

Large scale DFT calculations with SIESTA

Pablo Ordejón,^{*1} Daniel Sanchez-Portal,^{*2} Alberto Garcia,^{*3} Emilio Artacho,^{*4} Javier Junquera,^{*4} and Jose M. Soler^{*4}

^{*1} *Institut de Ciència de Materials de Barcelona, CSIC, Spain*

^{*2} *Department of Physics and Materials Research Laboratory, University of Illinois, USA*

^{*3} *Departamento de Física Aplicada II, Universidad del País Vasco, Spain*

^{*4} *Departamento de Física de Materia Condensada, C-III, Universidad Autónoma de Madrid, Spain*

In the context of linear scaling methods for electronic structure and molecular dynamics calculations, SIESTA was developed as a fully first-principles method able to deal with systems with an unprecedented number of atoms, with a modest computational workload. The method has allowed us to study a large variety of problems involving nanoscale materials, such as nanoclusters, nanotubes, biological molecules, adsorbates in surfaces, etc. Here we present a brief review of such applications.

Introduction

The contribution of first-principles calculations to several fields in physics, chemistry, materials sciences, and recently geology and biology is more important than ever. The main reasons are, on one hand, the steady increase in computer power, and on the other, the continuous progress in methodology (both in efficiency and in accuracy of algorithms and approximations). As larger and more complex systems are falling within the range of applicability of these methods, some barriers (formerly believed to be fundamental) are being reached. One of these is the scaling of the computational effort with system size. In the most favorable cases like with Density Functional Theory (DFT),^{1,2)} quantum mechanical formulations of the electronic structure of atomic systems scale as the cube of the number of atoms (or electrons) of the system.³⁾ This makes it very difficult to reach system sizes larger than a few hundreds of atoms, and is therefore a huge barrier for the study of problems in nanoscale materials. The way out of this trap was found several years ago, when a number of ideas suggested the possibility of developing approximate although accurate schemes to reduce the computational cost to linear scaling. These so called $O(N)$ methods⁴⁾ have matured since those first proposals, and now constitute a viable route for studying systems with unprecedented size.

The key for achieving linear scaling is the explicit use of locality, meaning by it the insensitivity of the properties of a region of the system to perturbations sufficiently far away from it.⁵⁾ A local language will thus be needed for the two different problems one has to deal with in a DFT-like method: building the self-consistent Hamiltonian, and solving it. Most of the initial effort was dedicated to the latter^{4,6)} using empirical or semi-empirical Hamiltonians. The SIESTA project⁷⁻¹⁰⁾ started in 1995 to address the former. Atomic-orbital basis sets were chosen as the local language, allowing for arbitrary basis sizes, what resulted in a general-purpose, flexible linear-scaling DFT program.⁹⁻¹¹⁾ A parallel effort has been the search for orbital bases that would meet the standards of precision of conventional first-principles calculations, but keeping as small a range as possible for maximum efficiency.¹⁰⁾

In this paper, we give a brief overview of several applications

of SIESTA to problems in different nanoscale systems. We first summarize the main features and approximations used in SIESTA, and then give a very brief description of several applications in carbon nanostructures (fullerenes and nanotubes), metallic nanostructures (transition and noble metals, and nanowires), biomolecules, and surfaces and disordered systems.

The SIESTA method

SIESTA is based on DFT, and can use both local-density (LDA)³⁾ and generalized-gradients (GGA) functionals,¹²⁾ including spin polarization, collinear and non-collinear.¹³⁾ The core electrons are replaced by norm-conserving pseudopotentials¹⁴⁾ factorized in the Kleinman-Bylander form,¹⁵⁾ including scalar-relativistic effects, and non-linear partial-core corrections.¹⁶⁾ The one-particle problem is then solved using linear combination of atomic orbitals (LCAO). There are no constraints either on the radial shape of these orbitals (which are treated numerically), or on the size of the basis, allowing for the full quantum-chemistry (QC) know-how¹⁷⁾ (multiple- ζ , polarization, off-site, contracted, and diffuse orbitals). Forces on the atoms and the stress tensor are obtained from the Hellmann-Feynman theorem (including Pulay corrections), and can be used for structure relaxations or molecular dynamics simulations of different types.

The DFT equations are solved using the self-consistent field (SCF) method. For a Hamiltonian, the one-particle Schrödinger equation is solved yielding the energy and density matrix for the ground state. This is performed either by diagonalization (cube-scaling, appropriate for systems under a hundred atoms or for metals) or with a linear-scaling algorithm. These have been extensively reviewed elsewhere.⁴⁾ SIESTA implements two $O(N)$ algorithms^{6,18)} based on localized Wannier-like wavefunctions.

Once the density matrix has been obtained, the SCF procedure continues with the calculation of a new Hamiltonian matrix. The matrix elements of the different terms of the Kohn-Sham Hamiltonian are calculated in one of two different ways.⁹⁾ The terms that involve integrals over two atoms only (kinetic energy, overlap, and other terms related with

the pseudopotential) are computed *a priori* as a function of the distance between the centers, and stored in tables to be interpolated later with very little use of time and memory. The other terms are calculated with the help of a uniform grid of points in real space. The smoothness of the intergrands determines how fine a grid is needed, and, of course, the finer the grid, the more expensive the calculation. We remark that the use of pseudopotentials, which eliminates the rapidly varying core charge, is essential to provide functions smooth enough to make the grid integration feasible. This fineness is measured by the energy of the shortest wave-length plane-wave that can be described with the grid, in analogy with plane-wave calculations.

The calculation of the Hamiltonian matrix elements sketched above has an $O(N)$ scaling *provided that* the range of overlap between the basis orbitals is finite. To achieve that, we use basis orbitals which strictly vanish beyond a cutoff radius¹⁰⁾ (instead of the usual approach of using decaying orbitals and neglecting matrix elements by whatever criterion). The main advantage is consistency: given a basis, the eigenvalue problem is solved for the *full* Hamiltonian. Thus, the procedure is numerically very stable even for short ranges, in contrast with the usual approach. In this and previous works, the radial parts of the finite-range orbitals were determined in the spirit of the method Sankey and Niklewski,¹⁹⁾ who proposed a scheme for minimal (single- ζ) bases that we have generalized to arbitrarily complete sets.¹⁰⁾ The single- ζ orbitals are obtained by solving the DFT atomic problem (including the pseudopotential) with the boundary condition for the orbitals of being zero beyond the cutoff radius, while remaining continuous.²⁰⁾ For the efficient generation of larger, more complete basis sets we have used the ideas developed within the QC community over the years, incorporating them into new schemes adapted to numerical, finite-range bases for linear scaling. Numerical multiple- ζ bases are constructed in the split-valence philosophy.^{9,10)} Our approach also allows polarization orbitals¹⁰⁾ which are obtained by numerically solving the problem of the isolated atom in the presence of a polarizing electric field.

Applications

Carbon nanostructures: Fullerenes and nanotubes. A preliminary version of SIESTA was first applied to study the shape of large hollow carbon fullerenes⁷⁾ up to C_{540} , the results contributing to establish that they do not tend to a spherical-shape limit but tend to facet around the twelve corners given by the pentagons. SIESTA has been also applied to carbon nanotubes. In a first study, structural, elastic and vibrational properties were characterized.²¹⁾ A second work was dedicated to their deposition on gold surfaces, and the STM images that they originate,²²⁾ specially addressing experiments on finite-length tubes. A third study has been dedicated to the opening of single-wall nanotubes with oxygen, and the stability of the open, oxidized tubes for intercalation studies.²³⁾

Metallic nanostructures: Clusters and wires. Gold nanoclusters of small sizes (up to Au_{75}) were found²⁴⁾ to be amorphous, or nearly so, even for sizes for which very favorable geometric structures had been proposed before. In a fur-

ther study the origin of this striking situation is explained in terms of local stresses.²⁵⁾ Chains of gold atoms have been studied addressing the experiments²⁶⁾ which show them displaying remarkably long interatomic spacings (4–5 Å). A first study²⁷⁾ arrives at the conclusion that a linear gold chain would break at interatomic spacings much smaller than the observed ones. A possible explanation of the discrepancy is reported elsewhere,²⁸⁾ in terms of a spinning zigzag structure for the monoatomic wires. The magnetic properties of low dimensional Fe systems were also studied.²⁹⁾ Both free clusters and supported on Ag(001) surfaces were considered. The effect of coordination and substrate were studied. Finally, the structure and energetics of binary Ni-Al clusters was studied, addressing the segregation of species towards the cluster surface.³⁰⁾

Biomolecules. Feasibility tests on DNA were performed in the early stages of the project, by relaxing a dry B-form poly(dC)-poly(dG) structure with a minimal basis.^{8,9)} In preparation of realistic calculations, a thorough study³¹⁾ of 30 nucleic acid pairs has been performed addressing the precision of the approximations and the DZP bases, and the accuracy of the GGA functional,¹²⁾ obtaining good results even for the hydrogen bridges. Based on that, a first study of dry A-DNA has been performed, with a full relaxation of the structure, and an analysis of the electronic characteristics.³²⁾

Surfaces and disordered systems. A molecular dynamics simulation was performed³³⁾ on the clean surface of liquid silicon close to the melting temperature, in which surface layering was found, *i.e.*, density oscillations of roughly atomic amplitude, like what was recently found to happen in the surface of other liquid metals.³⁴⁾ Unlike them, though, the origin for silicon was found to be orientational, reminiscent of directed octahedral bonding. Adsorption studies have also been performed on solid silicon surfaces, Ba on Si(100)³⁵⁾ and C_{60} on Si(111).³⁶⁾ Both works study adsorption geometries and energetics. For Ba, interactions among adsorbed atoms and diffusion features are studied. For C_{60} , STM images have been simulated and compared to experiments. Metastable phases of FeSi on Si(111) were studied,³⁷⁾ to explain recent photoemission data. Finally, the structure and electronic localization of defects in amorphous Si and C were analyzed,³⁸⁾ supporting the view of the dangling bond as the source of ESR signal in a-Si.

References

- 1) P. Hohenberg and W. Kohn: Phys. Rev. **135**, 864 (1964).
- 2) W. Kohn and L. J. Sham: Phys. Rev. **140**, 1133 (1965).
- 3) M. Payne et al.: Rev. Mod. Phys. **64**, 1045 (1992).
- 4) For review see: P. Ordejón: Comp. Mat. Sci. **12**, 157 (1998); S. Goedecker: Rev. Mod. Phys. **71**, 1085 (1999).
- 5) W. Kohn: Phys. Rev. Lett. **76**, 3168 (1996).
- 6) P. Ordejón, D. A. Drabold, R. M. Martin, and M. P. Grumbach: Phys. Rev. B **51**, 1456 (1995) and references therein.
- 7) P. Ordejón, E. Artacho, and J. M. Soler: Phys. Rev. B **53**, R10441 (1996).
- 8) P. Ordejón, E. Artacho, and J. M. Soler: Mat. Res. Soc. Symp. Proc. **408**, 85 (1996).
- 9) D. Sánchez-Portal, P. Ordejón, E. Artacho, and J. M. Soler: Int. J. Quant. Chem. **65**, 453 (1997).
- 10) E. Artacho, D. Sánchez-Portal, P. Ordejón, A. García, and J. M. Soler: Phys. Stat. Sol. (b) **215**, 809 (1999).
- 11) D. Sánchez-Portal, P. Ordejón, E. Artacho, and J. M. Soler: to be published.

- 12) J. P. Perdew, K. Burke, and, M. Ernzerhof: Phys. Rev. Lett. **77**, 3865 (1996).
- 13) T. Oda, A. Pasquarello, and R. Car: Phys. Rev. Lett. **80**, 3622 (1998).
- 14) N. Troullier and J. L. Martins: Phys. Rev. B **43**, 1993 (1991).
- 15) L. Kleinman and D. M. Bylander: Phys. Rev. Lett. **48**, 1425 (1982).
- 16) S. G. Louie, S. Froyen, and M. L. Cohen: Phys. Rev. B **26**, 1738 (1982).
- 17) S. Huzinaga et al.: *Gaussian Basis Sets for Molecular Calculations*, (Elsevier Science, Amsterdam, 1984); R. Poirier, R. Kari, and R. Csizmadia: *Handbook of Gaussian Basis Sets*, (Elsevier Science, Amsterdam, 1985) and references therein.
- 18) J. Kim, F. Mauri, and G. Galli: Phys. Rev. B. **52**, 1640 (1995).
- 19) O. F. Sankey and D. J. Niklewski: Phys. Rev. B **40**, 3979 (1989).
- 20) For attaining a desired degree of convergence, the cutoff radii of the different orbitals depend on the species, and also on the specific orbitals themselves. We have proposed a systematic way of finding the different cutoff radii as a function of a single parameter that defines the finite-range approximation: the energy shift.¹⁰⁾ All pseudoatomic orbitals shift upwards in energy when confined to a sphere; defining this energy shift gives all the required radii, guaranteeing a balanced basis.
- 21) P. Ordejón, D. Sánchez-Portal, E. Artacho, and J. M. Soler: Fuller. Sci. Technol., (in press); D. Sánchez-Portal, E. Artacho, J. M. Soler, A. Rubio, and P. Ordejón: Phys. Rev. B **59**, 12678 (1999).
- 22) A. Rubio, D. Sánchez-Portal, E. Artacho, P. Ordejón, and J. M. Soler: Phys. Rev. Lett. **82**, 3520 (1999).
- 23) M. S. C. Mazzoni, H. Chacham, P. Ordejón, D. Sánchez-Portal, J. M. Soler, and E. Artacho: Phys. Rev. B **60**, 2208 (1999).
- 24) I. L. Garzón, K. Michaelian, M. R. Beltrán, A. Posada-Amarillas, P. Ordejón, E. Artacho, D. Sánchez-Portal, and J. M. Soler: Phys. Rev. Lett. **81**, 1600 (1998).
- 25) J. M. Soler, M. R. Beltrán, K. Michaelian, I. L. Garzón, P. Ordejón, D. Sánchez-Portal, and E. Artacho: Phys. Rev. B **61**, 5771 (2000).
- 26) H. Ohnishi, Y. Kondo, and K. Takayanagi: Nature **395**, 780 (1998); A. I. Yanson, G. Rubio Bollinger, H. E. van den Brom, N. Agraït, and J. M. van Ruitenbeek: Nature **395**, 783 (1998).
- 27) J. A. Torres, E. Tosatti, A. dal Corso, F. Ercolessi, J. J. Kohanoff, F. Di Tolla, and J. M. Soler: Surf. Sci. Lett. **426**, L441 (1999).
- 28) D. Sánchez-Portal, J. Junquera, P. Ordejón, A. García, E. Artacho, and J. M. Soler: Phys. Rev. Lett. **83**, 3884 (1999).
- 29) Izquierdo, A. Vega, L. C. Balbás, D. Sánchez-Portal, E. Artacho, J. M. Soler, and P. Ordejón: Phys. Rev. B **61**, 13639 (2000).
- 30) M. Calleja, C. Rey, M. M. G. Alemany, L. J. Gallego, P. Ordejón, D. Sánchez-Portal, E. Artacho, and J. M. Soler: Phys. Rev. B **60**, 2020 (1999).
- 31) M. Machado, P. Ordejón, D. Sánchez-Portal, E. Artacho, and J. M. Soler: submitted to J. Phys. Chem. B.
- 32) E. Artacho, D. Sánchez-Portal, P. Ordejón, and J. M. Soler: to be published.
- 33) G. Fabricius, E. Artacho, D. Sánchez-Portal, P. Ordejón, D. A. Drabold, and J. M. Soler: Phys. Rev. B **60**, 16283 (1999).
- 34) M. J. Regan et al.: Phys. Rev. Lett. **54**, 9730 (1996) and references therein.
- 35) J. Wang, J. Hallmark, D. S. Marshall, W. J. Ooms, P. Ordejón, J. Junquera, D. Sánchez-Portal, E. Artacho, and J. M. Soler: Phys. Rev. B (Rapid Comm.), **60**, 4968 (1999).
- 36) J. I. Pascual, J. Gómez-Herrero, C. Rogero, A. Baró, D. Sánchez-Portal, E. Artacho, P. Ordejón, and J. M. Soler: Chem. Phys. Lett. **321**, 78 (2000).
- 37) J. Junquera, R. Weht, and P. Ordejón: to be published.
- 38) P. A. Fedders, D. A. Drabold, P. Ordejón, G. Fabricius, D. Sanchez-Portal, E. Artacho, and J. M. Soler: Phys. Rev. B **60**, 10594 (1999).