Calcul intensif dans les matériaux : exemple du code ab initio BigDFT

L. Genovese, B. Videau, T. Deutsch

L_Sim – CEA Grenoble

October 17, 2017
Atomistic Simulations

Material Science – Chemistry

- Theory – Experiment – Simulation
- Hardware – Computers
- Algorithms

- Atomistic Simulations
  - Force fields (Molecular Dynamics)
  - Semi-Empirical (tight-binding) methods
  - Density Functional Theory
  - Quantum Chemistry (wavefunction-based methods)
  - Quantum Monte-Carlo

- More than 40,000 DFT users
Atomistic Simulations

Material Science – Chemistry

- Theory – Experiment – Simulation
- Hardware – Computers
- Algorithms

- Atomistic Simulations
  - Force fields (Molecular Dynamics)
  - Semi-Empirical (tight-binding) methods
  - Density Functional Theory
  - Quantum Chemistry (wavefunction-based methods)
  - Quantum Monte-Carlo

- More than 40,000 DFT users
**Large use of the Density Functional Theory**

**RESEARCH**

**RESEARCH ARTICLE SUMMARY**

**DFT METHODS**

Reproducibility in density functional theory calculations of solids


INTRODUCTION: The reproducibility of results is one of the underlying principles of science. An observation can only be accepted by the scientific community when it can be confirmed by independent studies. However, reproducibility does not come easily. Recent works have painfully exposed cases where previous conclusions were not upheld. The scrutiny of the scientific community has also turned to research involving computer programs, finding that reproducibility depends more strongly on implementation than commonly thought. These problems are especially relevant for property predictions of crystals and molecules, which hinge on precise computer implementations of the governing equation of quantum physics.

RATIONALE: This work focuses on density functional theory (DFT), a particularly popular quantum method for both academic and industrial applications. More than 60 000 DFT papers are published each year, and DFT is now increasingly used in an automated fashion to build large databases or apply multiple techniques with limited human supervision. Therefore, the reproducibility of DFT results underlies the scientific credibility of a substantial fraction of current work in the natural and engineering sciences. A plethora of DFT computer codes are available, many of them differing considerably in their details of implementation, and each yielding a certain "precision" relative to other codes. How is one to decide which of these codes will produce the correct result, and which does not? We devised a procedure to assess the precision of DFT methods and used this to demonstrate reproducibility among many of the most widely used DFT codes. The essential part of this assessment is a pairwise comparison of a wide range of methods with respect to their predictions of the equations of state of the elemental crystals. This effort required the combined expertise of a large group of code developers and expert users.

RESULTS: We calculated equation-of-state data for four classes of DFT implementations, totaling 40 methods. Most codes agree very well, with pairwise differences that are comparable to those between different high-precision experiments. Even in the case of pseudopotential approaches, which largely depend on the atomic potentials used, a similar precision can be obtained if using the full potential. The remaining deviations are due to subtle effects, such as specific numerical implementations or the treatment of relativistic terms.

CONCLUSION: Our work demonstrates that the precision of DFT implementations can be determined, even in the absence of one absolute reference code. Although this was not the case 5 to 10 years ago, most of the commonly used codes and methods are now found to predict essentially identical results. The established precision of DFT codes not only ensures the reproducibility of DFT predictions but also puts several past and future developments on a firmer footing. Any newly developed methodology can now be tested against the benchmark to verify whether it reaches the same level of precision. New DFT applications can be shown to have used a sufficiently precise method. Moreover, high-precision DFT calculations are essential for developing improvements to DFT methodology, such as new density functionals, which may further increase the predictive power of the simulations.

The list of author affiliations is available in the full article online.

*Corresponding author. E-mail: stefan.cottenier@rug.be (S.C.)

On our website: Read the full article at http://dx.doi.org/10.1126/science.aad3000

Read the full article online: http://dx.doi.org/10.1126/science.aad3000

**Number of publications per year (1975-2014) on topics "density functional" or "DFT". From Mavropoulos, 2015.**
Outline

1. Introduction: Ab initio and wavelet
2. BigDFT code
3. Parallelization (MPI + OpenMP)
4. GPU acceleration
5. Linear scaling
6. Perspectives
Ab initio calculations with DFT

**Several advantages**

- **Ab initio:** No adjustable parameters
- **DFT:** Quantum mechanical (fundamental) treatment

**Main limitations**

- Approximated approach (electron correlations)
- Requires high computer power, limited to few hundreds atoms in most cases

**Wide range of applications**

Nanoscience, biology, materials
Finding the most stable atomic configuration

Need a large number of calculations

Different methods linked with BigDFT:
- Minima Hopping (S. Goedecker, Basel)
- Activation-Relaxation Technique (N. Mousseau, Montreal)

Applied on different systems

Benefit from high flexibility and performances

Interesting for potential synthesis pathways
A basis for nanosciences: the BigDFT project

STREP European project: BigDFT(2005-2008)
Four partners, 15 contributors:

Aim: To develop an ab-initio DFT code based on Daubechies Wavelets for large system calculations, distributed freely (GNU-GPL license)

References

“Daubechies wavelets as a basis set for density functional pseudopotential calculations”,

“Daubechies wavelets for linear scaling density functional theory”, S. Mohr, L. Genovese, T. Deutsch,
Goal

Wavelets
an ideal basis for electronic structure calculations – flexible, systematic etc.

(Linear-scaling) DFT
allows us to access very large system sizes via the use of a localized minimal basis set

we want to combine the two...
Goal

Massively parallel architectures

... and run calculations on large, realistic systems, using massively parallel machines
Why do we use wavelets in BigDFT?

Adaptivity

One grid, two resolution levels in BigDFT:

- 1 scaling function ("coarse region")
- 1 scaling function and 7 wavelets ("fine region")

- Ideal for big inhomogeneous systems
- Efficient Poisson solver, capable of handling different boundary conditions – free, wire, surface, periodic
- Established code with many capabilities
A brief description of wavelet theory

Two kind of basis functions

A Multi-Resolution real space basis
The functions can be classified following the resolution level they span.

Scaling Functions
The functions of low resolution level are a linear combination of high-resolution functions

\[ \phi(x) = \sum_{j=-m}^{m} h_j \phi(2x - j) \]

Centered on a resolution-dependent grid: \( \phi_j = \phi_0(x - j) \).
A brief description of wavelet theory

Wavelets

They contain the DoF needed to complete the information which is lacking due to the coarseness of the resolution.

\[ \frac{1}{2} \phi(2x) = \sum_{j=-m}^{m} \tilde{h}_j \phi(x - j) + \sum_{j=-m}^{m} \tilde{g}_j \psi(x - j) \]

Increase the resolution without modifying grid space

SF + W = same DoF of SF of higher resolution

\[ \psi(x) = \sum_{j=-m}^{m} g_j \phi(2x - j) \]

All functions have compact support, centered on grid points.
Adaptivity of the mesh

Atomic positions (H$_2$O)
Adaptivity of the mesh

Fine grid (high resolution)
Adaptivity of the mesh

Coarse grid (low resolution)
Systematic basis set

Two parameters for tuning the basis
- The grid spacing
- The extension of the grid

Convergence of a methane molecule
## Deltatest benchmark: $\Delta = 1.0$

Science 351, 6280 (2016)

### For first three rows

<table>
<thead>
<tr>
<th>Code</th>
<th>Version</th>
<th>Basis</th>
<th>Electron treatment</th>
<th>$\Delta$-value</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>WIEN2k</td>
<td>13.1</td>
<td>LAPW/APW+lo</td>
<td>all-electron</td>
<td>0 meV/atom</td>
<td>S. Cottenier</td>
</tr>
<tr>
<td>BigDFT</td>
<td>1.7.6</td>
<td>Daubechies wavelets</td>
<td>HGHk-semicore and NLCC 2015 norm-conserving</td>
<td>0.1 meV/atom</td>
<td>BigDFT [11]</td>
</tr>
<tr>
<td>Elk</td>
<td>3.1.5</td>
<td>APW+lo</td>
<td>all-electron</td>
<td>0.2 meV/atom</td>
<td>Elk [14]</td>
</tr>
<tr>
<td>VASP</td>
<td>5.2.12</td>
<td>plane waves</td>
<td>PAW 2015 GW-ready (5.4)</td>
<td>0.2 meV/atom</td>
<td>K. Lejaeghere</td>
</tr>
<tr>
<td>Quantum ESPRESSO</td>
<td>5.1</td>
<td>plane waves</td>
<td>SSSP Accuracy mixed NC/US/PAW potential library</td>
<td>0.2 meV/atom</td>
<td>QuantumESPRESSO [12]</td>
</tr>
<tr>
<td>FLEUR</td>
<td>0.26</td>
<td>LAPW (+lo)</td>
<td>all-electron</td>
<td>0.2 meV/atom</td>
<td>FLEUR [9]</td>
</tr>
<tr>
<td>FHI-aims</td>
<td>081213</td>
<td>tier2 numerical orbitals</td>
<td>all-electron (relativistic atomic_zora scalar)</td>
<td>0.3 meV/atom</td>
<td>ASE [2]</td>
</tr>
</tbody>
</table>

Screenshot of DeltaTest webpage as of 24/02/16, elements up to Ar, new NLCC - HGH - NC - PSP (S. Saha)

Laboratoire de Simulation Atomistique  http://inac.cea.fr/L_Sim T. Deutsch
A code **both** for Solid-State and Quantum Chemistry

- 3D periodic, Surfaces and Free BC (← Poisson Solver)
- Very high precision (analytic KS operators)
- Usage of analytic HGH pseudopotentials
- AE accuracy, benchmarked in G2-1, S22, DeltaTest

**Present functionalities**

Traditional functionalities for GS Kohn-Sham DFT (including metals, Hybrid Functionals), LR-TDDFT, empirical VdW

Exhaustive library of Structural Prediction, \(O(N)\) calculations

**Under implementation**

Non orthorhombic cells, Systems embedded in electrostatic environments

Resonant States extraction

JCP 144, 014103 (2016),

Laboratoire de Simulation Atomistique  http://inac.cea.fr/L_Sim  T. Deutsch
**BigDFT**

**Overview**

- Code
- Bugs
- Blueprints
- Translations
- Answers

Registered 2013-02-13 by 🚀 Luigi Genovese

**BigDFT** is an ab initio code based on Daubechies wavelets.

Such functions have features which make them a powerful and promising basis set for application in materials science. These are a compact support multiresolution basis, and form one of the few examples of systematic real space basis sets. They are an optimal basis for expanding localised information. The real space description they provide allows to build an efficient, clean method to treat systems in complex environments, like surfaces geometries or system with a net charge. The mathematical properties of the formalism are optimal to build a robust, highly optimised code, conceived for systems of few hundred atoms, with excellent efficiency on parallel computers.

BigDFT is a DFT massively parallel electronic structure code using a wavelet basis set with the capability to use a linear scaling method. Wavelets form a real space basis set distributed on an adaptive mesh (two levels of resolution in our implementation).

GTH or GHG pseudopotentials are used to remove the core electrons.

The code BigDFT is available in ABINIT v5.5 and higher but can also be downloaded in a standalone version from the website.

Thanks to our Poisson solver based on a Green function formalism, periodic systems, surfaces and isolated systems can be simulated with explicit boundary conditions.

The Poisson solver can also be downloaded and used independently and is integrated in ABINIT, Octopus and CP2K.

The code is free software, available under GNU-GPL license and the BigDFT developer community encourages anyone willing to contribute to join the team.

🔗 Change branding

違反 Home page 📚 Wiki
Each section of BigDFT is, when appropriate, defined as a module with its own build system and compilation instructions. At present:

- FUTILE 1.0
- CheSS 1.0
- PSolver 1.8
Completely separated from BigDFT code

**Chebyshev Sparse Solvers**  Fundamental for $O(N)$ BigDFT
- Can be built completely independently from BigDFT
- Comes with a set of tests and examples
- Integration in ELSI package (PEXSI, OMM, ELPA,...)

Comparison with PEXSI in BigDFT

<table>
<thead>
<tr>
<th>Matrix Size</th>
<th>Spectral Width (eV)</th>
<th>Sparsity S (%)</th>
<th>Sparsity H (%)</th>
<th>Sparsity K (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Runtime (seconds)</th>
<th>Matrix Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>1000</td>
</tr>
<tr>
<td>100</td>
<td>1000</td>
</tr>
<tr>
<td>150</td>
<td>1000</td>
</tr>
<tr>
<td>200</td>
<td>1000</td>
</tr>
</tbody>
</table>

CheSS  PEXSI
Outline

1. Introduction: Ab initio and wavelet
2. BigDFT code
3. Parallelization (MPI + OpenMP)
4. GPU acceleration
5. Linear scaling
6. Perspectives
Massively parallel (MPI + OpenMP)

Two kinds of parallelisation
- By orbitals (Hamiltonian application, preconditioning)
- By components (overlap matrices, orthogonalisation)

A few (but large) packets of data
More demanding in bandwidth than in latency
- Better data locality (Hamiltonian application and orthonormality)
- Optimal speedup (eff. $\sim 85\%$), also for big systems

Cubic scaling code
For systems bigger than 500 atomes (1500 orbitals) : orthonormalisation operation is predominant ($N^3$)
Orbital distribution scheme

Used for the application of the hamiltonian

The hamiltonian (convolutions) is applied separately onto each wavefunction

\[ \psi_1 \]
\[ \psi_2 \]
\[ \psi_3 \]
\[ \psi_4 \]
\[ \psi_5 \]

\{ MPI 0 \}
\{ MPI 1 \}
\{ MPI 2 \}
Coefficient distribution scheme

Used for scalar product & orthonormalisation

BLAS routines (level 3) are called, then result is reduced

<table>
<thead>
<tr>
<th></th>
<th>MPI 0</th>
<th>MPI 1</th>
<th>MPI 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>ψ₁</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ψ₂</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ψ₃</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ψ₄</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ψ₅</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Communications are performed via **MPI_ALLTOALLV**
Using GPUs in a Big systematic DFT code

Nature of the operations

- Operators approach via convolutions
- Linear Algebra due to orthogonality of the basis
- Communications and calculations do not interfere

A number of operations which can be accelerated

Evaluating GPU convenience

Three levels of evaluation

1. Bare speedups: GPU kernels vs. CPU routines
   Does the operations are suitable for GPU?

2. Full code speedup on one process
   Amdahl’s law: are there hot-spot operations?

3. Speedup in a (massively?) parallel environment
   The MPI layer adds an extra level of complexity
No Hot-spot operations

Different code sections can be ported on GPU up to 20x speedup for some operations, 7x for the full parallel code.
No Hot-spot operations

Different code sections can be ported on GPU up to **20x speedup** for some operations, 7x for the full parallel code.

Bull-Fourier 2009 Prize


Use of Titan (Oak Ridge) with 18,000 GPUs
### Sample BigDFT run: Graphene, 4 C atoms, 52 kpts

No. of Flop: $8.053 \cdot 10^{12}$

<table>
<thead>
<tr>
<th></th>
<th>MPI</th>
<th>1</th>
<th>1</th>
<th>4</th>
<th>1</th>
<th>4</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPU</td>
<td>NO</td>
<td>NV</td>
<td>NV</td>
<td>ATI</td>
<td>ATI</td>
<td>NV + ATI</td>
<td></td>
</tr>
<tr>
<td>Time (s)</td>
<td>6020</td>
<td>300</td>
<td>160</td>
<td>347</td>
<td>197</td>
<td>109</td>
<td></td>
</tr>
<tr>
<td>Speedup</td>
<td>1</td>
<td>20.07</td>
<td>37.62</td>
<td>17.35</td>
<td>30.55</td>
<td>55.23</td>
<td></td>
</tr>
<tr>
<td>GFlop/s</td>
<td>1.34</td>
<td>26.84</td>
<td>50.33</td>
<td>23.2</td>
<td>40.87</td>
<td>73.87</td>
<td></td>
</tr>
</tbody>
</table>

Supercomputer in a workstation

---

[www.bigdft.org](http://www.bigdft.org)

Introduction

BigDFT code

MPI/OpenMP

GPU

Linear

Perspectives

---

[Labouratoire de Simulation Atomistique](http://inac.cea.fr/L_Sim)

T. Deutsch
PS is used for the Exact Exchange operator
Recently accelerated via GPU
conjunction with GPU-Direct (improves communications)
PS is used for the Exact Exchange operator
Recently accelerated via GPU
conjunction with GPU-Direct (improves communications)
PS is used for the Exact Exchange operator

Recently accelerated via GPU
conjunction with GPU-Direct (improves communications)
PS is used for the Exact Exchange operator
Recently accelerated via GPU conjunction with GPU-Direct (improves communications)
Hybrid Functionals for large systems (2016) $\gamma=$PBE0/PBE

**UO$_2$ systems:**

<table>
<thead>
<tr>
<th>Atoms</th>
<th>Orbitals</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>200</td>
</tr>
<tr>
<td>96</td>
<td>1432</td>
</tr>
<tr>
<td>324</td>
<td>5400</td>
</tr>
<tr>
<td>768</td>
<td>12800</td>
</tr>
<tr>
<td>1029</td>
<td>17150</td>
</tr>
</tbody>
</table>

**324 atoms system (1800 processes)**

<table>
<thead>
<tr>
<th>Timing for 2 SCI (seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1400</td>
</tr>
<tr>
<td>1200</td>
</tr>
<tr>
<td>1000</td>
</tr>
<tr>
<td>800</td>
</tr>
<tr>
<td>600</td>
</tr>
<tr>
<td>400</td>
</tr>
<tr>
<td>200</td>
</tr>
</tbody>
</table>

**Perspectives**

- GPU
- Linear

www.bigdft.org

Introduction

BigDFT code

MPI/OpenMP
Scaling of BigDFT

“Traditional” BigDFT code

We can reach systems containing up to a few hundred electrons thanks to wavelet properties and efficient parallelization: (MPI + OpenMP + GPU)

Varying the number of atoms $N$

DFT operations scale differently:
- $O(N \log N)$: Poisson solver
- $O(N^2)$: convolutions
- $O(N^3)$: linear algebra

and have different prefactors:
- $c_{O(N^3)} \ll c_{O(N^2)} \ll c_{O(N \log N)}$

For bigger systems the $O(N^3)$ will dominate

(first) motivation for a new approach
Local orbitals and linear scaling

**KS orbitals**
Linear combinations of support functions $\phi_\alpha(r)$:

$$\Psi_i(r) = \sum_\alpha c_i^\alpha \phi_\alpha(r)$$
- localized around atoms
- expanded in wavelets
- optimized in-situ

**Density Matrix**
Defined via the kernel $K^{\alpha\beta}$ in the $\phi_\alpha(r)$ basis:

$$\rho(r, r') = \sum_i f_i \Psi_i(r)\Psi_i(r') = \sum_{\alpha, \beta} \phi_\alpha(r) K^{\alpha\beta} \phi_\beta(r')$$

Extended Kohn-Sham orbitals
- cubic scaling, high accuracy

Localized support functions (LCAO)
- linear scaling, low accuracy

Localized adaptive support functions
- linear scaling, high accuracy

**Localized form for the Density and the Hamiltonian**
Localization $\rightarrow$ sparse matrices $\rightarrow O(N)$
Comparison with the cubic version

Energy and forces with accuracy of a systematic approach

- 20 min for 18,000 atoms
- CPU Time and memory \( \propto \) number of atoms
- Precise DFT computing for thousands of atoms at Institute-Scale \((10^2 \rightarrow 10^3\) CPU cores)

Different levels of precision \( \Rightarrow \) cutoff radii

Without fine-tuning converges to absolute energy differences of the order of 10 meV/atom, and almost exact forces.

High flexibility, like the cubic code

- Charged systems, various BC (free, surfaces, periodic)
- System sizes: 100 - 30K atoms \( \sim 100\) k Basis functions
Recent improvements

Algorithm is robust and reliable on a variety of systems

**Accurate and efficient linear scaling DFT calculations with universal applicability**

S. Mohr, L. E. Ratcliff, L. Genovese, D. Caliste, P. Boulanger, S. Goedecker and T. Deutsch


DOI: 10.1039/c5cp00437c

Included in the Real-space numerical grid methods in quantum chemistry themed issue of *PCCP*

Guest-edited by Luca Frediani (The Arctic University of Norway) and Dage Sundholm (University of Helsinki)
Why Large Scale DFT?

Present-day situation

- Quantum Chem. Methods
- Hybrid DFT
- DFT
- LS-DFT
- FMO

- CHEMISTRY
- MATERIALS SCIENCE
- BRIDGING THE LENGTHSCALE GAP
- BIOLOGY

- QM
- MM
- QM/MM

- Force Fields
- Coarse-graining
- Model - MD

- www.bigdft.org
Summary and future directions

Linear-Scaling DFT calculations based on wavelets

- Robust convergence, high accuracy and flexibility (BC)
- Reduction in degrees of freedom → large systems via moderate sized machines (∼ TFlop/s) Lab-scale
- Optimal mapping between KS DoF and atoms
- Different level of descriptions (controlling the precision)
  QM ⊃ Fragments ⊃ Atomic charges
- Opens up new possibilities

Challenges and future directions

- Explore interplay environment ↔ electronic excitations (CDFT, QM/MM, statistics…)
- Provide high quality back end for extraction of atomic multipoles from QM calculations
- Towards a control of the level of theory (QM/QM)
Order N methods  Stephan Mohr, Laura E. Ratcliff, Paul Boulanger

Group of Stefan Goedecker  B. Schaefer, A. Ghazemi, S. Saha, G. Fisicaro, A. Degomme, J. Flores-Livas (Basel University)

Link with ABINIT and python bindings  Damien Caliste (CEA)

Resonant states  M. Morinière, I. Duchemin (CEA), B. Nectoux, E. Cancés (CERMICS)

Optimized convolutions  B. Videau, J.-F. Méhaut (LIG, computer scientists, Grenoble)