Linear-Scaling Density Functional Theory with Tens of Thousands of Atoms: ONETEP

Nicholas D.M. Hine1, Peter D. Haynes1
1. Blackett Laboratory and Thomas Young Centre, Imperial College London, Exhibition Road, London SW7 2AZ, U.K.

Abstract

We present recent improvements to the ONETEP code. ONETEP is an all-electron electronic-structure package for total energy calculations within density-functional theory. Its main distinguishing features are true linear scaling, in that the total computational effort scales only linearly with system size, and plane-wave accuracy, in that the convergence of the total energy is systematically improvable by increasing cutoffs. We present recent improvements to the parallel performance of the code, and thus in effect considerable increases in the scope and scale of feasible calculations with ONETEP, especially in solids. On parallel computers comprising large clusters of commodity servers, recent improvements make calculations of tens of thousands of atoms in a solid possible even for small numbers of cores (10–100). Efficient scaling with number of atoms is demonstrated up to 32,768 atoms on 64 cores, and efficient scaling with number of cores is demonstrated up to 512 cores for 32,768 atoms.

ONETEP Theory

ONETEP was developed from the beginning as a parallel code and its efficient scaling and performance on isolated molecules, nanotubes and similar systems with a high degree of sparsity has been well-documented. Recent work has focused on improving performance in solids, where communications bottlenecks in large systems has previously limited the useful applicability of the code.

Matrix Algebra

One time-limiting component of ONETEP calculations is sparse matrix algebra, especially during kernel optimization. As Fig 3 shows, the pattern of filling of the sparse matrices representing 

and 

can be highly structured, allowing considerable optimization of the communication and computation patterns. Recent improvements include:

- Dense matrix algebra to replace sparse algebra in small systems or systems such as metals where local truncation is not possible. See Fig 4.
- Reordered, non-blocking symmetric commutative kernels to allow different node-to-node pairings to take different lengths of time.
- Reduced total volume by communicating only blocks of multiphase contributions to matrix products.
- Loop-unrolled, block multiplication hand-coded for common block sizes (1,4,9).

Combined, these developments have dramatically improved both the speed and scaling (with system size and number of parallel processes) of matrix algebra.

Parallel Optimisation

ONETEP combines an O(N) scaling with 'plane-wave' accuracy, in that the convergence of the total energy is systematically improvable by increasing cutoffs. The localized basis in ONETEP comprises 'Nonorthogonal Generalized Wannier Functions' (NGWFs) expressed in terms of a basis of periodic host-like limited delta functions, or plane functions (see Fig 2) strictly localised to spherical regions of radius \( R_{\alpha} \). These plane functions, with coefficients \( C_{\alpha} \), are centered on the grid points \( \mathbf{r}_i \) of a regular grid specified by a plane-wave cutoff energy \( E_{\text{cut}} \).

The minimisation of the energy occurs via nested loops. The outer loop minimizes the energy with respect to the coefficients \( C_{\alpha} \)

while inner loop, performed at fixed \( C_{\alpha} \), minimizes the energy with respect to the kernel elements \( K_{\alpha\beta} \)

(8)

Recent Applications

GaAs Nanorods

'Self-assembly' is a promising route to constructing working nanotube devices. Water-soluble GaAs nanowires which display spontaneous polarisation due to the polar bonding and lack of inversion symmetry, have been observed forming a variety of self-assembled structures. Linear-scaling DFT with ONETEP allows a window on the complex interplay between bonding and long-range electrostatic effects (which can be treated with cutoff convolutions) interacted to model these systems.

Defect Formation Energies

Understanding defects and defect clusters in crystalline materials is a tough challenge for electronic structure methods due to the requirement of embedding (often large) localised systems in periodic hosts. As a very simple example, even formation energy of the comparitively simple neutral vacancy in Germanium can be treated only at the local density approximation level. Therefore, while the bulk crystal is an insulator, there is a donor centre at the Fermi level in an unsmoothed defect superrl. In small cells with a traditional DFT approach, this manifests itself as a band whose occcupation varies as a function of \( k \) and may require a folding of occupation numbers to provide results converged with respect to supercell size.

References

Improvements to the ONETEP code have lowered the 'computation point', the number of atoms beyond which a linear scaling algorithm becomes required to provide useful results. This has already been achieved for (100) atoms for a self-doped system such as a nanotube, but it has now been brought down to the regime of feasible calculations for solids as well. For example, in silicon semiconductors such as Si, Ge and III–Vs, we observe the crossover point relative to CASTEP to be around 600 – 1000 atoms. We have also demonstrated the efficiency and accuracy of the code to two important applications: self-assembly in GaAs nanowires and defect formation energies in semiconductors.