Introduction to the SIESTA code

Javier Junquera

UC
UNIVERSIDAD
DE CANTABRIA

Most important references followed in this lecture

INSTITUTE OF PHYSICS PUBLISHING — JOURNAL OF PHYSICS: CONDENSED MATTER

J. Phys.: Condens. Matter **14** (2002) 2745–2779 — PII: S0953-8984(02)30737-9

The SIESTA method for *ab initio* order-N materials simulation

José M Soler[1], Emilio Artacho[2], Julian D Gale[3], Alberto García[4], Javier Junquera[1,5], Pablo Ordejón[6] and Daniel Sánchez-Portal[7]

SIESTA: Recent developments and applications

Cite as: J. Chem. Phys. **152**, 204108 (2020); doi: 10.1063/5.0005077
Submitted: 18 February 2020 • Accepted: 20 April 2020 •
Published Online: 27 May 2020

View Online — Export Citation — CrossMark

Alberto García,[1,a] Nick Papior,[2,b] Arsalan Akhtar,[3,c] Emilio Artacho,[4,5,6,7,d] Volker Blum,[8,9,e] Emanuele Bosoni,[1,f] Pedro Brandimarte,[5,g] Mads Brandbyge,[10,h] J. I. Cerdá,[11,i] Fabiano Corsetti,[4,j] Ramón Cuadrado,[3,k] Vladimir Dikan,[1,l] Jaime Ferrer,[12,13,m] Julian Gale,[14,n] Pablo García-Fernández,[15,o] V. M. García-Suárez,[12,13,p] Sandra García,[3,q] Georg Huhs,[16,r] Sergio Illera,[3,s] Richard Korytár,[17,t] Peter Koval,[18,u] Irina Lebedeva,[4,v] Lin Lin,[19,20,w] Pablo López-Tarifa,[21,x] Sara G. Mayo,[22,y] Stephan Mohr,[16,z] Pablo Ordejón,[3,aa] Andrei Postnikov,[23,ab] Yann Pouillon,[15,ac] Miguel Pruneda,[3,ad] Roberto Robles,[21,ae] Daniel Sánchez-Portal,[5,21,af] Jose M. Soler,[22,24,ag] Rafi Ullah,[4,25,ah] Victor Wen-zhe Yu,[8,ai] and Javier Junquera[15,aj]

30 years of SIESTA

What is SIESTA?

(http://www.icmab.es/siesta)

Method and its implementation in a **simulation code of atomic systems**

Solves numerically the **quantum mechanic equations**
that rule the behaviour of the **electrons**

Allows the determination of the **atom dynamics** and
simulate physical and chemical processes that
happen at the atomic scale

EVERYTHING FROM FIRST-PRINCIPLES

First-principles calculations are free of parameters but not free of approximations

Use a set of "accepted" approximations
to solve the corresponding equations on a computer

NO EMPIRICAL INPUT

Quantum mechanics enables the study of materials at the atomic level without experiments

What are the main approximations?

Born-Oppenhaimer

Decouple the movement of the electrons and the nuclei.

Treatment of electron-electron interactions.

Wavefunction theory and density functional theory (DFT)

Pseudopotentials

Treatment of the (nuclei + core) — valence.

Basis set

Restrict the electronic wave function to the space of linear combination of a finite number of basis functions

Numerical evaluation of matrix elements

Efficient and self-consistent computations of Hamiltonian and Overlap matrices.

Supercells

To deal with periodic systems

What makes SIESTA different?
Efficiency

SIESTA is a first-principles code, based on Density Functional Theory…

…as many others

Wien-2K	VASP	FHI-Aims	Gaussian
ELK	Abinit	CP2K	CRYSTAL
	Quantum Espresso	CONQUEST	ADF
	CASTEP	PLATO	Qchem
	GPAW	Open-MX	Nwchem
	Octopus	DMOL	Turbomol
		BigDFT	
		Quantum ATK	

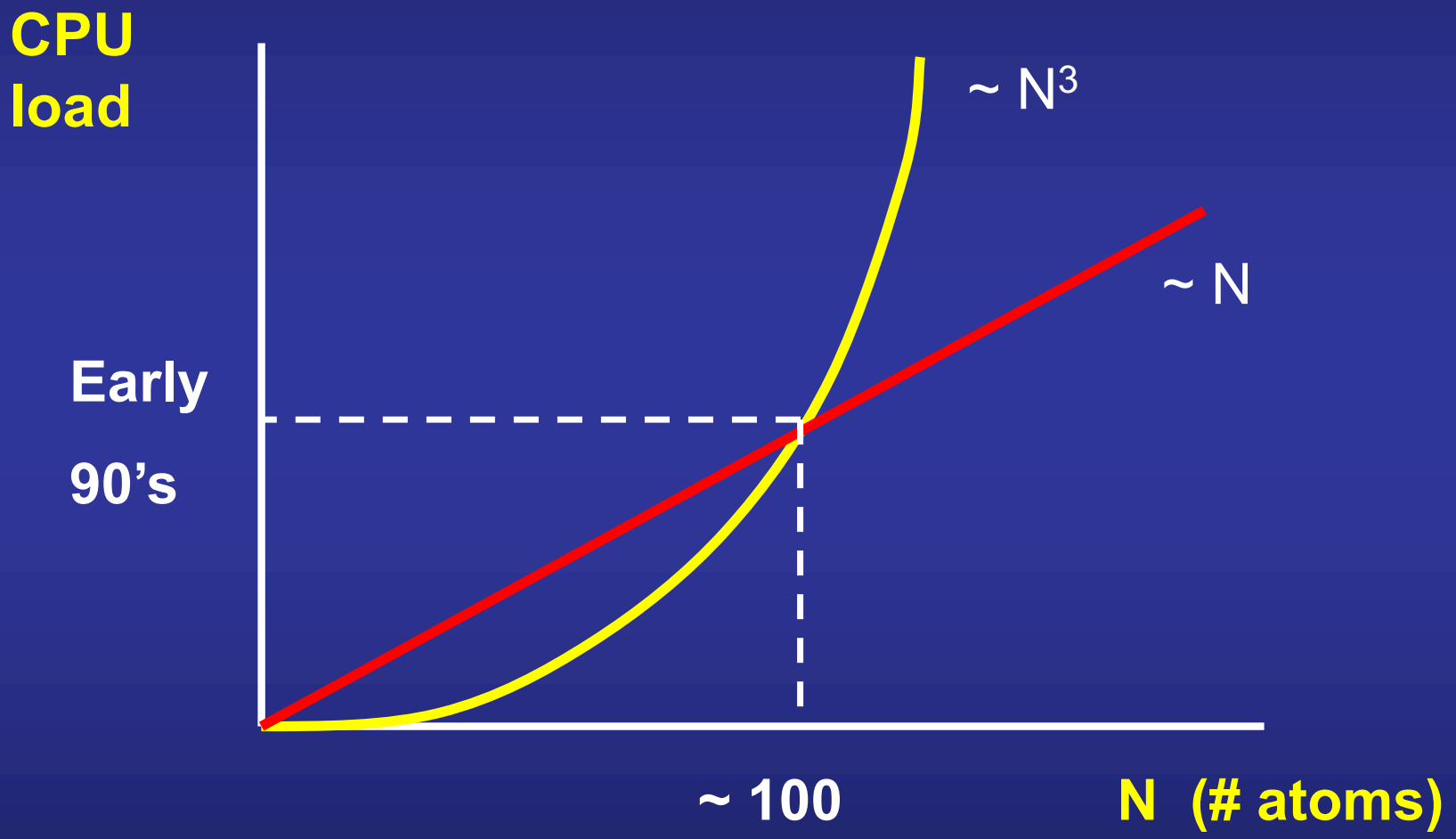
(And many more, apologize if some are missing here…)

Aim from inception: Efficiency

To do larger simulation boxes in modest computational platforms

Pioneer fully self-consistent LINEAR-SCALING DFT code (or Order-N)

Order-N methods: The computational load scales linearly with the system size



G. Galli and M. Parrinello, Phys. Rev Lett. 69, 3547 (1992)

Based on atomic-like orbitals as basis sets (LCAO: Linear Combination of Atomic Orbitals)

$$\psi_i(\vec{r}) = \sum_{\mu} \phi_\mu(\vec{r}) c_{\mu i} \qquad \mu \equiv \{Ilmn\}$$

$$\phi_{Ilmn}(\vec{r}) = R_{Iln}(|\vec{r}_I|) Y_{lm}(\hat{r}_I)$$

ADVANTAGES

- **Very efficient**

(number of basis functions needed is usually very small).

Rule of thumb: 3-5 functions per electron vs ~100 PW per electron

- **Large reduction of CPU time and memory**
- **Straightforward physical interpretation (population analysis, projected density of states,…)**

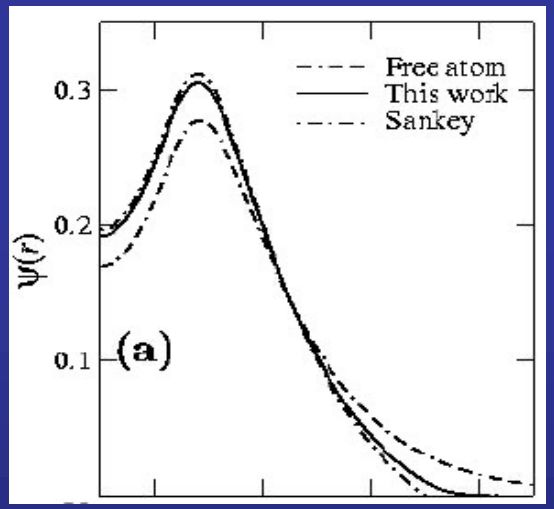
DISADVANTAGES

- **…Lack of systematic for convergence (not unique way of enlarge the basis set)**
- **Human and computational effort searching for a good basis set before facing a realistic project**
- **Responsibility on the user**

Finite support basis functions: strictly localized numerical atomic orbitals

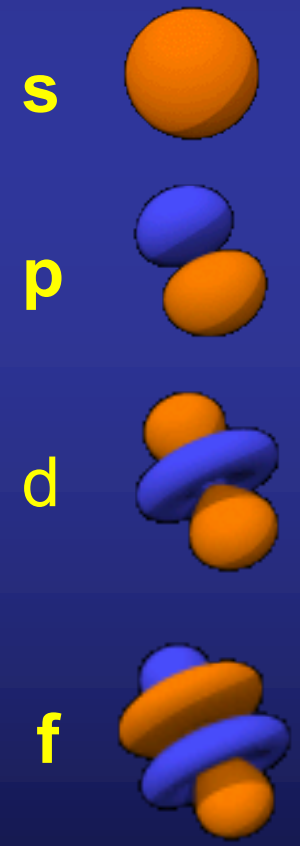
$$\phi_{Ilmn}(\vec{r}) = R_{Iln}(|\vec{r}_I|)\, Y_{lm}(\hat{r}_I)$$

Radial part:
degree of freedom to play with



Following Sankley and Niklewsky
Phys. Rev. B 40, 3979 (1989)

Spherical harmonics:
well defined (fixed) objects



Converging the basis size: from quick and dirty to highly converged calculations

Single-ζ (minimal or SZ)

One single radial function per angular momentum shell occupied in the free–atom

Improving the quality

Radial flexibilization:

Add more than one radial function within the same angular momentum than SZ

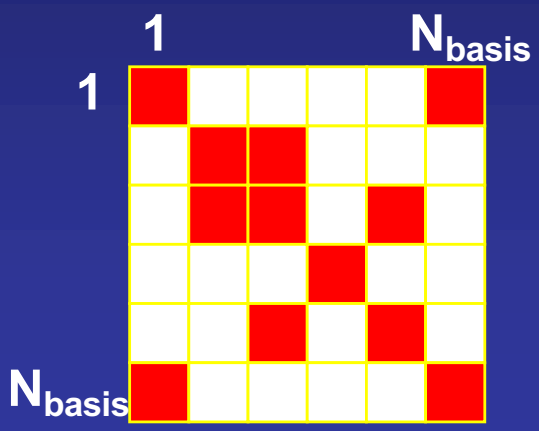
Multiple-ζ

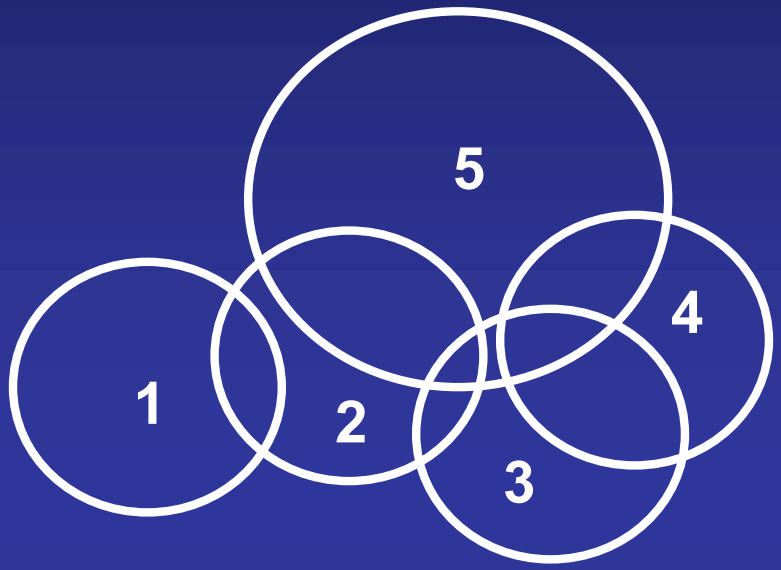
Angular flexibilization:

Add shells of different atomic symmetry (different l)

Polarization

Order-N methods rely heavily on the sparsity of the Hamiltonian and overlap matrices





1 with 1 and 2

2 with 1,2,3, and 5

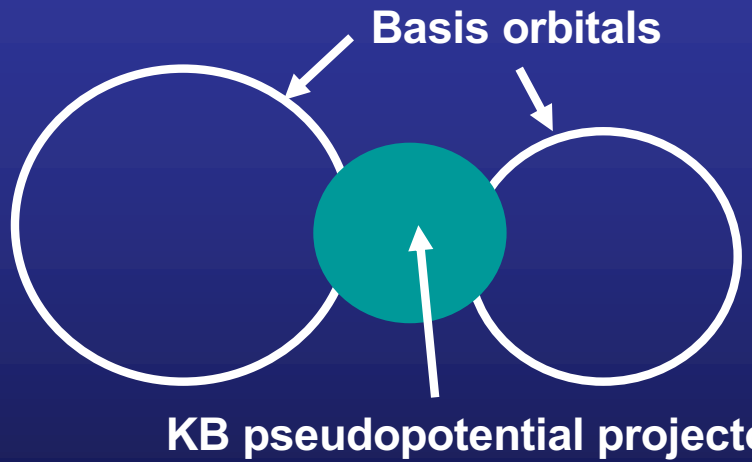
3 with 2,3,4, and 5

4 with 3,4 and 5

5 with 2,3,4, and 5

Sparse ≡ many entrances of the matrix are zero

Non-overlap interactions



$S_{\mu\nu}$ and $H_{\mu\nu}$ are sparse

$\rho_{\mu\nu}$ is not strictly sparse but only a sparse subset is needed

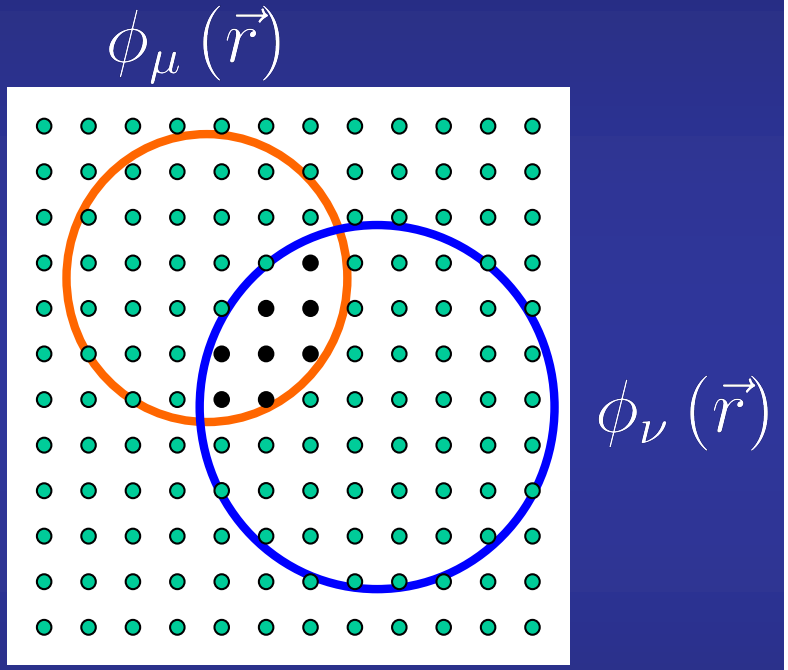
Computation of the density in real space, Hamitonian and Overlap matrix elements: always O(N)

Three dimensional discrete grid to compute Hartree, exchange correlation and neutral atom potentials (related with pseudopotentials)

$$\rho\left(\vec{r}\right) = \sum_{\mu\nu} \rho_{\mu\nu} \phi_{\nu}^{*}\left(\vec{r}\right) \phi_{\mu}\left(\vec{r}\right)$$

$$\rho_{\mu\nu} = \sum_{i} c_{\mu i} n_i c_{i\nu}$$

Density matrix



Find all the atomic orbitals that do not vanish at a given grid point

(in practice, interpolate the radial part from numerical tables)

Once the density is known, we compute the potentials

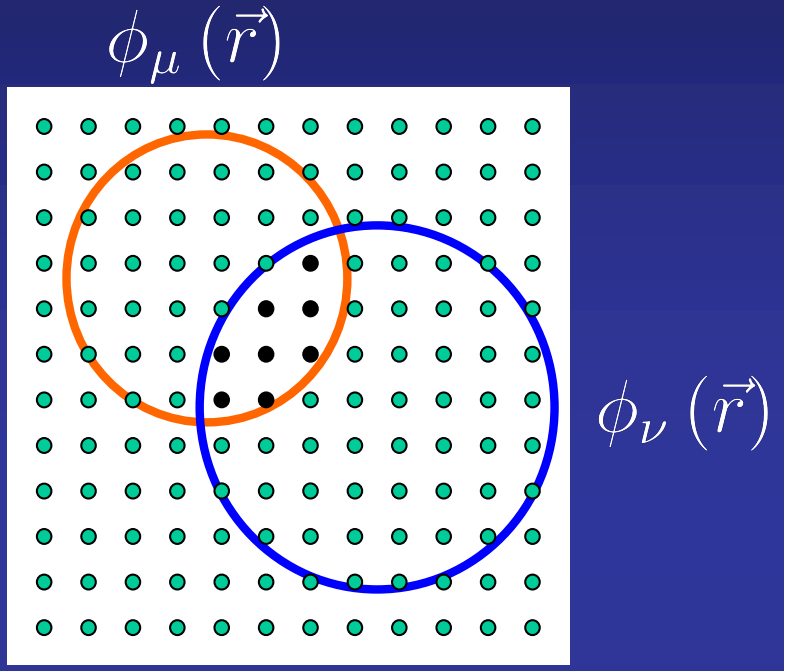
EVERYTHING O(N)

$$\rho\left(\vec{r}\right) \rightarrow V^{xc}\left(\vec{r}\right)$$

$$\delta\rho\left(\vec{r}\right) \xrightarrow{FFT} \delta V^{H}\left(\vec{r}\right)$$

Computation of the density in real space, Hamitonian and Overlap matrix elements: always O(N)

$$\rho\left(\vec{r}\right) = \sum_{\mu\nu} \rho_{\mu\nu} \phi_{\nu}^{*}\left(\vec{r}\right) \phi_{\mu}\left(\vec{r}\right)$$



For the computation of the Hamiltonian matrix elements,

we add together all the grid contributions and perform the integral

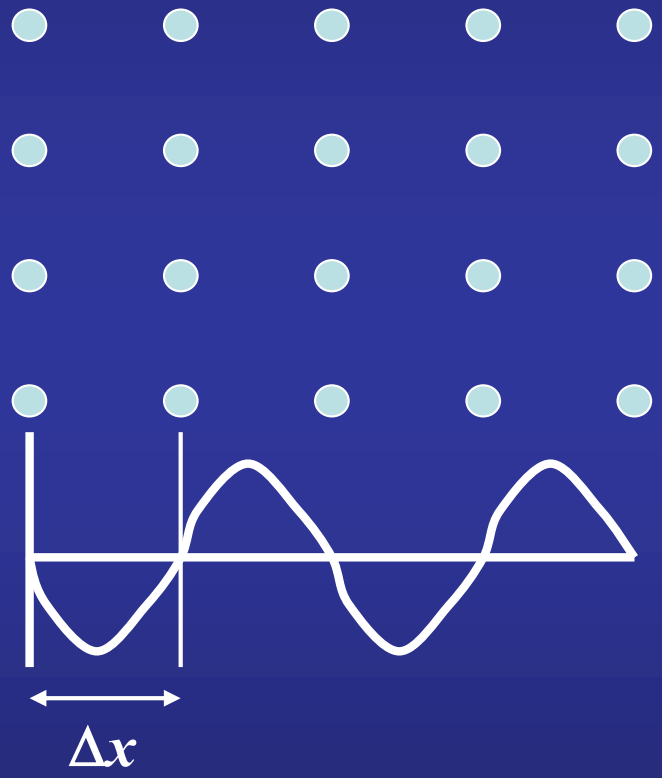
$$V\left(\vec{r}\right) = V^{NA}\left(\vec{r}\right) + \delta V^{H}\left(\vec{r}\right) + V^{xc}\left(\vec{r}\right)$$

$$\int d\vec{r}\ \phi_{\nu}^{*}\left(\vec{r}\right) V\left(\vec{r}\right) \phi_{\nu}\left(\vec{r}\right) \approx \sum_{i} \phi_{\nu}^{*}\left(\vec{r}\right) V\left(\vec{r}\right) \phi_{\nu}\left(\vec{r}\right) \Delta\vec{r}$$

Volume per grid point

Fineness of the grid controlled by a single parameter, the "MeshCutoff"

E_{cut} : maximum kinetic energy of the plane waves that can be represented in the grid without aliasing



$$\Delta x \Longrightarrow k_c = \frac{\pi}{\Delta x} \Longrightarrow E_c = \frac{\hbar^2 k_c^2}{2m_e}$$

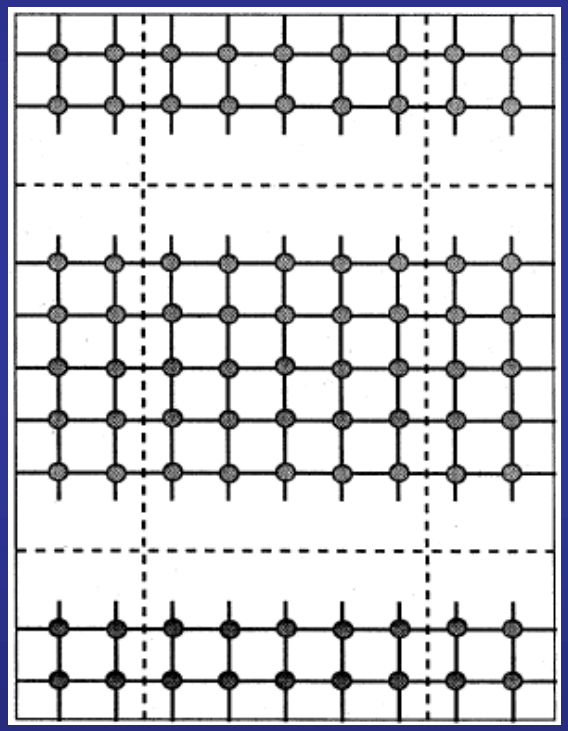
In the grid, we represent the density ⇒ grid cutoff not directly comparable with the plane wave cutoff to represent wave functions

(Strictly speaking, the density requires a value four times larger)

Periodic boundary conditions. For simulations of aperiodic systems: supercell approach

Example:

The supercell approach for surfaces: the slab geometry



The semi-infinite bulk is represented by a slab with two surfaces

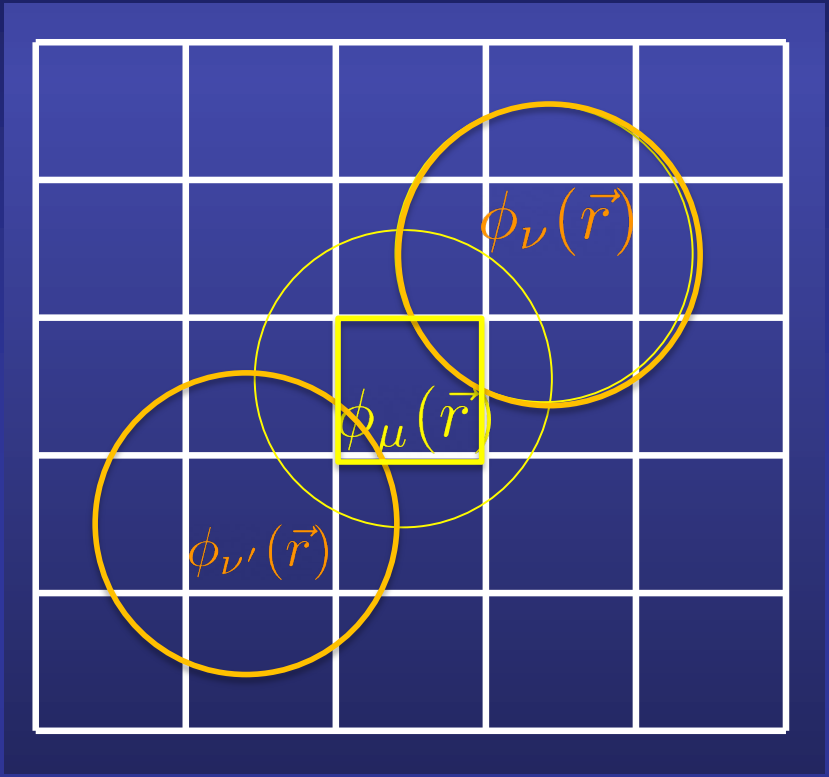
The slab has to be large enough that the two surfaces do not interact with each other

The vacuum between periodic replicas has also to be large enough, specially in charged or polarized slabs

M. C. Payne *et al.*, Rev. Mod. Phys. 64, 1045 (1992)

Usually, semiconductor and insulators require larger supercells than metals

Brillouin zone sampling



$\phi_\mu(\vec{r})$ **Orbital in the unit cell**

$\phi_\nu(\vec{r})$ **and** $\phi_{\nu'}(\vec{r})$
are equivalent orbitals related by a lattice vector

$H_{\mu\nu'}$ **:All non-zero matrix elements in real space between a orbital in the unit cell and the periodic replicas of orbital** ν

$$H_{\mu\nu}(\vec{k}) = \sum_{\nu' \equiv \nu} H_{\mu\nu'} e^{i\vec{k}\cdot\left(\vec{R}_{\nu'} - \vec{R}_{\mu}\right)}$$

Once the hamiltonian and the overlap matrices are build: solve a generalized eigenvalue problem

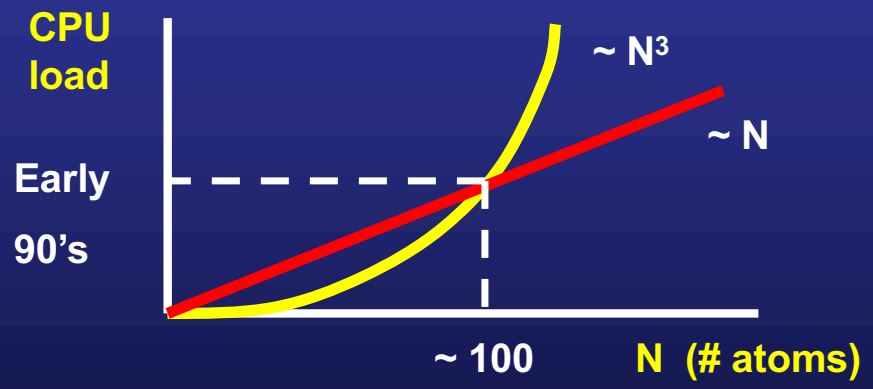
$$\begin{bmatrix} H \end{bmatrix} \begin{bmatrix} C \end{bmatrix} = E_{n\vec{k}} \begin{bmatrix} S \end{bmatrix} \begin{bmatrix} C \end{bmatrix}$$

The solver step takes most of the CPU time

Originally: linear scaling solvers

Minimization of an energy functional based on the localization of solutions

Not valid for metals or "dirty" gap systems



Now: Various solver options with various scalings

Standard diagonalization techniques:

- use pre-packaged libraries (scalapack)
- $O(N^3)$ in time, $O(N^2)$ in memory
- Both eigenvectors and eigenvalues

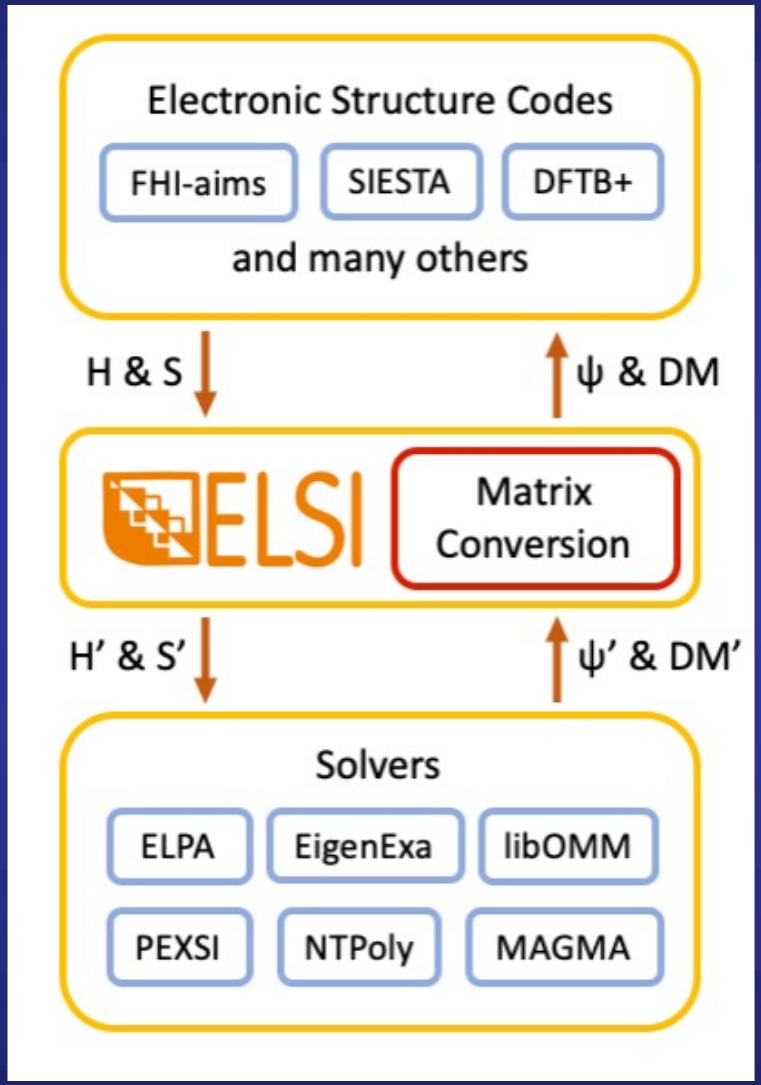
ELPA

CHESS

PEXSI

Many others

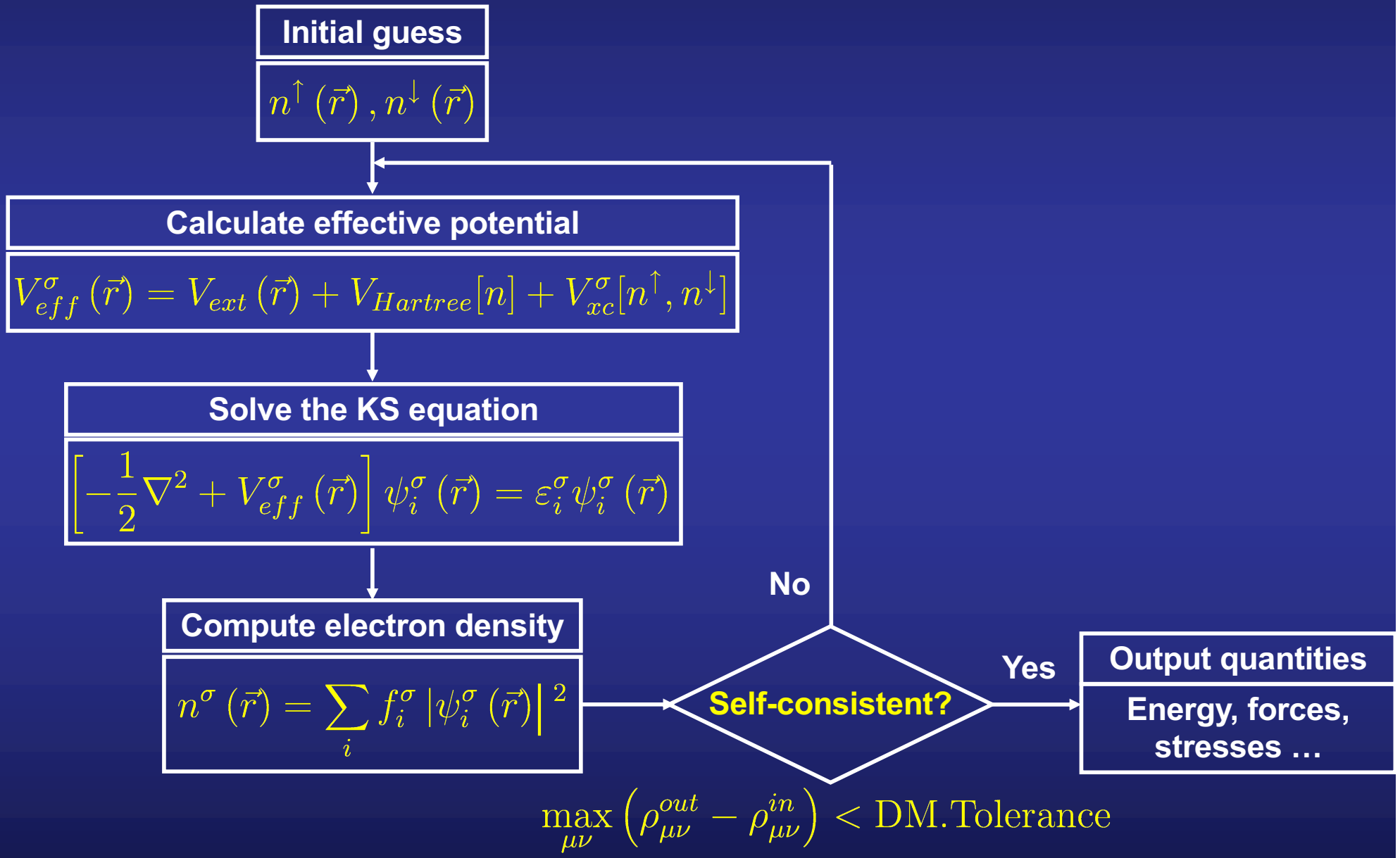
Solver structure for performance and features: Use external libraries



https://elsi-interchange.org

Interface with SIESTA: collaboration with Victor Yu (Duke University)

The Kohn-Sham equations must be solved self-consistently
The potential (input) depends on the density (output)



Atomic forces and stresses obtained by direct diferentiation of the energy expression

$$\vec{F}_I = -\frac{\partial E^{KS}}{\partial \vec{R}_I}$$

$$\sigma_{\alpha\beta} = \frac{\partial E^{KS}}{\partial \epsilon_{\alpha\beta}} \qquad \epsilon_{\alpha\beta} \equiv \text{strain tensor}$$

"One piece of energy ⇒ one piece of force and stress"

Calculated as the analytical derivatives of the energy

Pulay corrections, related with the dependency of the basis set on atomic positions, automatically included

Calculated only in the last self-consistent step

Different ensembles, different Lagrangians, different Conserved magnitudes.

- *NVE (Verlet): Microcanonical.*
- *Integrates Newtons equations of motion, for N particles, in a fixed volume V.*
- *Natural time evolution of the system: E is a constant of motion*

- *NVT (Nose): Canonical*
- *System in thermal contact with a heat bath.*
- *Extended Lagrangian:*
- *N particles + Thermostat, mass Q.*

- *NPE (Parrinello-Rahman) (isobarical)*
- *Extended Lagrangian*
- *Cell vectors are dynamical variables with an associated mass.*

- *NPT (Nose-Parrinello-Rahman)*
- *2 Extended Lagrangians*
- *NVT+NPE.*

SIESTA capabilities

SIESTA: Recent developments and applications

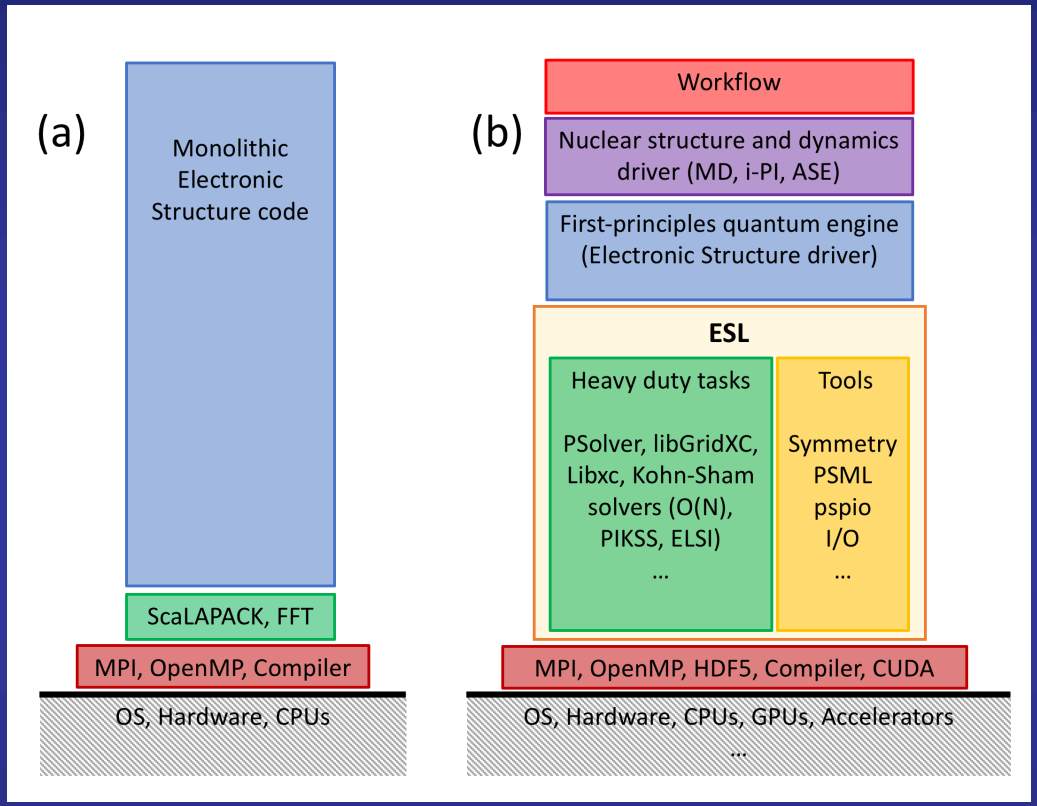
Cite as: J. Chem. Phys. 152, 204108 (2020); doi: 10.1063/5.0005077
Submitted: 18 February 2020 • Accepted: 20 April 2020 •
Published Online: 27 May 2020

View Online Export Citation CrossMark

Alberto García,[1,a] Nick Papior,[2,b] Arsalan Akhtar,[3,c] Emilio Artacho,[4,5,6,7,d] Volker Blum,[8,9,e] Emanuele Bosoni,[1,f] Pedro Brandimarte,[5,g] Mads Brandbyge,[10,h] J. I. Cerdá,[11,i] Fabiano Corsetti,[4,j] Ramón Cuadrado,[3,k] Vladimir Dikan,[1,l] Jaime Ferrer,[12,13,m] Julian Gale,[14,n] Pablo García-Fernández,[15,o] V. M. García-Suárez,[12,13,p] Sandra García,[3,q] Georg Huhs,[16,r] Sergio Illera,[3,s] Richard Korytár,[17,t] Peter Koval,[18,u] Irina Lebedeva,[4,v] Lin Lin,[19,20,w] Pablo López-Tarifa,[21,x] Sara G. Mayo,[22,y] Stephan Mohr,[16,z] Pablo Ordejón,[3,aa] Andrei Postnikov,[23,ab] Yann Pouillon,[15,ac] Miguel Pruneda,[3,ad] Roberto Robles,[21,ae] Daniel Sánchez-Portal,[5,21,af] Jose M. Soler,[22,24,ag] Rafi Ullah,[4,25,ah] Victor Wen-zhe Yu,[8,ai] and Javier Junquera[15,aj]

- PSML pseudopotentials
- LDA+U
- TRANSIESTA (ballistic charge transport at the nanoscale)
- TDDFT in real time
- Full Spin-Orbit coupling, non-collinear magnetism
- Various eigensolvers with different scalings
- Density functional perturbation theory
- TDDFT in frequency space and GW post-processing
- Multiscale (second-principles) and QM/MM
- Several analysis and post-processing tools
- Wannierization
- Hybrid functionals
- Dynamical Mean Field Theory (coming soon)
- Superconductivity and electron phonon coupling

SIESTA modularity:
A change in the paradigm



The CECAM electronic structure library

M. J. T. Oliveira *et al*.
J. Chem. Phys. 153, 024117 (2020)

Norm conserving pseudopotential library in Pseudopotential Markup Language (PSML) format

Remove interoperability problems
(the same pseudopotential operator can be used by different codes)

Computer Physics Communications 227 (2018) 51–71

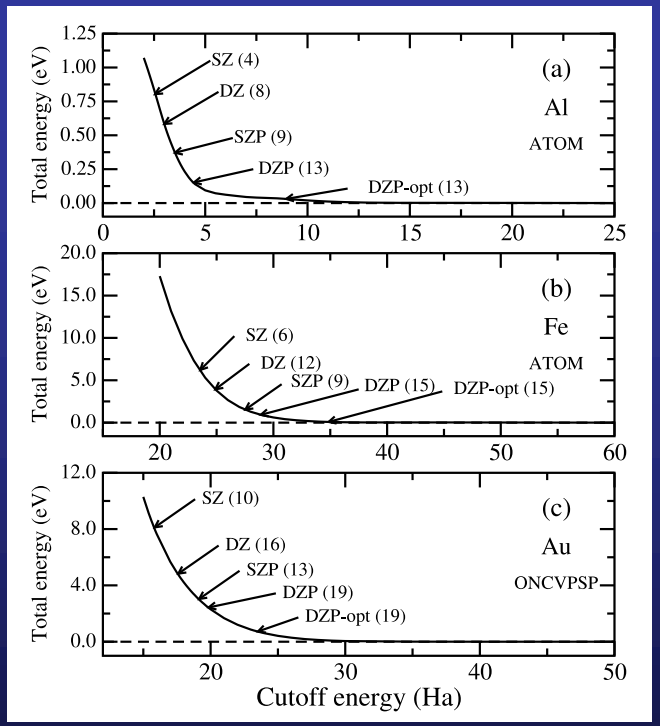
Contents lists available at ScienceDirect

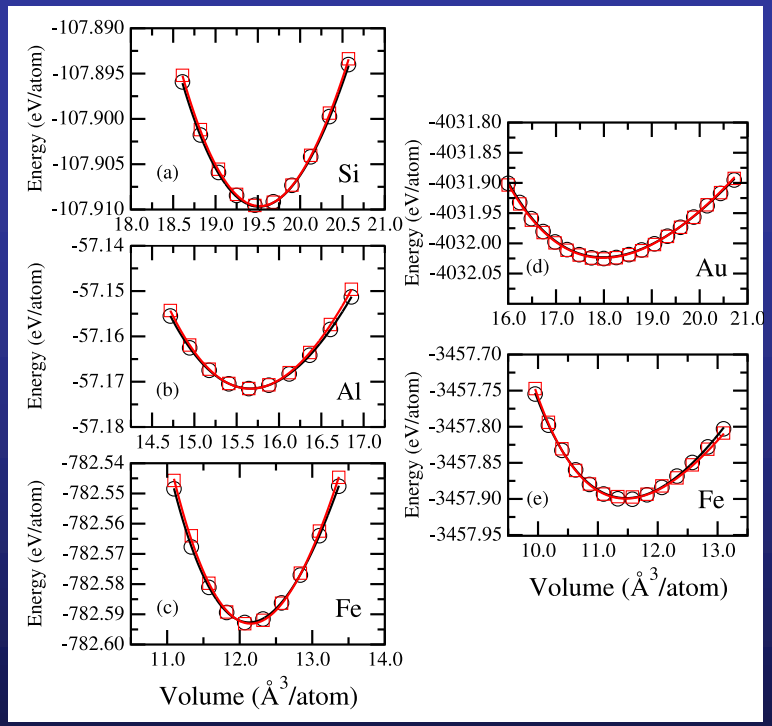
Computer Physics Communications

journal homepage: www.elsevier.com/locate/cpc

The PSML format and library for norm-conserving pseudopotential data curation and interoperability☆

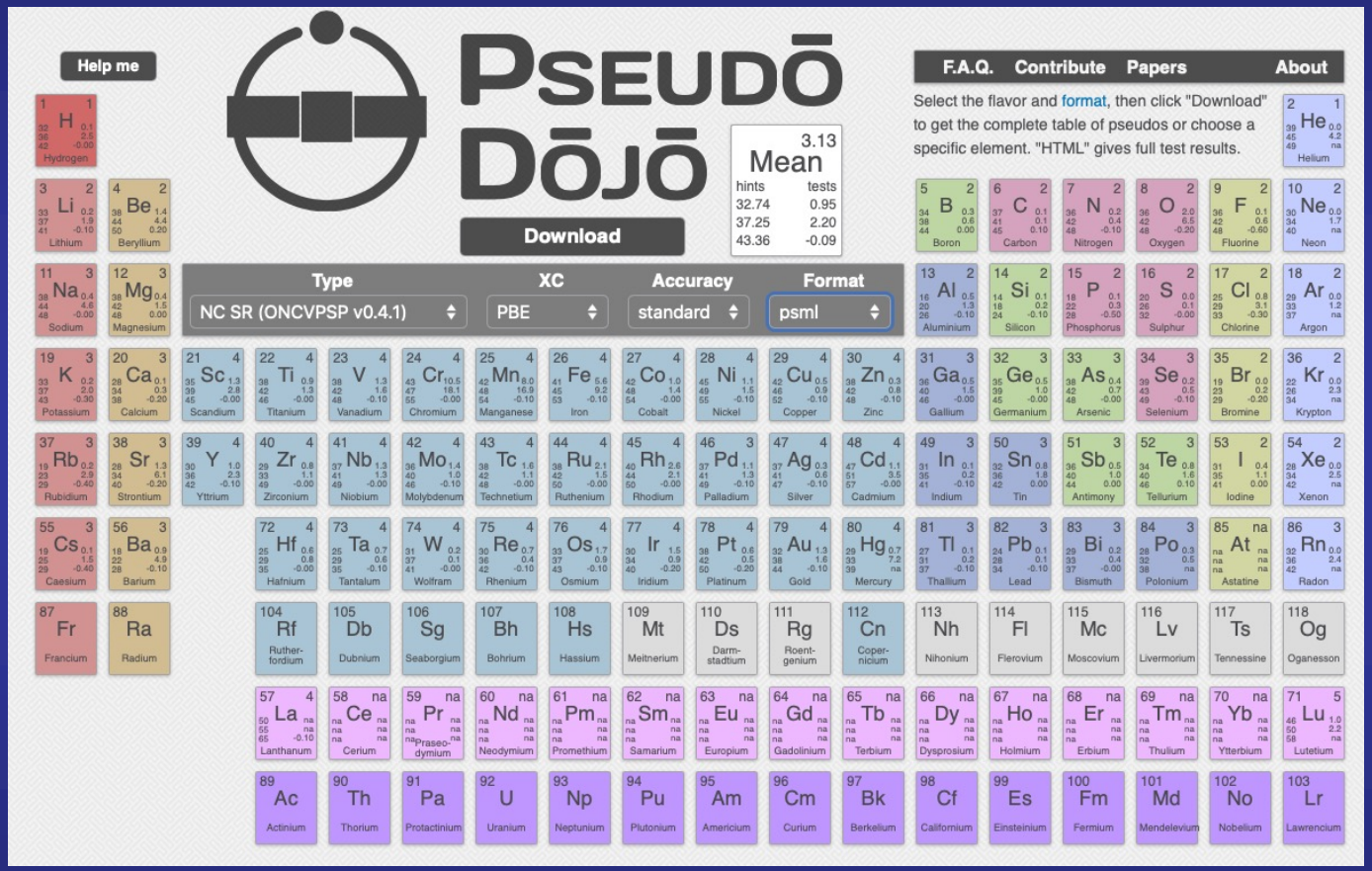
Alberto García [a,*], Matthieu J. Verstraete [b], Yann Pouillon [c], Javier Junquera [c]





Norm conserving pseudopotential library in Pseudopotential Markup Language (PSML) format

http://www.pseudo-dojo.org

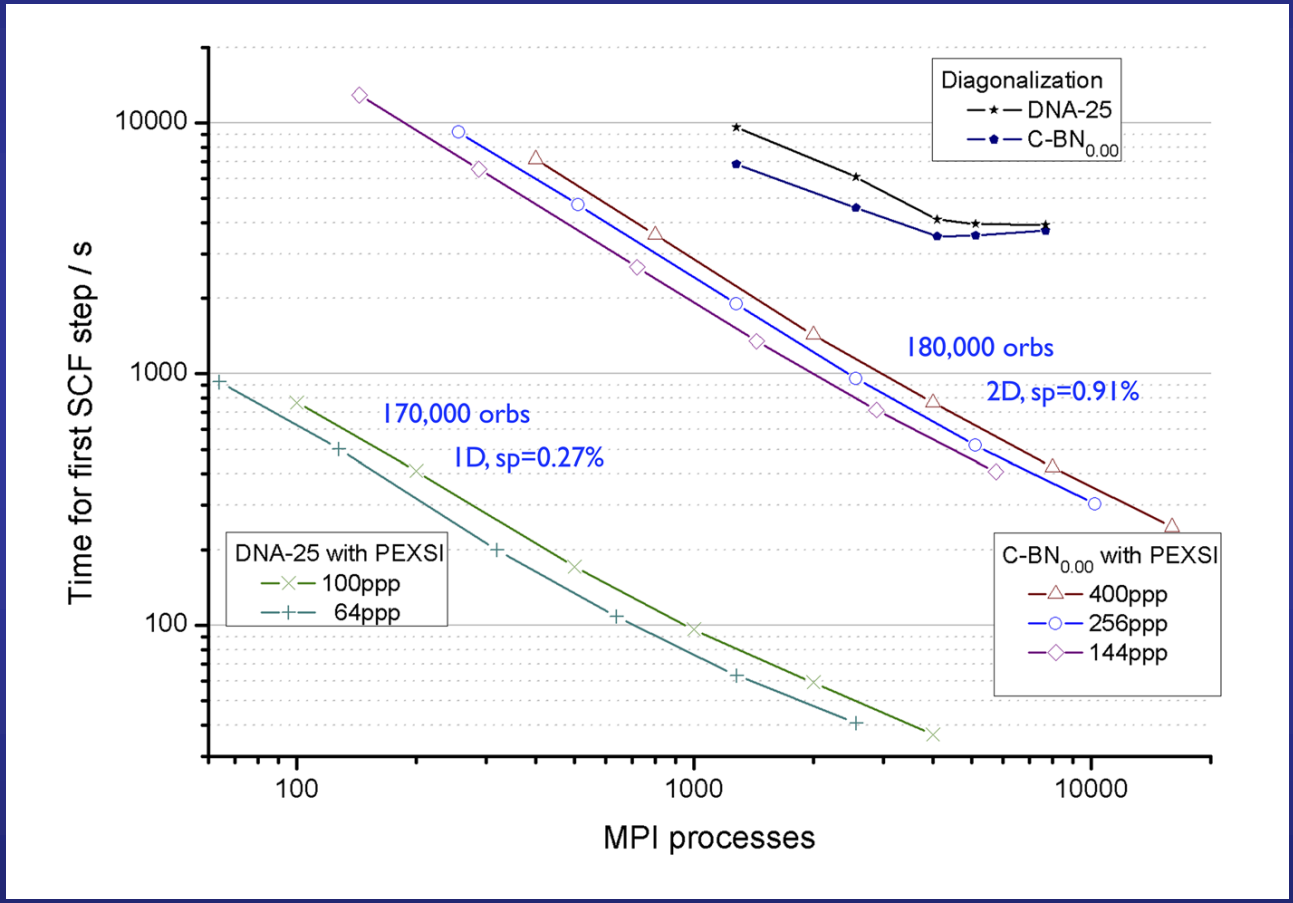


Periodic table of curated pseudopotentials

The testing of the pseudopotential is a responsability of the user,
But the availability of reliable and accurate norm-conserving pseudopotential lowers the barrier

SIESTA parallelization

Both distributed [message-passing (MPI)] and shared-memory (OpenMP) parallelization options implemented



SIESTA base line efficiency can be scaled up to every larger systems

SIESTA is a very user friendly code

```
SystemName          Water molecule
SystemLabel         h2o
NumberOfAtoms       3
NumberOfSpecies     2

%block ChemicalSpeciesLabel
 1  8  O      # Species index, atomic number, species label
 2  1  H
%endblock ChemicalSpeciesLabel

AtomicCoordinatesFormat  Ang
%block AtomicCoordinatesAndAtomicSpecies
 0.000  0.000  0.000  1
 0.757  0.586  0.000  2
-0.757  0.586  0.000  2
%endblock AtomicCoordinatesAndAtomicSpecies
```

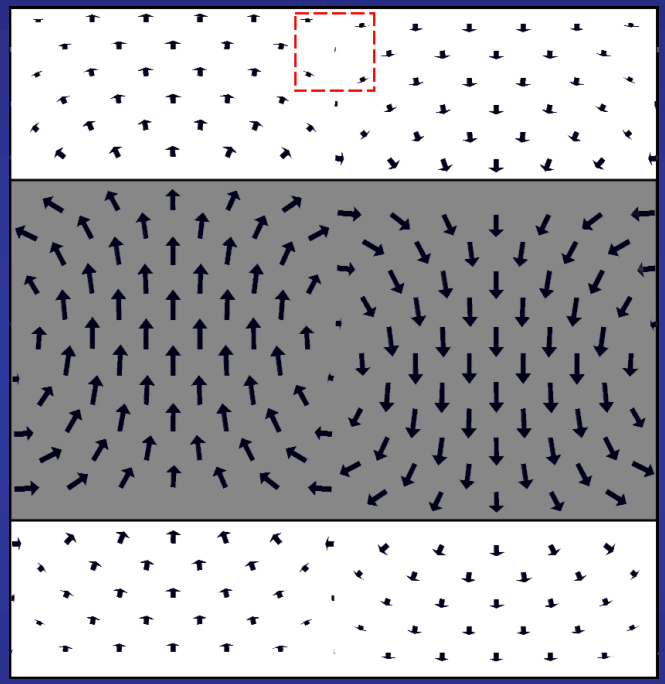
SIESTA applications in extremely varied contexts

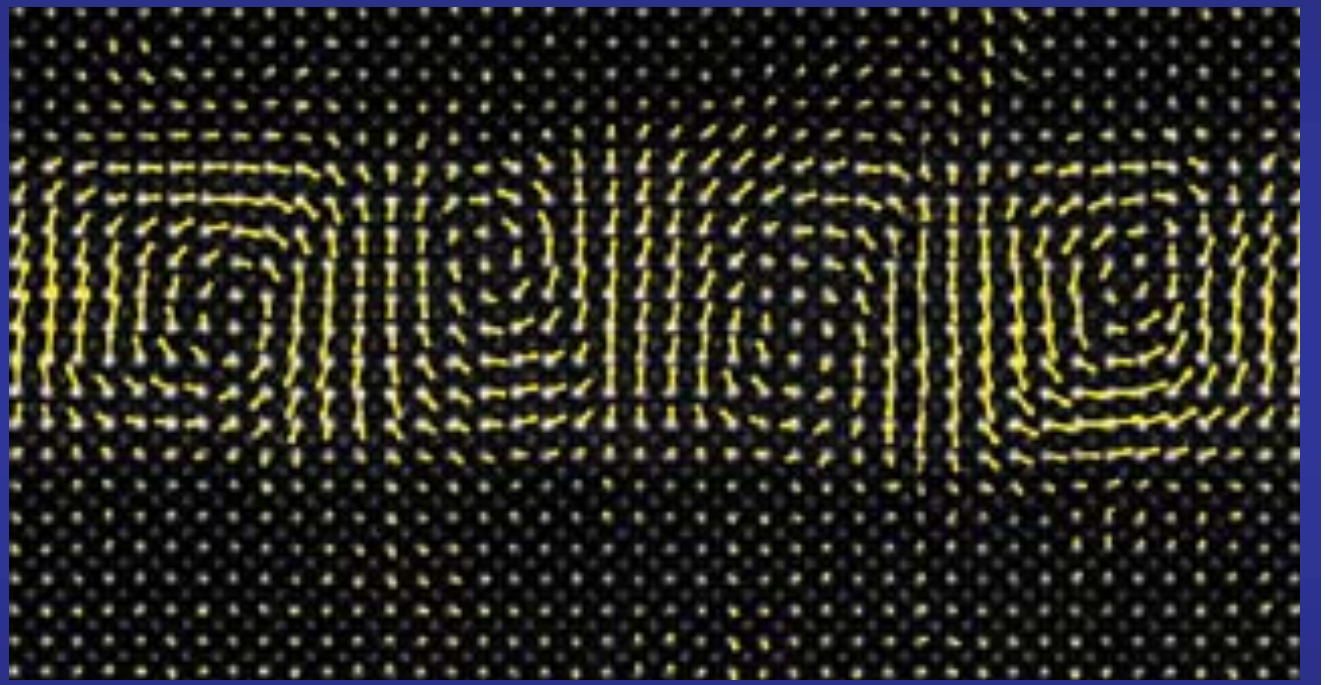
- Physics
- Chemistry
- Material Science
- Biology
- Geology
- Nanoscience
- Engineering
- Pharmacy

Difficult to follow (over 2000 citations per year)

☐	The SIESTA method for ab initio order-N materials simulation JM Soler, E Artacho, JD Gale, A García, J Junquera, P Ordejón, ... Journal of Physics: Condensed Matter 14 (11), 2745	11746	2002
	Density-functional method for nonequilibrium electron transport M Brandbyge, JL Mozos, P Ordejón, J Taylor, K Stokbro Physical Review B 65 (16), 165401	4806	2002
	Self-consistent order-N density-functional calculations for very large systems P Ordejón, E Artacho, JM Soler Physical Review B 53 (16), R10441	2788	1996
	Density-functional method for very large systems with LCAO basis sets D Sánchez-Portal, P Ordejon, E Artacho, JM Soler International journal of quantum chemistry 65 (5), 453-461	1851	1997
	Linear-scaling ab-initio calculations for large and complex systems E Artacho, D Sánchez-Portal, P Ordejón, A Garcia, JM Soler physica status solidi (b) 215 (1), 809-817	1181	1999
☐	Numerical atomic orbitals for linear-scaling calculations J Junquera, Ó Paz, D Sánchez-Portal, E Artacho Physical Review B 64 (23), 235111	1133	2001
☐	The SIESTA method; developments and applicability E Artacho, E Anglada, O Diéguez, JD Gale, A García, J Junquera, ... Journal of Physics: Condensed Matter 20 (6), 064208	599	2008
☐	Systematic generation of finite-range atomic basis sets for linear-scaling calculations E Anglada, JM Soler, J Junquera, E Artacho Physical Review B 66 (20), 205101	318	2002

Material science: topologically non-trivial phases in $PbTiO_3/SrTiO_3$ superlattices



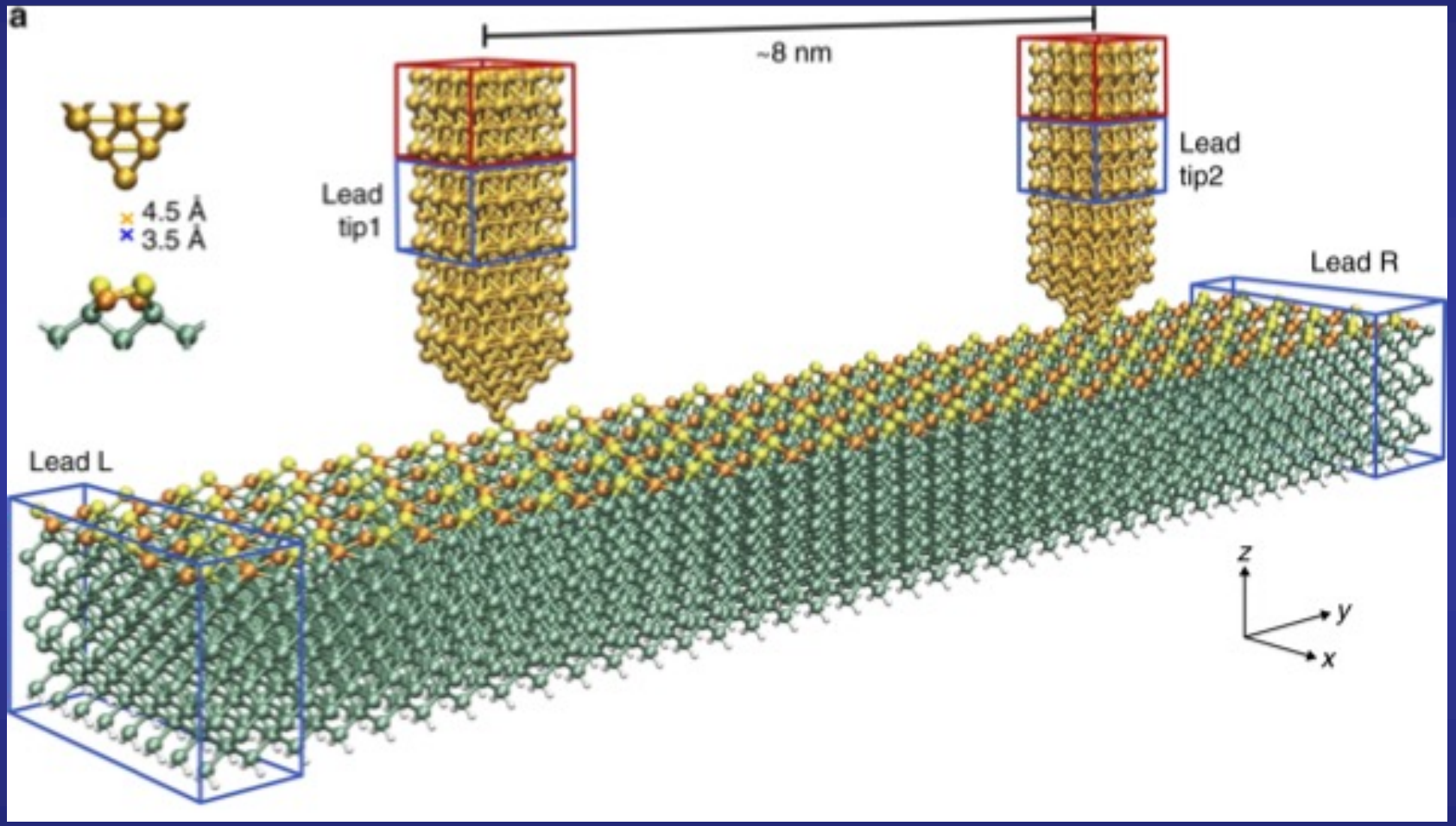


P. Aguado-Puente and J. Junquera
Phys. Rev. B 85, 184105 (2012)

A. Yadav *et al*. Nature 530, 198 (2016)

Predictive power of the simulations

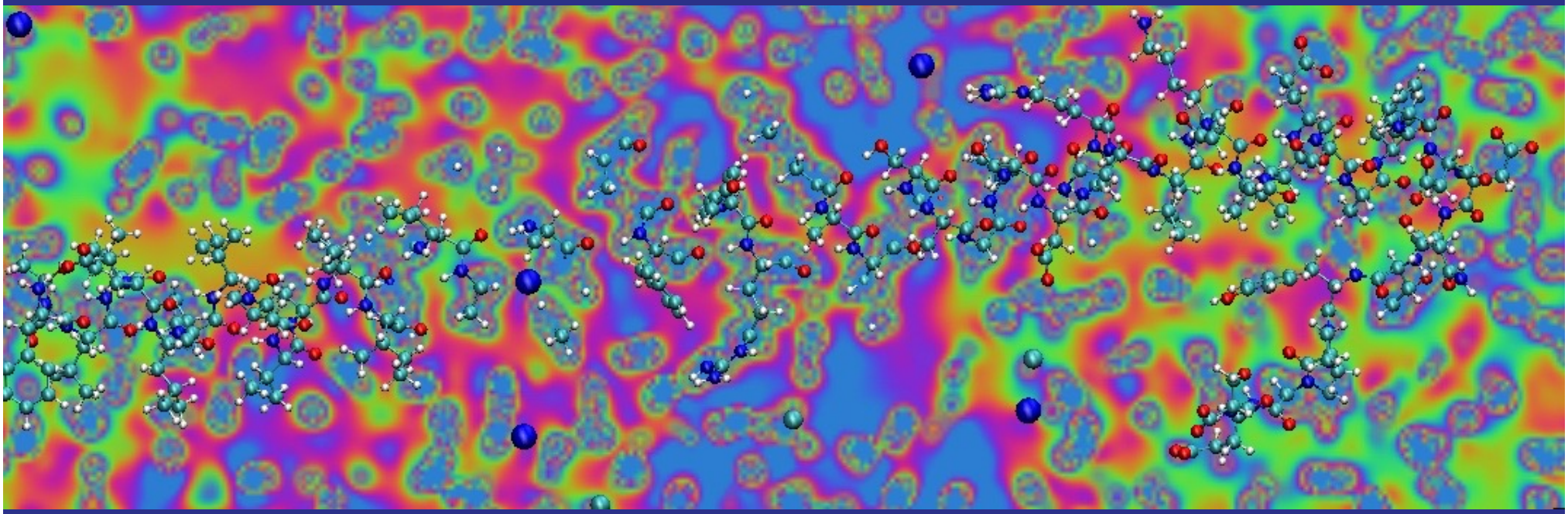
Nanotechnology:
quantum mechanical transport simulations



Four terminal transport on Ge Surface and gold tips (4924 atoms)

M. Kolmer *et al*. Nat. Commun. 10, 1573 (2019)

Biology:
Electrostatics around pilin protein in wet conditions

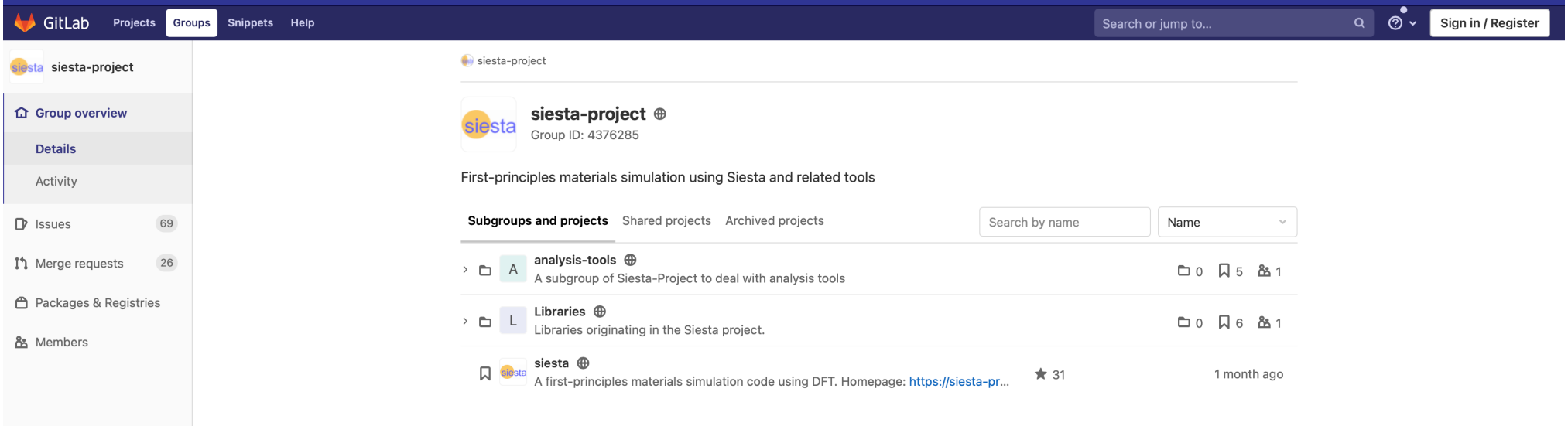


A natural nano electric wire
(4580 atoms)

G. T. Feliciano *et al*. J. Phys. Chem. A 116, 8023 (2012)

Distribution: fully open source (GPL) since 2016

http://gitlab.com/siesta-project



Siesta support

https://departments.icmab.es/leem/siesta/

siesta

Home About Code Documentation Pseudopotentials The Team News Support

SIESTA is both a method and its computer program implementation, to perform efficient electronic structure calculations and ab initio molecular dynamics simulations of molecules and solids. SIESTA's efficiency stems from the use of a basis set of strictly-localized atomic orbitals. A very important feature of the code is that its accuracy and cost can be tuned in a wide range, from quick exploratory calculations to highly accurate simulations matching the quality of other approaches, such as plane-wave methods.

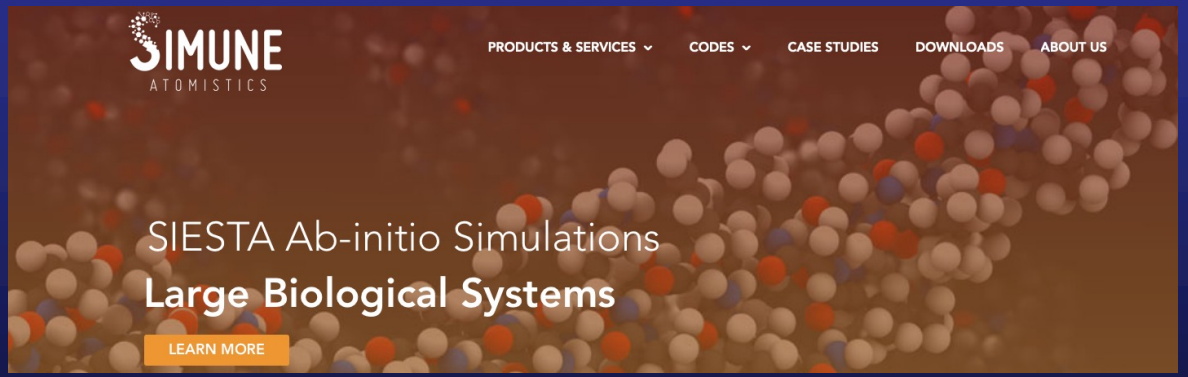
The possibility of treating large systems with some first-principles electronic-structure methods has opened up new opportunities in many disciplines. The SIESTA program is open source and has become quite popular, being increasingly used by researchers in geosciences, biology, and engineering (apart from those in its natural habitat of materials physics and chemistry). Currently there are several thousand users all over the world, and the paper describing the method (J. Phys. Cond. Matt. 14, 2745 (2002)) has received more than 8000 citations so far.

For an overview of recent developments, and sample applications of SIESTA, see J. Chem. Phys. 152, 204108 (2020).

SIESTA is one of the flagship codes of the MaX Center of Excellence.

Devoted Youtube channel will be launched next week

SIMUNE: a Company that offers profesional support
https://www.simuneatomistics.com



Siesta support

https://personales.unican.es/junqueraj/siesta-tutorial.html

Self-explained SIESTA tutorial

Updated April 2016

Here we present a collection of theoretical lectures and self-explained SIESTA exercises compiled during the years on many schools and tutorials. I would like to acknowledge the collaboration of the SIESTA team during the preparation, testing, and cleaning of the exercises:

Emilio Artacho	CIC Nanogune	San Sebastián (Spain)
Julian Gale	Curtin University of Technology	Perth (Australia)
Alberto García	Institut de Ciència de Materials de Barcelona (ICMAB-CSIC)	Barcelona (Spain)
José Soler	Universidad Autónoma de Madrid	Madrid (Spain)
Pablo Ordejón	Institut Català de Nanociència i Nanotecnologia (ICN2)	Barcelona (Spain)
Daniel Sánchez-Portal	Unidad de Física de Materiales, Centro Mixto CSIC-UPV/EHU	San Sebastián (Spain)

The present exercises are just simple examples to show how-to run different capabilities of SIESTA. Before running a calculation for production, test the pseudopotentials, basis sets, and perform the convergence tests (mesh cutoff, k-grid sampling, etc). Some of these parameters have been chosen for you to speed up the calculations, and might not be converged.

Please, report any error or mistake that you could detect on these exercises to: javier.junquera -- unican.es

Openings

Some openings for PhD fellowships funded by the University of Cantabria are available. If you are a Master student who wish to work in the field, please take a look at this call

The Theory of Condensed Matter Group at the University of Cantabria is willing to support personal applications to PhD and post-doctoral fellowships offered by funding agencies from Spain and Europe (through the Marie Skłodowska-Curie actions)

Some theoretical lectures

Set of self-explained SIESTA exercises

Back to Javier Junquera's home page