\mathbf{r}^0, t)$. Following the quasicontinuum paradigm[8;9], we define this mapping based on the displacement of *repnuclei*:

$$\Phi(\mathbf{r}^0, t) = \sum_{J \in \mathcal{B}} \varphi(\mathbf{r}^0 | \mathbf{R}_J^0)\ \Phi(\mathbf{R}_J^0, t), \tag{1}$$

where $\mathcal{B}$ represents the index set associated with the *repnuclei*. Once the displacements of the *repnuclei* $\Phi(\mathbf{R}_J^0, t)$ for $J \in \mathcal{B}$ are available, the displacement of any point in the nanostructure is obtained by interpolation using the shape functions $\varphi(\ \cdot\ |\ \cdot\ )$.

THE QUASICONTINUUM APPROACH FRAMEWORK

Finding the stable configuration of a nanostructure (called hereafter the *Ionic Problem*) reduces to minimizing the total energy E_{tot} with respect to the positions of the nuclei. More precisely, the equilibrium configuration of a nanostructure is provided by that distribution of the nuclei that minimizes the energy

$$E_{tot} = E_e + E_{ext} + E_{nn},$$

where E_{nn} is the internuclear interaction energy, E_e is the electronic ground-state energy for the corresponding nuclear distribution, and E_{ext} is the electron-nuclei interaction energy defined as

$$E_{ext}(\hat{\rho}(\mathbf{r}); \{\mathbf{R}_I\}) \quad = \quad -\int \hat{\rho}(\mathbf{r})\, V_{ext}\,(\mathbf{r}; \{\mathbf{R}_I\})\, d\mathbf{r} = -\sum_{A=1}^{M} \int \frac{Z_A \hat{\rho}(\mathbf{r})}{||\mathbf{r} - \mathbf{R}_A||}\, d\mathbf{r},$$

where $V_{ext}(\mathbf{r}; \{\mathbf{R}_I\}) = \sum\limits_{A=1}^{M} \frac{Z_A \hat{\rho}(\mathbf{r})}{||\mathbf{r}-\mathbf{R}_A||}\, d\mathbf{r}$ is the ionic potential, which depends on the positions of the nuclei $\{\mathbf{R}_I\}$.

The electronic structure computation is approached here as the solution of a constrained minimization problem according to the Hohenberg-Kohn theorem[10], $\min_\rho E_{tot}[\rho; \{\mathbf{R}_I\}]$, subject to $\int \rho(\mathbf{r}) \mathrm{d}\mathbf{r} = N_e$, where N_e represents the number of electrons present in the system. The solution to this problem depends parametrically on the locations of the nuclei $\mathbf{R}_I$, $I = 1, \ldots, M$, a consequence of the Born-Oppenheimer approximation. Subsequently, the optimization of nuclei positions in the entire system is the solution of $\min_{\{\mathbf{R}_I\}} E[\{\mathbf{R}_I\}; \rho]$.

Let us consider the optimization problem

$$\min_{\{\mathbf{R}_I\}} E_{tot} = E_e + E_{ext} + E_{nn}$$

subject to the constraint that for a nuclear configuration $\{\mathbf{R}_I\}$ the energy E_e is the electron ground-state energy. Under this assumption, the first-order optimality conditions yield

$$\mathbf{F}_K = \frac{\partial E_{ext}}{\partial \mathbf{R}_K} + \frac{\partial E_{nn}}{\partial \mathbf{R}_K} = \mathbf{0},$$

where $\mathbf{F}_K$, $K = 1, \ldots, M$, is interpreted as the force acting on nucleus K and

$$E_{nn} \quad = \quad \sum_{A=1}^{M} \quad \sum_{B=A+1}^{M} \frac{Z_A Z_B}{||\mathbf{R}_B - \mathbf{R}_A||}.$$

This result is proved in the Appendix, and for each nucleus K leads to

$$\mathbf{F}_K = \int \hat{\rho}(\mathbf{r}) \frac{\mathbf{r} - \mathbf{R}_K}{||\mathbf{r} - \mathbf{R}_K||^3} \, \mathrm{d}\mathbf{r} + \sum_{A=1, A \neq K}^{M} Z_A \frac{\mathbf{R}_A - \mathbf{R}_K}{||\mathbf{R}_A - \mathbf{R}_K||^3} = \mathbf{0} \ ,$$

which thus allows one to solve the nuclear equilibrium problem by using only the solution of the electron density problem and not the values and the derivatives of the kinetic and exchange energy functionals. Once the electron density is available, the equilibrium conditions $\mathbf{F}_K = \mathbf{0}$, $K = 1, \ldots, M$, can be imposed right away. The major computational consequence of this result is that *how the actual ground state electron density $\hat{\rho}(\mathbf{r})$ was obtained is irrelevant*; there is also no need to have an explicit energy functional of the electron density ρ. Moreover, the gradient of the energy with respect to the atomic positions at the current electron density has the same property[11]. Therefore the electron density can be computed with a stand-alone software package that requires only the current atomic positions.

When a local quasicontinuum approach is used, the equilibrium conditions are imposed only for *repnuclei*, that is, only for $J \in \mathcal{B}$. The positions of the remaining atoms in the system is then expressed in terms of the positions of the *repnuclei*. The *repnuclei* become the nodes of an atomic mesh, and interpolation is used to recover the position of the remaining nuclei. Concretely, if the atomic mesh is denoted by $\mathcal{M}$, τ is an arbitrary cell in this mesh, $\mathcal{V}(\tau)$ represents the set of nodes associated with cell τ, and φ_L is the shape function associated with node L in cell τ,

$$\mathbf{F}_J = \int \hat{\rho}(\mathbf{r}) \frac{\mathbf{r} - \mathbf{R}_J}{||\mathbf{r} - \mathbf{R}_J||^3} \, \mathrm{d}\mathbf{r} + \sum_{\tau \in \mathcal{M}} \sum_{A \in \tau} Z_A \frac{\sum\limits_{L \in \mathcal{V}(\tau)} \mathbf{R}_L \varphi_L(\mathbf{R}_A) - \mathbf{R}_J}{||\sum\limits_{L \in \mathcal{V}(\tau)} \mathbf{R}_L \varphi_L(\mathbf{R}_A) - \mathbf{R}_J||^3} = \mathbf{0} \ , \quad J \in \mathcal{B} \ . \tag{2}$$

This effectively reduces the dimension of the problem from $3\,M$ (the (x, y, z) coordinates of the nuclei), to $3\,M_{rep}$, where M_{rep} is the number of nodes in the atomic mesh, which is the number of elements in $\mathcal{B}$. The sum in the expression of $\mathbf{F}_J$ above is most likely not going to be the simulation bottleneck (solving the electron problem for $\hat{\rho}$ is significantly more demanding), but fast-multipole methods[12;13;14] can be considered to speed the summation.

Denoting by $\mathbf{P}_i$, $i = 1, \ldots, M_{rep}$, the position of the representative nucleus n_i, we can group the set of nonlinear equations of Eq. (2) into a nonlinear system that is solved for the relaxed configuration of the structure:

$$\begin{aligned}
&\mathbf{f}_1(\mathbf{P}_1, \mathbf{P}_2, \ldots \mathbf{P}_{M_{rep}}) = \mathbf{0} \\
&\mathbf{f}_2(\mathbf{P}_1, \mathbf{P}_2, \ldots \mathbf{P}_{M_{rep}}) = \mathbf{0} \\
&\ldots \\
&\mathbf{f}_{M_{rep}}(\mathbf{P}_1, \mathbf{P}_2, \ldots \mathbf{P}_{M_{rep}}) = \mathbf{0}
\end{aligned}$$

where $\mathbf{f}_J$ is obtained based on Eq. (2). The solution of this system is found by a Newton-like method. Evaluating the Jacobian information is straightforward and not detailed here. We note that in Eq. (2) a connection is made back to Eq. (1); the position of an arbitrary nucleus A in cell τ is computed based on interpolation using the nodes $\mathcal{V}(\tau)$, one of many choices available (one could consider *repnuclei* from neighboring cells, for instance). Effectively, this provides in Eq. (1) an expression for $\Phi(\cdot, t)$, which depends only on $J \in \mathcal{V}(\tau)$ rather than $J \in \mathcal{B}$.

THE ELECTRONIC PROBLEM

The *electronic problem* refers to the computation of the ground-state electron density given the positions of nuclei in the nanostructure. Scaling considerations and accuracy requirements established DFT as the most viable candidate for handling this task. In a general form, density functional is written as

$$E_e[\rho] = T[\rho(\mathbf{r})] + E^{Har}[\rho(\mathbf{r})] + E^{xc}[\rho(\mathbf{r})],$$

where $T[\rho(\mathbf{r})]$ is the kinetic energy functional, $E^{Har}[\rho(\mathbf{r})]$ is the electron-electron Coulomb repulsion energy, and $E^{xc}[\rho(\mathbf{r})]$ is the exchange and correlation energy. The ground-state electron density is the function $\hat{\rho}(\mathbf{r})$ that

minimizes $E_e[\rho]+E_{ext}(\rho(\mathbf{r});\{\mathbf{R}_I\})$ with respect to the electron density subject to a charge conservation constraint, $\int \hat{\rho}(\mathbf{r})d\,\mathbf{r} - N = 0$, and the requirement that the density stay nonnegative.

The orbital-free DFT (OFDFT) methods (see, for instance,[15]), based on the explicit approximations to the unknown exact functional are attractive as they are numerically easier to formulate and solve than the most widely used Kohn-Sham approach (KS-DFT)[1] and there is no need for orbital localization and orthonormalization. Compared to the KS-DFT approaches that typically require the solution of a nonlinear algebraic eigenvalue problem, the OFDFT approaches result in optimization problems based on a methodology that scales linearly and is relatively simple to implement. The main difficulty lies in the quality of approximate functionals, since the exact functionals are not known. Efforts to find accurate functionals have been quite successful for several simple metal systems. OFDFT has recently been used in molecular dynamics simulations for accurate representation of interatomic forces in order to reproduce and provide an explanation for calorimetry results in Na clusters[16], for the studies of several thousand atoms near a metallic grain boundary[17], for predict of the dislocation nucleation during nanoindentation of Al_3Mg[18] used in combination with the quasicontinuum method[9;19], and for the metal-insulator transition in a two-dimensional array of metal nanocrystal quantum dots[20].

Approximations for the exchange and correlation energy functionals, $K\,[\rho]$, are discussed, for instance, in[21] and[22]. Providing suitable expressions for the kinetic energy functional remains a challenging task and, because of reduced transferability, is the factor that prevents widespread use of the approach. The simplest explicit functional is due to Thomas and Fermi[23;24]

$$T_{TF}\,[\rho] = C_F \int \rho^{\frac{5}{3}}(\mathbf{r})\ \mathtt{d}\mathbf{r},$$

where C_F is a constant. This crude approximation has been improved on by the von Weizsäcker form of the kinetic energy functional[25],

$$T_{vW}[\rho] = T_{TF}[\rho] + \frac{1}{8}\int \frac{|\nabla\rho(\mathbf{r})|^2}{\rho(\mathbf{r})}\ \mathtt{d}\mathbf{r},$$

which has been further improved on[26;27;28;15]

$$\begin{aligned} T_{Hyb}[\rho] &= T_{vW}[\rho] + \textstyle\sum_\alpha \lambda_\alpha T_\alpha \\ T_\alpha &= \int\int \rho^\alpha(\mathbf{r})\rho^\alpha(\mathbf{r}')w_\alpha(\mathbf{r}-\mathbf{r}';\rho(\mathbf{r}))\ \mathtt{d}\mathbf{r}\ \mathtt{d}\mathbf{r}' \end{aligned},$$

where the function $w_\alpha(\mathbf{r}-\mathbf{r}';\rho(\mathbf{r}))$ is an electron density-dependent kernel that is formulation dependent.

Within the framework of OFDFT for electronic structure computation a model reduction approach is pursued that relies on the near-periodicity assumption introduced above[29;30]. The entire domain of interest is first meshed and divided into subdomains. Using a finite element approach one can express the kinetic, exchange correlation, Coulomb, and electron-nuclei interaction energies in terms of the values of the electron density at grid points. Since the bulk of a nanostructure often displays quasiperiodic conditions, not all the density grid point values will be considered as "degrees of freedom". Instead, in order to reduce the overall dimension of the optimization problem, only the density value at grid points of so-called *active* subdomains are considered actual degrees of freedom. Among the active subdomains is a subset of *reconstruction* subdomains, which are used in recovering the value of the electron density in the nonreference subdomains. The latter are called *passive* subdomains. If there are no passive subdomains, no reconstruction process is involved, and the proposed approach becomes an OFDFT domain partitioning scheme in which all the degrees of freedom are accounted for, and the subdomains are treated in parallel.

The value of the density in the passive subdomains is implicitly accounted through the interpolation operator acting on the reconstruction subdomains in a self-consistent manner. In its simplest representation, the reconstruction of the electron density in a passive subdomain $\mathtt{D}_i$ (see Figure 1) is characterized by two sets of parameters: the reconstruction weights $\vartheta^\alpha(i)$, and the reconstruction vectors $\mathbf{T}^\alpha(i)$, where a Greek superscript is used to indicate the index of a reconstruction subdomain $\mathtt{Y}_\alpha$. The reconstruction vector $\mathbf{T}^\alpha(i)$ takes the point $\mathbf{r}$ in subdomain $\mathtt{D}_i$ to its image in the reconstruction subdomain $\mathtt{Y}_\alpha$, and $\vartheta^\alpha(i)$ is the weight with which the subdomain $\mathtt{Y}_\alpha$ participates in the reconstruction of the electron density in subdomain $\mathtt{D}_i$. Generalizing this idea, if $\bar{R}$ is a function that depends

on the electron density, the proposed *reconstruction ansatz* calls for a computation of the value of $\bar{R}$ at a point $\mathbf{r}$ that belongs to a passive subdomain D_i as a linear combination of values of the function $\bar{R}$ evaluated at suitably chosen points in the reconstruction subdomains, which are determined based on the underlying near-periodicity of the material assumption. Referring to Figure 1, $\bar{R}(\mathbf{r})$ in subdomain D_i is expressed by interpolation in terms of values $\bar{R}(\mathbf{r}^\alpha)$, for $\mathbf{r}^\alpha \in \mathsf{Y}_\alpha, \alpha \in \{1, \ldots, 7\}$:

$$\bar{R}(\mathbf{r}) = \sum_{\alpha \in \mathcal{R}(i)} \vartheta^\alpha(i)\, \bar{R}(\mathbf{r}^\alpha),$$

where $\mathcal{R}(i)$ represents the union of all reconstruction subdomains Y_α involved in the reconstruction of subdomain D_i, and the reconstruction weights ϑ are determined based on the type of interpolation considered. The deformation field factors into the reconstruction scheme. Concretely, in the proposed *reconstruction ansatz* $\bar{R}(\Phi(\mathbf{r}^0, t))$ is replaced in passive subdomains with a linear combination of values in the reconstruction subdomains taking into account the underlying near-periodicity of the material:

$$\bar{R}(\Phi(\mathbf{r}^0, t)) = \sum_{\alpha \in \mathcal{R}(i)} \vartheta^\alpha(i) \bar{R}(\Phi(\mathbf{r}^0 + \mathbf{T}^\alpha(i), t)) \ .$$

Here $\mathbf{T}^\alpha(i)$ is the reconstruction vector that defines in the reference configuration the corresponding point out of subdomain Y_α that is engaged in the reconstruction of D_i. In a perfect crystal the reconstruction vector would be chosen based on the primitive vectors of the Bravais lattice (see, for instance,[31]). Referring back to the example presented in Figure 1, $\mathcal{R}(4) = \mathcal{R}(5) = \mathcal{R}(6) = \{3, 4\}$; in other words, the reconstruction of the subdomains D_4, D_5, and D_6 is based on values of the density in subdomains Y_3 and Y_4. Similarly, $\mathcal{R}^{-1}(\alpha)$ represents the set of all the subdomains that have the values of the density reconstructed based on values from Y_α; for instance, $\mathcal{R}^{-1}(3) = \{3, 4, 5, 6\}$, in other words, the reconstruction subdomain Y_3 is implicated in the reconstruction of D_3, D_4, D_5, and D_6. In general, the subdomains D_i may be thought to be of identical shape, in which case the interpolation approach is reminiscent of the gap-tooth method[32] where the reference subdomains are the "teeth". Herein, however, the reconstruction by interpolation of the density is also carried out in the gaps, and not only at the boundary of the teeth, because of the long-range electrostatic interactions.

It is reasonable to expect that there will be parts of the nanostructure where the reconstruction approach is not applicable because of the breakdown of the near-periodicity assumption. In these cases, all subdomains spanning such volumes will be active, effectively leading to a domain decomposition approach to OFDFT calculations.

Formulation Framework

The calculation of electron energy E_e requires the evaluation of integrals of the form

$$I[\Theta] = \int \Theta(\mathbf{r})\, \mathrm{d}\mathbf{r}.$$

The integrand Θ will be partitioned into two components: $\Theta(\mathbf{r}) = \bar{R}(\rho(\mathbf{r}), \mathbf{r})\, \bar{L}(\mathbf{r})$. $\bar{R}$ depends on the value of the density at the location $\mathbf{r}$ and possibly on the spatial component $\mathbf{r}$ itself (to simplify the notation, without any loss of generality, this component will be denoted by $\bar{R}(\mathbf{r})$). The component $\bar{L}$ does not depend on the electron density ρ. For instance, in the case of the electron-nuclei interaction,

$$I[\Theta] = E_{ne} = -\int \rho(\mathbf{r}) \sum_{A=1}^{M} \frac{Z_A}{||\mathbf{r} - \mathbf{R}_A||}\, \mathrm{d}\mathbf{r} \Rightarrow \left\{ \begin{array}{lcl} \bar{R}(\rho(\mathbf{r}), \mathbf{r}) & = & \rho(\mathbf{r}) \\ \bar{L}(\mathbf{r}) & = & \sum\limits_{A=1}^{M} \frac{-Z_A}{||\mathbf{r} - \mathbf{R}_A||} \end{array} \right. .$$

The other energies for the Thomas-Fermi-Dirac functional can be cast into this form as well, with the double integral component being treated separately. With $D^t = \Phi(D^0, t)$,

$$E_e = \int\limits_{D^t} \bar{R}(\mathbf{r})\, \bar{L}(\mathbf{r})\, \mathrm{d}\mathbf{r} = \int\limits_{D^0} \bar{R}(\Phi(\mathbf{r}^0, t))\, \bar{L}(\Phi(\mathbf{r}^0, t))\, J(\mathbf{r}^0, t)\, \mathrm{d}\mathbf{r}^0,$$

where $J(\mathbf{r}^0,t) = \left|\det(\nabla_0\,\Phi(\mathbf{r}^0,t))\right|$. The notation $R(\mathbf{r}^0) = \bar{R}(\Phi(\mathbf{r}^0,t))$ and $L(\mathbf{r}^0) = \bar{L}(\Phi(\mathbf{r}^0,t))J(\mathbf{r}^0,t)$ will be used; and although there is a time dependency involved, it will be omitted for brevity. Likewise, the zero superscript, which indicates that the integration is with respect to the initial configuration, will be dropped to simplify the notation. With this, E_e requires the computation of quantities like

$$I[\Theta] = \int\limits_{D^0} R(\mathbf{r})L(\mathbf{r})\,\mathrm{d}\mathbf{r}.$$

As far as the nomenclature is concerned, at a point $\mathbf{r}$, the $R(\mathbf{r})$ component is reconstructed according to the proposed ansatz and thus computed as a linear combination of functions evaluated at *remote* points. The component $L(\mathbf{r})$ is evaluated at the *local* point $\mathbf{r}$. This partitioning is used to compute the integral $I[\Theta]$ in terms of electron density values from the active subdomains using a suitably chosen quadrature rule:

$$\begin{aligned} I[\Theta] &= \sum_{i=1}^{u} \sum_{\tau_{i,j}\in\mathcal{M}(\mathtt{D}_i)} \sum_{k\in Q(i,j)} w_{i,j,k} R(\mathbf{r}_{i,j,k}) L(\mathbf{r}_{i,j,k}) \\ &= \sum_{i=1}^{u} \sum_{k\in Q(i)} w_{i,k}\, R(\mathbf{r}_{i,k})\, L(\mathbf{r}_{i,k}) \end{aligned},$$

where u is the total number of subdomains and, for $k \in Q(i,j)$, $\mathbf{r}_{i,j,k}/w_{i,j,k}$, represent the quadrature points/weights in cell $\tau_{i,j}$ of mesh $\mathcal{M}(\mathtt{D}_i)$ for computing the integral $I[\Theta]$ on the subdomain $\mathtt{D}_i$. Note that in order to keep the notation simple, the j subscript associated with the cell has been dropped. Reconstruction is applied to get $R(\mathbf{r}_{i,k})$:

$$R(\mathbf{r}_{i,k}) = \sum_{\alpha\in\mathcal{R}(i)} \vartheta^\alpha(i) R(\mathbf{r}_{i,k} + \mathbf{T}^\alpha(i)) = \sum_{\alpha\in\mathcal{R}(i)} \vartheta^\alpha(i) R(\mathbf{r}^\alpha_{i,k}).$$

This is the case when the same subdomains are involved in the reconstruction of the value of $\mathbf{r}_{i,k}$ for $k \in Q(i)$ and might not be the case if the partitioning of the overall domain in subdomains $\mathtt{D}_i$ and $\mathtt{Y}_\alpha$ is not done carefully.

In what follows the *clout* $C^\alpha_n(i)$ of a node n in the mesh $\mathcal{M}(\mathtt{Y}_\alpha)$ relative to the subdomain $\mathtt{D}_i$ represents the set of indices k for which the associated quadrature point $\mathbf{r}_{i,k} \in \mathtt{D}_i$, when subjected to the reconstruction translation, falls within a cell of $\mathcal{M}(\mathtt{Y}_\alpha)$ for which n is a node. Using this notation and linear shape function-based interpolation, we obtain

$$\begin{aligned} &\sum_{k\in Q(i)} w_{i,k}\, L(\mathbf{r}_{i,k})\, R(\mathbf{r}_{i,k}) \\ &= \sum_{k\in Q(i)} w_{i,k}\, L(\mathbf{r}_{i,k}) \left[\sum_{\alpha=1}^{p} \vartheta^\alpha(i) \sum_{n\in\mathcal{N}(\tau^\alpha(\mathbf{r}_{i,k}))} R(\mathbf{r}^\alpha_n)\, \varphi^\alpha_n(\mathbf{r}^\alpha_{i,k}) \right] \\ &= \sum_{\alpha\in\mathcal{R}(i)} \sum_{n\in\mathcal{M}(\mathtt{Y}_\alpha)} R(\mathbf{r}^\alpha_n)\, \vartheta^\alpha(i) \sum_{k\in C^\alpha_n(i)} w_{i,k}\, L(\mathbf{r}_{i,k})\, \varphi^\alpha_n(\mathbf{r}^\alpha_{i,k}), \end{aligned}$$

where $\tau^\alpha(\mathbf{r}_{i,k})$ is a function that returns the cell in the mesh $\mathcal{M}(\mathtt{Y}_\alpha)$ in which the quadrature point $\mathbf{r}_{i,k} \in \mathtt{D}_i$ falls when subjected to the reconstruction translation, and $\mathcal{N}(\tau)$ returns the set of node points associated with the cell τ. Typically, a node n has several cells that it belongs to, and a shape function is associated to each pair (node n, cell it belongs to). This aspect is acknowledged, but for simplicity the notation does not reflect this dependency. Defining

$$\begin{aligned} \kappa^{\alpha\leftarrow i}_n &= \vartheta^\alpha(i) \sum_{k\in C^\alpha_n(i)} w_{i,k}\, L(\mathbf{r}_{i,k})\, \varphi^\alpha_n(\mathbf{r}^\alpha_{i,k}), \\ \kappa^\alpha_n[L] &= \sum_{i\in\mathcal{R}^{-1}(\alpha)} \kappa^{\alpha\leftarrow i}_n, \end{aligned}$$

the dependency of the kernel at node n in the subdomain $\mathtt{Y}_\alpha$ is explicitly indicated to depend on the expression of the local function component L: $\kappa^\alpha_n = \kappa^\alpha_n[L]$. The integral and its derivative with respect the value of the electron

density at a node n of the mesh $\mathcal{M}(\mathsf{Y}_\alpha)$ are expressed as

$$\begin{aligned} I[\Theta] &= \sum_{\alpha=1}^{p} \sum_{n\in\mathcal{M}(\mathsf{Y}_\alpha)} \kappa_n^\alpha \, R(\mathbf{r}_n^\alpha) = \kappa[L]\cdot\mathbf{R}[\hat{\rho}], \\ \frac{\partial I[\Theta]}{\partial \hat{\rho}_n^\alpha} &= \kappa_n^\alpha[L] \, \frac{\partial R}{\partial \rho}(\hat{\rho}_n^\alpha), \end{aligned}$$

where

$$\begin{aligned} \kappa[L] &= \left[\kappa_1^1[L],\ldots,\kappa_{y(1)}^1[L],\ldots\ldots,\kappa_1^p[L],\ldots,\kappa_{y(p)}^p[L]\right], \\ \mathbf{R}[\hat{\rho}] &= \left[R(\mathbf{r}_1^1),\ldots,R(\mathbf{r}_{y(1)}^1),\ldots\ldots,R(\mathbf{r}_1^p),\ldots,R(\mathbf{r}_{y(p)}^p)\right]^T. \end{aligned}$$

Here $y(\alpha)$ represents the number of nodes in the reconstruction subdomain Y_α, and $R(\mathbf{r}_n^\alpha)$ is the value of the function R evaluated at the node n of the mesh $\mathcal{M}(\mathsf{Y}_\alpha)$. The notation $R[\hat{\rho}]$ emphasizes that this vector depends on the value of the density ρ but only at a discrete set of locations, that is, the nodes of the meshes $\mathcal{M}(\mathsf{Y}_\alpha)$, for $\alpha = 1,\ldots,p$. The kernel vector is constant and evaluated once; the vector $R[\hat{\rho}]$ changes with the value of the density and in an iterative process should be evaluated at each iteration.

A matrix-vector notation describes the above procedure more concisely. For a subdomain D_i and a reconstruction subdomain Y_α, $\alpha\in\mathcal{R}(i)$, a quadrature matrix is defined to capture the concept of a clout associated with a node n in Y_α relative to the subdomain D_i. Thus, $\mathbf{Q}^{\alpha\leftarrow i;\mathcal{O}} \in \mathbb{R}^{q(i;\mathcal{O})\times y(\alpha)}$ has as many rows as there are quadrature points $q(i;\mathcal{O})$ in the subdomain D_i, and a number of columns equal to the number of nodes $y(\alpha)$ in the reconstruction subdomain Y_α, for $\alpha\in\mathcal{R}(i)$. The superscript $\mathcal{O}$ is necessary to differentiate between different quadrature types in the case of a double integral, as will be the case shortly. The notation suggests that this matrix refers to the outermost integral; for a double integral a superscript $\mathcal{I}$ is used to refer to a quantity defined in relation to the innermost integral. Note also that the number of quadrature points $q(i;\mathcal{O})$ depends on what quadrature rule is considered for integration and that the factor $\vartheta^\alpha(i)$ that indicates the weight of the subdomain Y_α in the reconstruction of the subdomain D_i is also rolled into the expression for $\mathbf{Q}^{\alpha\leftarrow i;\mathcal{O}}$. For a quadrature point $\mathbf{r}_{i,k}\in\mathsf{D}_i$, the entry (k,n) is nonzero provided $k\in C_n^\alpha(i)$. Therefore, the clout of a node n is the set of rows with nonzero entries in the column associated with this node. A nonzero entry assumes the form

$$\mathbf{Q}^{\alpha\leftarrow i;\mathcal{O}}[k,n] = \vartheta^\alpha(i)\, w_{i,k}\, \varphi_n^\alpha(\mathbf{r}_{i,k}^\alpha).$$

Defining

$$\tilde{\mathbf{L}}_i^{\mathcal{O}} = \left[\begin{array}{ccc} L(\mathbf{r}_{i,1}^{\mathcal{O}}) & \ldots & L(\mathbf{r}_{i,q(i;\mathcal{O})}^{\mathcal{O}}) \end{array}\right],$$

then, for $i\in\mathcal{R}^{-1}(\alpha)$, we have

$$\kappa^{\alpha\leftarrow i} = \tilde{\mathbf{L}}_i^{\mathcal{O}}\, \mathbf{Q}^{\alpha\leftarrow i;\mathcal{O}}. \tag{3}$$

Approximation of a double integral will now be established for the Coulomb integral:

$$J\,[\rho] = \frac{1}{2}\int\int \frac{\rho(\mathbf{r})\;\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\;\mathrm{d}\mathbf{r}\;\;\mathrm{d}\mathbf{r}'.$$

Defining first

$$L(\mathbf{r}) = \int \frac{\rho(\mathbf{r}')}{|\mathbf{r}'-\mathbf{r}|}\,\mathrm{d}\mathbf{r}',$$

we can approximate the Coulomb integral as

$$J\,[\rho] = \frac{1}{2}\sum_{\alpha=1}^{p} \sum_{n\in\mathcal{M}(\mathsf{Y}_\alpha)} \rho_n^\alpha \sum_{i\in\mathcal{R}^{-1}(\alpha)} \vartheta^\alpha(i) \sum_{k\in C_n^\alpha(i)} w_{i,k}\, L(\mathbf{r}_{i,k})\, \varphi_n^\alpha(\mathbf{r}_{i,k}^\alpha).$$

Furthermore,

$$L(\mathbf{r}_{i,k}) = \sum_{\beta=1}^{p} \sum_{m \in \mathcal{M}(\Upsilon_\beta)} \kappa_m^\beta[i,k]\ \rho_m^\beta,$$

where the notation $\kappa_m^\beta[i,k]$ indicates that the kernel κ_m^β[i,k] corresponds to the local function $|\mathbf{r}' - \mathbf{r}_{i,k}|^{-1}$. Using the notation

$$K_{nm}^{\alpha\beta} = \sum_{i \in \mathcal{R}^{-1}(\alpha)} \sum_{k \in C_n^\alpha(i)} \vartheta^\alpha(i)\, w_{i,k}\, \kappa_m^\beta[i,k]\ \varphi_n^\alpha(\mathbf{r}_{i,k}^\alpha)$$

leads to

$$\begin{aligned} J[\rho] &= \frac{1}{2}\,\rho^T \mathbf{K} \rho \\ \mathbf{K} &= \begin{pmatrix} \mathbf{K}^{11} & \mathbf{K}^{12} & \dots & \mathbf{K}^{1p} \\ \mathbf{K}^{21} & \mathbf{K}^{22} & \dots & \mathbf{K}^{2p} \\ \dots & \dots & \dots & \dots \\ \mathbf{K}^{p1} & \mathbf{K}^{p2} & \dots & \mathbf{K}^{pp} \end{pmatrix} \\ \mathbf{K}^{\alpha\beta} &= \sum_{i \in \mathcal{R}^{-1}(\alpha)} \sum_{j \in \mathcal{R}^{-1}(\beta)} \mathbf{K}^{\alpha \leftarrow i, \beta \leftarrow j} = \left[K_{nm}^{\alpha\beta} \right] \quad n = 1, \dots, y(\alpha),\ \ m = 1, \dots, y(\beta) \end{aligned}$$

with $\mathbf{K}^{\alpha \leftarrow i, \beta \leftarrow j}$ yet to be defined. Corresponding to the quadrature point associated with the outer integral, $\mathbf{r}_{i,k}^{\mathcal{O}}$, a row vector is defined as

$$\tilde{\mathbf{L}}_j^{\mathcal{I}}[i,k] = \left[\ |\mathbf{r}_{i,k}^{\mathcal{O}} - \mathbf{r}_{j,1}^{\mathcal{I}}|^{-1} \quad \dots \quad |\mathbf{r}_{i,k}^{\mathcal{O}} - \mathbf{r}_{j,q(j;\mathcal{I})}^{\mathcal{I}}|^{-1}\ \right]. \tag{4}$$

Then,

$$\kappa^{\beta \leftarrow j}[i,k] = \tilde{\mathbf{L}}_j^{\mathcal{I}}[i,k]\ \mathbf{Q}^{\beta \leftarrow j;\mathcal{I}} \in \mathbb{R}^{1 \times y(\beta)}.$$

Define

$$\tilde{\mathbf{L}}^{i,\mathcal{O};j,\mathcal{I}} = \begin{bmatrix} \tilde{\mathbf{L}}_j^{\mathcal{I}}[i,1] \\ \dots \\ \tilde{\mathbf{L}}_j^{\mathcal{I}}[i, q(i;\mathcal{O})] \end{bmatrix}$$

and

$$\kappa^{\beta \leftarrow j}[i] = \begin{bmatrix} \kappa^{\beta \leftarrow j}[i,1] \\ \dots \\ \kappa^{\beta \leftarrow j}[i, q(i;\mathcal{O})] \end{bmatrix} = \tilde{\mathbf{L}}^{i,\mathcal{O};j,\mathcal{I}}\ \ \mathbf{Q}^{\beta \leftarrow j;\mathcal{I}}.$$

Then,

$$\mathbf{K}^{\alpha \leftarrow i, \beta \leftarrow j} = \left[\mathbf{Q}^{\alpha \leftarrow i;\mathcal{O}}\right]^T \ \ \left[\tilde{\mathbf{L}}^{i,\mathcal{O};j,\mathcal{I}}\right] \ \ \left[\mathbf{Q}^{\beta \leftarrow j;\mathcal{I}}\right]. \tag{5}$$

Note that $\kappa^{\beta \leftarrow j}[i] \in \mathbb{R}^{q(i;\mathcal{O}) \times y(\beta)}$ and $\mathbf{K}^{\alpha \leftarrow i, \beta \leftarrow j} \in \mathbb{R}^{y(\alpha) \times y(\beta)}$. Implementation details for the parallel evaluation of the method's associated kernels are discussed in[33].

The Optimization Problem

The formalism introduced for the computation of an integral $I[\Theta] = \int \Theta(\mathbf{r})\, \mathrm{d}\mathbf{r}$ hinges on the partitioning $\Theta(\mathbf{r}) = R(\mathbf{r})\, L(\mathbf{r})$ and has been applied to the Thomas-Fermi-Dirac DFT, leading to the following optimization problem:

$$\begin{aligned} \min E_{TOT} &= -C_X \kappa \cdot \hat{\rho}^{\frac{4}{3}} + C_F \kappa \cdot \hat{\rho}^{\frac{5}{3}} + \kappa_{ne} \cdot \hat{\rho} + \frac{1}{2}\,\hat{\rho}^T \mathbf{K} \hat{\rho} \\ 0 &= \kappa \cdot \hat{\rho} - N_e \\ 0 &\le \hat{\rho} \end{aligned}$$

where $\hat{\rho} = \left[\hat{\rho}_1^1, \ldots, \hat{\rho}_{y(1)}^1, \ldots, \hat{\rho}_1^p, \ldots, \hat{\rho}_{y(p)}^p\right]^T$ and

$$\kappa = \kappa[1] \qquad \kappa_{ne} = \kappa \left[\sum_{A=1}^{M} \frac{Z_A}{|\mathbf{r} - \mathbf{R}_A|}\right].$$

Defining for an exponent $c \in \mathbb{R}_+$ a diagonal matrix

$$\mathbf{D}[\rho^c] = diag\left(\left(\rho_1^1\right)^c, \ldots, \left(\rho_{y(1)}^1\right)^c, \ldots, \left(\rho_p^1\right)^c, \ldots, \left(\rho_{y(p)}^p\right)^c\right),$$

we obtain the gradient of the cost function

$$\nabla E_{TOT} = \mathbf{g}[\hat{\rho}] = \kappa^T \left(\frac{5}{3} C_F \mathbf{D}[\hat{\rho}^{\frac{2}{3}}] - \frac{4}{3} C_X \mathbf{D}[\hat{\rho}^{\frac{1}{3}}]\right) + \kappa_{ne}^T + \frac{1}{2}\hat{\rho}^T \left(\mathbf{K} + \mathbf{K}^T\right).$$

The Hessian is evaluated as

$$\mathbf{H}[\hat{\rho}] = \mathbf{H}_d[\hat{\rho}] + \frac{1}{2}\left(\mathbf{K} + \mathbf{K}^T\right),$$

where

$$\begin{array}{rcll} \mathbf{H}_d[\hat{\rho}] & = & diag & \left(\mathbf{H}^1[\hat{\rho}], \ldots, \mathbf{H}^p[\hat{\rho}]\right), \\ \mathbf{H}^\alpha[\hat{\rho}] & = & diag & \left(\kappa_1^\alpha \left(\frac{2}{3} C_F \left(\hat{\rho}_1^\alpha\right)^{-\frac{1}{3}} - \frac{1}{3} C_X \left(\hat{\rho}_1^\alpha\right)^{-\frac{2}{3}}\right), \ldots, \right. \\ & & , \ldots, & \left. \kappa_{y(\alpha)}^\alpha \left(\frac{2}{3} C_F \left(\hat{\rho}_{y(\alpha)}^\alpha\right)^{-\frac{1}{3}} - \frac{1}{3} C_X \left(\hat{\rho}_{y(\alpha)}^\alpha\right)^{-\frac{2}{3}}\right)\right). \end{array}$$

The value of the electron density should always remain positive, and therefore the minimization is best approached in the framework of bound constrained optimization. Bound-constrained optimization problems (BCOPs) have the form

$$\min\{f(x) : l \leq x \leq u\},$$

where $f : \mathbb{R}^n \mapsto \mathbb{R}$ is a nonlinear function with continuous first- and second-order derivatives, the vectors l and u are fixed, and the inequalities are taken componentwise. A classical result[34] shows that the bound-constrained optimization problem has a unique solution on the feasible region

$$\Omega = \{x \in \mathbb{R}^n : l \leq x \leq u\}$$

when the function $f : \mathbb{R}^n \mapsto R$ is strictly convex. This result holds for unbounded Ω, and the components of l and u are allowed to be infinite. For the projection operator

$$[T_\Omega d]_i = \begin{cases} d_i & \text{if} \quad x_i \in (l_i, u_i) \\ \min\{d_i, 0\} & \text{if} \quad x_i = l_i \\ \max\{d_i, 0\} & \text{if} \quad x_i = u_i \end{cases},$$

x^* is a solution of the BCOP if and only if the projected gradient $T_\Omega \nabla f(x^*) = 0$. Given a tolerance τ, an approximate solution to the BCOP is any $x \in \Omega$ such that

$$\|T_\Omega \nabla f(x)\| \leq \tau.$$

Note that this holds whenever x is sufficiently close to x^*. Algorithms for solving these problems are usually generalizations of well-known methods for unconstrained optimization. For unconstrained optimization, Newton's method, for example, solves a linear system involving the Hessian matrix of second derivatives and the gradient vector. Each iteration of active-set methods fixes a set of variables to one of their bounds and solves an unconstrained minimization problem using the remaining variables. A set of three algorithms used in conjunction with the electronic structure computation problem is presented and discussed in[33]. These algorithms are part of the Toolkit for Advanced Optimization (TAO) library [35;36]. TAO provides optimization software for the solution of scientific applications on high-performance architectures. These applications include minimizing energy functionals that arise in differential equations and molecular geometry optimization. Various software packages are available for solving these problems, but TAO provides the portability and scalability necessary for parallel optimization on high performance computers (Linux clusters, IBM BG/L, etc.).

PROPOSED COMPUTATIONAL SETUP

Given a nanostructure of known atomic composition, not necessarily monoatomic or single-crystal, the goal is to determine the electron density distribution as well as the positions of the nuclei, that is, the mapping Φ. Here we do not consider dynamics of the nuclei.

As indicated in Fig. 2, the proposed solution has three principal modules: the *preprocessing* stage, the *electronic problem* and the *ionic problem*. *Preprocessing* is carried out once at the beginning of the simulation. A domain D^0 is selected to include the nanostructure investigated. The partitioning of D^0 into u subdomains $\mathsf{D}_i, i = 1, \ldots, u$, is done to mirror the underlying periodicity of the structure. The subdomains $\mathsf{D}_{\chi(1)}$ through $\mathsf{D}_{\chi(p)}$ become the active subdomains and, as in Figure 1, they are denoted by Y_1 through Y_p. A set of values of the electron density is required at the nodes of the discretization mesh; the initial guess for the electron density could be an overlap of isolated atom electron densities throughout the nanostructure or, when practical, could be obtained based on periodic boundary conditions assumption by computing it in a domain D_j and then cloning for the remaining domains D_k. *Preprocessing* concludes with the initialization of the deformation map Φ to identity.

With a suitable norm, the new electron density ρ^{new} is compared to ρ^{init}, and the computation restarts the *electronic problem* after setting $\rho^{init} = \rho^{new}$ unless the corrected and initial values of the electron density are close. This iterative process constitutes the first inner loop of the algorithm.

The *ionic problem* uses the newly computed electron density to reposition the nuclei. The nonlinear system in Eq. (2) provides the position of the *repnuclei*; the other nuclei are positioned based on the quasicontinuum paradigm according to Eq. (1). The nonlinear system is solved by an iterative method that leads to the second inner loop, which has four steps: (a) evaluate the integral of Eq. (2); when necessary, evaluate its partial with respect to $\mathbf{P}_i$; (b) evaluate the double sum of Eq. (2), which is based on a partitioning of the structure, and, when necessary, evaluate its partial with respect to the position of the representative atoms; (c) carry out a quasi-Newton step to update the positions $\mathbf{P}_i$ of the M_{rep} representative nuclei; and (d) go back to (a) if no converge results.

The precision in determining the positions of the nuclei is directly influenced by the accuracy of the electron density $\rho(\mathbf{r})$. Thus, an important issue, not addressed by this work, is the sensitivity of the solution of the nonlinear system in Eq. (2) with respect to $\rho(\mathbf{r})$. It remains to be determined what level of approximation of the electron density suffices for solving the ionic problem at a satisfactory level of accuracy. After determining the position of the nuclei, the algorithm computes the new deformation mapping Φ according to Eq. (1). If the overall change in the position of *repnuclei* at the end of the ionic problem is smaller than a threshold value, the computation stops; otherwise the new distribution of the nuclei is the input to a new electronic problem (second stage of the algorithm). In summary, the algorithm passes through the preprocessing stage once. It then solves the electronic problem (the first inner loop) and proceeds to the ionic problem (the second inner loop). The outer loop (electronic problem, followed by ionic problem) stops when there is no significant change in the position of the *repnuclei*.

PRELIMINARY NUMERICAL RESULTS

The fact that quasicontinuum method represents a meaningful reduction model approach has been established and documented in[7;8;9;19]. The focus of the numerical experiments presented here is on model reduction as applied to the electronic problem. The approach proposed for the solution of the electronic problem has been investigated in the context of undeformed topologies. In other words, for the deformation gradient $\nabla_0\ \Phi(\mathbf{r}^0, t)$, $J(\mathbf{r}^0, t) = \det(\nabla_0\ \Phi(\mathbf{r}^0, t)) = 1$.

String of Atoms Example

Our first example is a three-dimensional variation of the one dimensional case analyzed in the previous section. The size of each of the 3D subdomains surrounding a hydrogen atom is $3 \times 3 \times 3$ (all units henceforth are atomic units). A full simulation with no reconstruction is provided as the reference solution. Two scenarios with seven and five active subdomains were subsequently considered; all meshes in this numerical experiment are uniform. In the first scenario, the subdomains D_1, D_2, D_3, D_7, D_{11}, D_{12}, and D_{13} were active; only D_3, D_7, and D_{11} were

Active Subdomains	13	7	5
Number of Iterations	605	245	221
Total Energy	-14.257	-14.256	-14.256

Table 1: Uniform mesh summary of the results. TAO-BLMVM optimization constraints are 10^{-6} for absolute and 10^{-5} for relative convergence tolerance.

used for reconstruction. In the second scenario, the subdomains D_1, D_2, D_7, D_{12}, and D_{13} were active; only D_2, D_7, and D_{12} were used for reconstruction. For this test, the number of nodes/cells in the active subdomains is as follows: 28561/22464 for the nonreconstruction case (13/13), 15379/12096 for the 7/13, and 10985/8640 for the 5/13 case. All meshes considered herein, uniform or variable, are made up of hexahedrons. Figure 3 displays the relative errors; shown are only the regions where the relative error is larger than 5%. The results show a slight improvement in the seven-subdomain case; as the number of active subdomains increases, the quality of the results improve. Because of the dimension reduction, the size of the optimization problem decreases, thereby leading to a reduction in the number of iterations. Moreover, each iteration is computationally less expensive. The large relative errors are explained by the small values assumed by the electron density away from the nuclei where in practice it is expected to be zero. This and the boundary artifacts explain the accumulation of the 5% relative error isosurfaces far away from the nuclei and close to the boundary of the solution domain. The differences in total energy are small for both the 7 and 5 active subdomain cases (about 0.007%; see Table 1). The results reported were obtained by running in parallel with 13 processes on a Linux cluster.

Slab of Atoms Example

Figure 4 shows the results obtained for the 5×5 subdomain 3D slab. Of the 25 subdomains considered for this simulation, one subdomain per nucleus of a hydrogen atom, only nine subdomains of darker color were considered active and used for reconstruction purposes. Figure 4a displays the electron density distribution on a mid-Z slice for the reconstructed domain (9/25). Figure 4b displays the subdomain structure of the slab, and Fig. 4c shows the relative error produced through reconstruction. Compared to the reference case, the relative error in the total value of the electronic energy was 0.03%. The number of nodes/cells for the 5×5 case with all subdomains active was 33275/25000. For the 9/25 reconstruction scheme, the number of unknowns was reduced from 33275 to 11979. The 3D simulation was run in parallel using 25 processes on a Linux cluster.

Nonuniform Mesh Results

Our third test investigated the effect of mesh adaptivity. An example consisting of a string of five hydrogen atoms was run in parallel on IBM BlueGene/L using five processes with no reconstruction. The solution on a uniform mesh is plotted in Fig. 5a; the variable mesh solution is presented in Fig. 5b. Although in both cases the number of mesh points is comparable, the total energy in the nonuniform case has a slightly smaller value, which indicates that it corresponds to a more relaxed distribution of the electron density. The peak electron density values are also higher for the variable mesh case because of a refined mesh capable of capturing fast variations in the vicinity of the nuclei. The energy values are slightly different in the two situations (a difference of 12%, from -5.8 to -5.2). In Fig. 5 a "smearing" effect is noticed in the constant-size mesh, where the relatively higher values of the electron density occupy larger volumes but with lower peaks. Both simulations use the same optimization settings (absolute and relative convergence tolerance). In each of the five subdomains, the number of nodes/cells was 10999/7712 for the variable mesh and 11661/9216 for the uniform mesh.

The number of iterations in the nonuniform mesh case is much larger (2181 as opposed to 212). However, the nonuniform mesh results were obtained without using any acceleration strategy. The poor convergence speed can be addressed by a multigrid approach or by providing Hessian information, which, while straightforward in the proposed approach, is not implemented yet.

When one brings into the picture the reconstruction component, the trend noticed above persists. For the 13-atom example run with 7 active subdomains, the uniform mesh size scenario led to an energy of -14.257 in 245 iterations. The variable mesh case led to -15.54 in 3299 iterations, which is an order of magnitude increase in the number of iterations.

1 Conclusions

A methodology is proposed to address the CPU-intensive task of electron structure computation for a large system. The approach combines a model reduction paradigm and parallel computation capabilities to increase the size and reduce the simulation time associated with large simulations. The entire domain of interest is first meshed and divided into subdomains. The kinetic, exchange correlation, Coulomb, and electron-nuclei interaction energies are expressed in terms of grid values of the electron density in a subset of so-called active subdomains. The resulting form of the energy is minimized subject to the charge conservation constraint. The implementation leverages a domain-decomposition paradigm, and for parallel simulation support it builds on top of the MPICH2 library and the Toolkit for Advanced Optimization. One salient feature of the proposed approach is that the function and gradient evaluations, as well as the optimization stage, are run in parallel. The reconstruction errors were shown to depend on the extent of model reduction. For a test problem consisting of a three-dimensional string of one-electron atoms, the proposed approach led to a threefold reduction in the number of iterations for convergence, while maintaining small values of relative error for the total energy and the electron density in the regions of interest (boundary artifacts led to larger values in these boundary regions).

The method could be improved in three ways. First, and most importantly, more advanced forms of the kinetic and exchange and correlation energy functionals need to be chosen, and the effective core potentials for many electron atoms have to be implemented. Second, for larger problems, cut-off techniques and fast-multipole methods[37] need to be considered. These would ease memory limitations and allow the simulation of large reconstruction tests that go beyond the current proof-of-concept applications. Third, the reconstruction approach should be extended to the DFT Kohn-Sham approach because it has a significantly larger user base than OFDFT.

Acknowledgment

This work was supported in part by the Mathematical, Information, and Computational Sciences Division subprogram of the Office of Advanced Scientific Computing Research, Office of Science, and in part by the Office of Basic Energy Sciences-Materials Sciences, both under U.S. Department of Energy Contract No. DE-AC02-06CH11357. Emil Constantinescu and Toby Heyn are acknowledged for their support in generating the three-dimensional results reported in the paper, and we thank Nick Schafer for reading the manuscript and providing feedback. Use of computer resources from Argonne National Laboratory Computing Resource Center and US DOE National Energy Research Scientific Computing Center is gratefully acknowledged.

References

[1] W. Kohn, L. J. Sham: *Phys. Rev.* **140** (1965) A1133

[2] S. Goedecker, G. E. Scuseria: *Computing in Science and Engineering* **5** (2003) 14

[3] J. M. Soler, E. Artacho, J. D. Gale, A. Garcia, J. Junquera, P. Ordejón, D. Sánchez-Portal: *J. Phys.: Condens. Matter* **14** (2002) 2745

[4] C.-K. Skylaris, P. D. Haynes, A. A. Mostofi, M. C. Payne: *The Journal of Chemical Physics* **122** (2005) 084119

[5] D. R. Bowler, R. Choudhury, M. J. Gillan, T. Miyazaki: *phys. stat. sol. (b)* **243** (2006) 989

[6] T. Kerdcharoen, K. Morokuma: *Chem. Phys. Lett.* **355** (2002) 257

[7] M. Fago, R. L. Hayes, E. A. Carter, M. Ortiz: *Phys. Rev.* **B70** (2004) 100102:1

[8] J. Knap, M. Ortiz: *Journal of the Mechanics and Physics of Solids* **49** (2001) 1899

[9] E. Tadmor, M. Ortiz, R. Phillips: *Philosophical Magazine A* **73** (1996) 1529

[10] P. Hohenberg, W. Kohn: *Phys. Rev.* **136** (1964) B864

[11] M. Anitescu, D. Negrut, T. Munson, P. Zapol: *Density functional theory-based nanostructure investigation: Theoretical considerations*: Tech. Rep. ANL/MCS–P1252–0505, Argonne National Laboratory, Argonne, Illinois (2005)

[12] A. W. Appel: *J. Sci. Stat. Comput.* **6** (1985) 85

[13] L. Greengard: *The rapid evaluation of potential fields in particle systems*: MIT Press, Cambridge, Massachusets (1987)

[14] H. G. Petersen, D. Soelvason, J. W. Perram, E. R. Smith: *J. Chem. Phys.* **101** (1994) 8870

[15] L.-W. Wang, E. A. Carter: In: *Theoretical methods in condensed phase chemistry Progress in Theoretical Chemistry and Physics*, ed. S. D. Schwartz, pp. 117–184. Kluwer, Dordrecht (2000)

[16] A. Aguado, J. M. Lopez: *Physical Review Letters* **94** (2005) 233401

[17] S. C. Watson, P. A. Madden: *Phys Chem Comm* **1** (1998) 1

[18] R. Hayes, G. Ho, M. Ortiz, E. Carter: *Philosophical Magazine* **86** (2006) 2343

[19] R. E. Miller, E. B. Tadmor: *Journal of Computer-Aided Materials Design* **9** (2002) 203

[20] S. C. Watson, E. A. Carter: *Computer Physics Communications* **128** (2000) 67

[21] W. Koch, M. C. Holthausen: *A Chemist's Guide to Density Functional Theory*: John Wiley & Sons Inc., New York, second edn. (2001)

[22] R. M. Martin: *Electronic structure: basic theory and practical methods*: Cambridge University Press, Cambridge, U.K. (2004)

[23] L. H. Thomas: *Proc.Camb.Phil.Soc.* **23** (1927) 542

[24] E. Fermi: *Rend. Accad. Lincei* **6** (1927) 602

[25] C. F. von Weizsacker: *Z. Phys.* **96** (1935) 431

[26] K. M. Carling, E. A. Carter: *Modelling Simul. Mater. Sci. Eng.* **11** (2003) 339

[27] M. Foley, P. A. Madden: *Phys. Rev.* **B53** (1996) 10589

[28] F. Perrot: *J. Phys.: Condens. Matter* **6** (1994) 431

[29] D. Negrut, M. Anitescu, T. Munson, P. Zapol: In: *Proceedings of IMECE 2005, ASME International Mechanical Engineering Congress and Exposition* (2005)

[30] M. Anitescu, D. Negrut, P. Zapol, A. El-Azab: *Mathematical Programming* (2006): conditionally accepted

[31] N. Ashcroft, N. Mermin: *Solid State Physics*: W.B. Saunders Company, Philadelphia (1976)

[32] Y. Kevrekidis, C. W. Gear, J. Li: *Physics Letters A* **190** (2003)

[33] D. Negrut, M. Anitescu, A. El-Azab, S. Benson, P. Zapol: *Journal of Computational Physics, submitted* (2006)

[34] J. Nocedal, S. J. Wright: *Numerical Optimization*: Springer-Verlag, New York (1999)

[35] S. J. Benson, L. C. McInnes, J. Moré, J. Sarich: *TAO user manual (revision 1.8)*: Tech. Rep. ANL/MCS-TM-242, Mathematics and Computer Science Division, Argonne National Laboratory (2005): http://www.mcs.anl.gov/tao

[36] S. J. Benson, L. C. McInnes, J. J. Moré: *ACM Transactions on Mathematical Software* **27** (2001) 361

[37] L. Greengard: *Science* **265** (1994) 909

APPENDIX

First Order Optimality Conditions

Theorem 1 *Consider the optimization problem*

$$\min_{\{\mathbf{R}_A\}} E_{tot} = E_e + E_{ext} + E_{nn},$$

subject to the constraint that for a nuclear configuration $\{\mathbf{R}_I\}$ the energy E_e is the electronic ground-state energy, and the electron density $\hat{\rho}$ that realizes this electronic ground energy additionally satisfies the charge constraint equation. Under these assumptions, the first-order optimality conditions for the optimization problem lead to

$$\mathbf{F}_K = \frac{\partial E_{ext}}{\partial \mathbf{R}_K} + \frac{\partial E_{nn}}{\partial \mathbf{R}_K} = \mathbf{0},$$

where $\mathbf{F}_K$ is interpreted as the force acting on nucleus K, and

$$E_{ext}(\mathbf{r}; \{\mathbf{R}_I\}) = -\sum_{A=1}^{M} \int \hat{\rho}(\mathbf{r})\, V_{ext}\left(\mathbf{r}; \{\mathbf{R}_I\}\right) d\mathbf{r} = -\sum_{A=1}^{M} \int \frac{Z_A \hat{\rho}(\mathbf{r})}{||\mathbf{r} - \mathbf{R}_A||}\, d\mathbf{r},$$

$$E_{nn} = \sum_{A=1}^{M} \sum_{B=A+1}^{M} \frac{Z_A Z_B}{||\mathbf{R}_B - \mathbf{R}_A||}.$$

Proof: The proof relies on the calculus of variations. Since $\hat{\rho}(\mathbf{r})$ is determined to minimize the electronic energy, there is a parametric dependency of this value on the ionic position: $\hat{\rho}(\mathbf{r}) = \rho(\mathbf{r}; \{\mathbf{R}_I\})$. After application of the chain rule, the optimality conditions for E_{tot} will read

$$\frac{\delta E_e}{\delta \rho} \frac{\partial \rho}{\partial \mathbf{R}_K} + \frac{\partial E_e}{\partial \mathbf{R}_K} + \frac{\partial E_{nn}}{\partial \mathbf{R}_K} = 0, \tag{6}$$

where $\mathbf{R}_K$ is the position of an arbitrary nucleus K.

The optimality conditions for minimizing the electronic energy as a functional of the electron density lead to

$$\frac{\delta E_e}{\delta \rho} + \lambda \frac{\delta g}{\delta \rho} = 0, \tag{7}$$

where λ is the Lagrange multiplier associated with the constraint

$$g[\rho] = 0$$

that the electron density must satisfy. For the problem at hand the charge conservation equation results in $g[\rho] = \int \rho(\mathbf{r})\, \mathrm{d}\mathbf{r} - N_e$. Therefore, the variation of $\rho(\mathbf{r})$ with respect to $\mathbf{R}_K$ must satisfy

$$\frac{\delta g}{\delta \rho} \frac{\partial \rho}{\partial \mathbf{R}_K} = 0.$$

Multiplying Eq. (7) from the right by $\frac{\partial \rho}{\partial \mathbf{R}_K}$ leads to $\frac{\delta E_e}{\delta \rho} \frac{\partial \rho}{\partial \mathbf{R}_K} = 0$, which, substituted back into (6) yields the desired optimality condition and thus completes the proof. Therefore, for each nucleus K in the system, the first order optimality condition leads to

$$\int \hat{\rho}(\mathbf{r}) \frac{\mathbf{r} - \mathbf{R}_K}{||\mathbf{r} - \mathbf{R}_K||^3}\, \mathrm{d}\mathbf{r} + \sum_{A=1, A\neq K}^{M} Z_A \frac{\mathbf{R}_A - \mathbf{R}_K}{||\mathbf{R}_A - \mathbf{R}_K||^3} = \mathbf{0}.$$

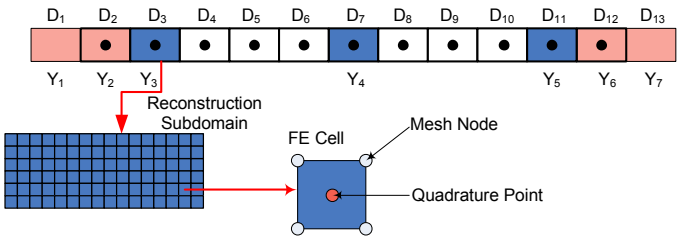


Figure 1: Partitioning of the computational domain: $\mathtt{D}_3$, $\mathtt{D}_7$ and $\mathtt{D}_{11}$ reconstruction subdomains; $\mathtt{D}_1$, $\mathtt{D}_2$, $\mathtt{D}_{12}$, and $\mathtt{D}_{13}$ active subdomains; $\mathtt{D}_4$, $\mathtt{D}_5$, $\mathtt{D}_6$, $\mathtt{D}_8$, $\mathtt{D}_9$, and $\mathtt{D}_{10}$ passive subdomains.

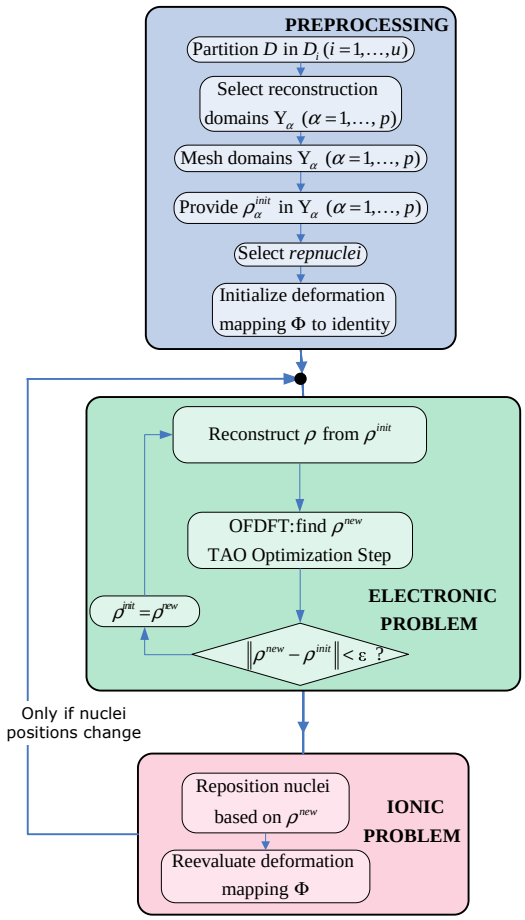
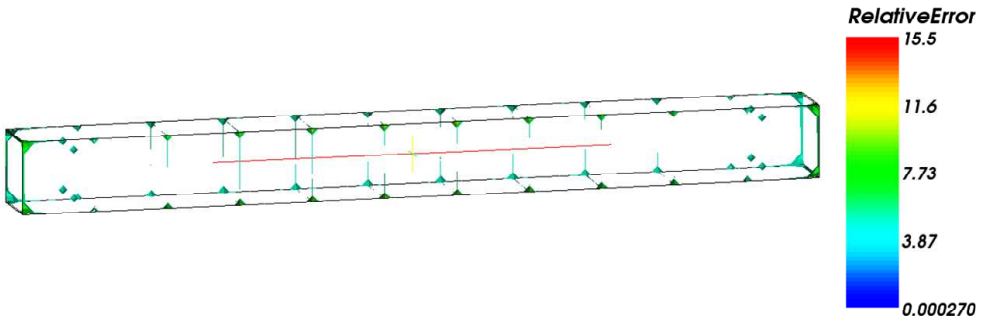
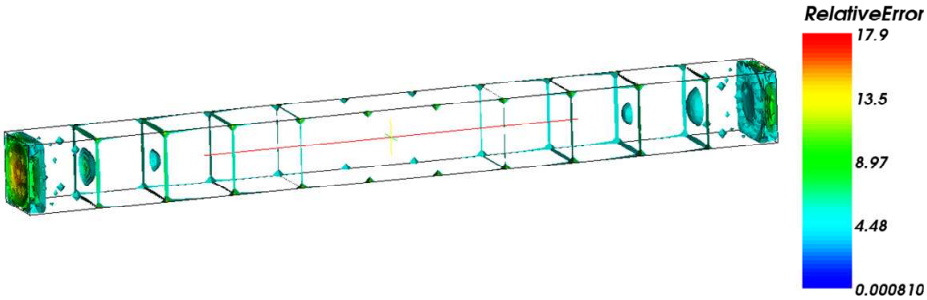


Figure 2: Computational flow.

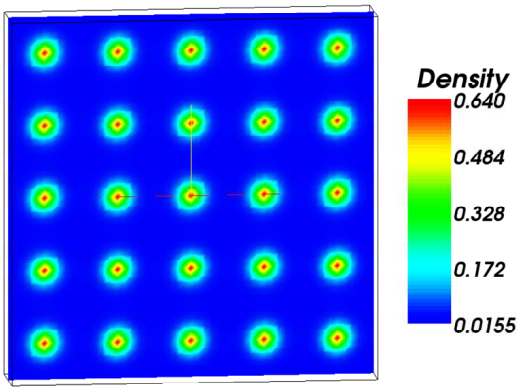


(a) Relative error for 13/7 (245)



(b) Relative error for 13/5 (221)

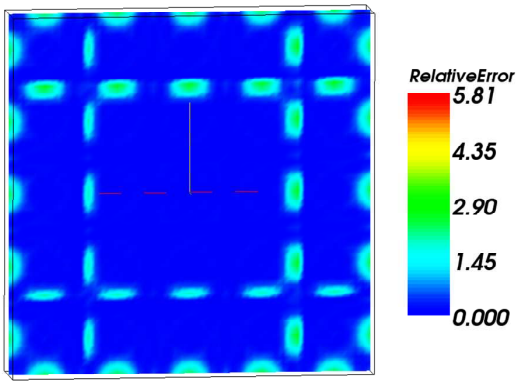
Figure 3: Relative error surface for the 13-subdomain scenarios using (a) 7 and (b) 5 active subdomains. In parentheses we show the number of optimization iterations.



(a) Electronic distribution for 25/25

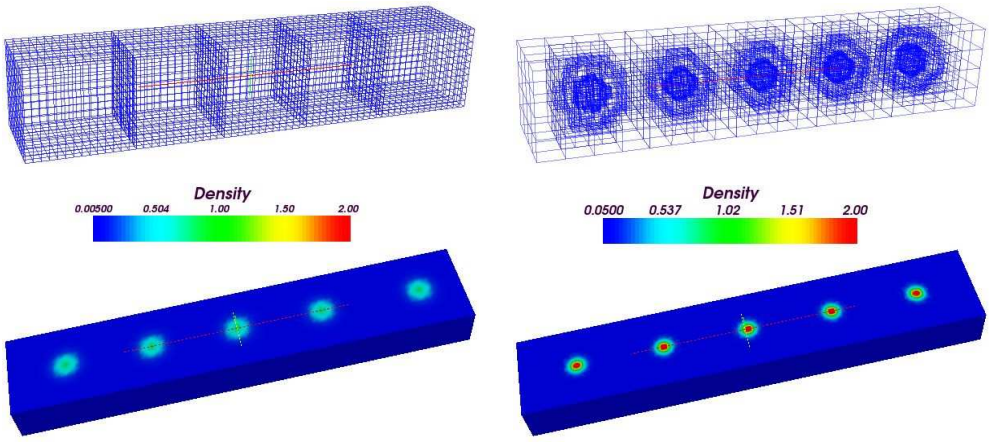
D_{21}	D_{22}	D_{23}	D_{24}	D_{25}
D_{16}	D_{17}	D_{18}	D_{19}	D_{20}
D_{11}	D_{12}	D_{13}	D_{14}	D_{15}
D_{6}	D_{7}	D_{8}	D_{9}	D_{10}
D_{1}	D_{2}	D_{3}	D_{4}	D_{5}

(b) Domain setup



(c) Relative error for 9/25

Figure 4: 5×5 slab simulation scenario results.



(a) fixed mesh (-5.1829/212) (b) nonuniform mesh (-5.8705/2181)

Figure 5: Density distribution for the 5-subdomain example using (a) a uniform mesh and (b) an adaptive mesh. Above each result we show the associated mesh. In parentheses we give the total energy/number of iterations.