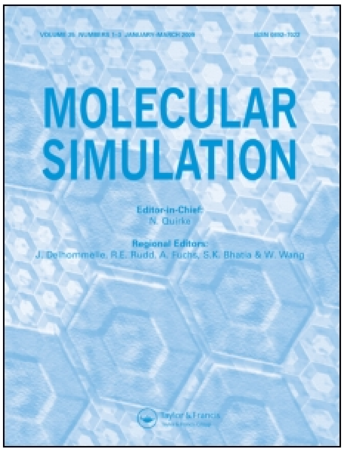

Molecular Simulation

Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

An ab initio Linear-Scaling Schem

P. D. Haynes[a]; M. C. Payne[a]
[a] Theory of Condensed Matter, Cavendish Laboratory, University of Cambridge, Cambridge, UK

To cite this Article Haynes, P. D. and Payne, M. C.(2000) 'An ab initio Linear-Scaling Schem', Molecular Simulation, 25: 3, 257 — 264

To link to this Article: DOI: 10.1080/08927020008044129

URL: http://dx.doi.org/10.1080/08927020008044129

Molecular Simulation, 2000, Vol. 25, pp. 257–264
Reprints available directly from the publisher
Photocopying permitted by license only

Published by license under
the Gordon and Breach Science
Publishers imprint.
Printed in Malaysia.

AN *AB INITIO* LINEAR-SCALING SCHEME

P. D. HAYNES* and M. C. PAYNE

*Theory of Condensed Matter, Cavendish Laboratory, University of Cambridge,
Madingley Road, Cambridge CB3 OHE, UK*

(Received April 1999; accepted May 1999)

In this paper we briefly survey the current state of *ab initio* calculations in terms of the accuracy and range of applicability of these methods for studying complex processes in real materials. We highlight some of the successes and limitations of these techniques and discuss the extent to which linear-scaling methods are able to extend the scope and scale of *ab initio* calculations. We argue that a combination of linear-scaling methods and hybrid modelling schemes is required to overcome many of the difficulties currently faced by conventional schemes, and present our own contributions towards the development of a robust and reliable linear-scaling method.

Keywords: Conventional techniques; linear-scaling techniques; hybrid modelling schemes; penalty-functional method; computer-simulation; electronic-structure calculations; *ab-initio* calculations; embedded cluster; density-functional theory

INTRODUCTION

The power of *ab initio* calculations has increased rapidly in recent years. In this paper we start by reviewing the present capability of first-principles calculations and point out some of the limitations of present conventional techniques. We then focus on linear-scaling techniques and discuss the extent to which they do represent a significant step forward in first-principles approaches. We review hybrid modelling schemes which also provide a route to much larger system-sizes and explain why linear-scaling approaches are the method of choice for the first-principles element of such schemes. Finally we briefly describe our work on the development of an *ab initio* linear-scaling code based upon a penalty-functional method.

---

*Corresponding author. e-mail: pdh1001@cam.ac.uk

CONVENTIONAL TECHNIQUES

The speed of progress in first-principles calculations over the fifteen years since Car and Parrinello's introduction of a unified scheme for density-functional theory (DFT) and molecular dynamics [1] has been breathtaking. The combination of the universal adoption of density-functional theory [2], significant improvements in numerical methods [3] and the availability of powerful computers has allowed first-principles studies of systems containing as many as a thousand atoms from the whole periodic table, typically yielding predictions of physical and chemical properties with an accuracy of a few percent. A critical aspect of calculations on systems of this size and complexity is the optimisation of atomic positions, since experiment is not capable of providing this information to sufficient accuracy. It is even possible to perform dynamical simulations in which the forces on the atoms are calculated from first principles and to study quantum effects in the ionic system using Feynman's path integral formulation [4].

While the power of modern first-principles calculations is impressive, it is important to appreciate that accessible system-sizes and time-scales are actually very limited: a system containing a thousand atoms typically has a volume of only 20 cubic angstroms. Although structural optimisation is possible in most first-principles codes, in general the true minimum energy ionic configuration will only be found if the connectivity of the ionic system is correctly specified. While dynamical simulations are in principle possible in any first-principles scheme that calculates ionic forces, in practice only the total energy pseudopotential technique has successfully used dynamics to address real scientific questions.

LINEAR-SCALING TECHNIQUES

The computational cost of conventional wavefunction-based total energy schemes scales at least as the cube of the number of atoms in the system. Many linear-scaling first-principles schemes have been proposed over the years [5–8]. While these schemes will allow calculations for very large numbers of atoms, there are other technical problems that may prevent the extraction of useful scientific information from such calculations. In particular, the length-scales and, more importantly, time-scales are still extremely limited. A system containing a million atoms is still only 200 cubic angstroms in volume, and the relevant time-scales (or equivalently the complexity of the phase space) increase very rapidly with system-size.

Furthermore, in such a large simulation most of the atoms will be very close to their bulk configuration and, even using a linear-scaling algorithm, considerable computational effort will be expended describing regions of the system where a full first-principles calculation is not required.

HYBRID MODELLING SCHEMES

A schematic illustration of a hybrid modelling scheme applied to a block of material containing two cracks is shown in Figure 1. *Ab initio* techniques are applied in the lightest shaded regions containing the crack tip, where accurate quantum-mechanical description of the bond-breaking process is essential. In the surrounding region, empirical interatomic potentials are used, and in the darkest shaded region continuum modelling is used. This approach allows each region of a system to be described by the modelling technique best-suited to the accuracy required in each region and also addresses the problem of multiple length-scales.

Hybrid modelling schemes incorporating empirical potentials in the atomistic region and a finite element description of the continuum regions have been developed [8a]. However, first-principles calculations have been incorporated in such schemes to a far lesser extent [8b]. The reason is

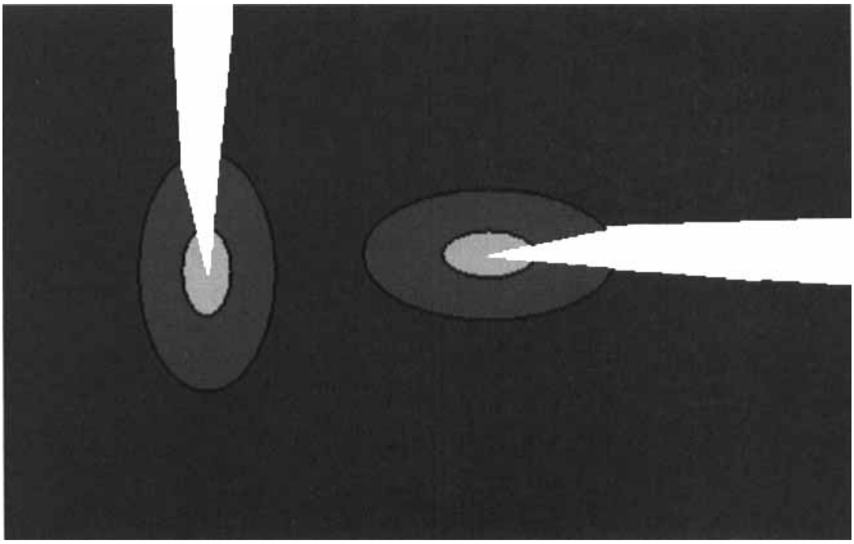

FIGURE 1 Schematic illustration of a hybrid modelling scheme.

primarily that the standard cluster and supercell based techniques are not easily included in such hybrid modelling schemes. Cluster models for extended systems are known to be very sensitive to the termination of the cluster, and supercells have serious topological restrictions due to the need to impose periodicity *e.g.*, it is impossible to model a single dislocation or surface in a supercell. In contrast, linear-scaling schemes need have no toplogical restrictions and provide a controlled termination of the quantum mechanics and thus may be incorporated into hybrid modelling schemes.

PENALTY-FUNCTIONAL METHOD

The cubic scaling of traditional *ab initio* methods arises in general either from the cost of diagonalizing the Hamiltonian or from orthogonalizing the wavefunctions in iterative schemes. Most linear-scaling techniques have been developed by reformulating the quantum-mechanical equations in terms of generalized Wannier functions or the single-particle density-matrix, instead of the extended single-particle wavefunctions. Enforcing the localization of the Wannier functions or truncating the density-matrix beyond a certain range results in a linear-scaling method. However, although many such schemes have been proposed [5–8], progress towards a truly general purpose scheme has been rather slow. Some of the problems encountered in implementing linear-scaling schemes are the existence of multiple minima in the energy surface, and slow convergence towards the electronic ground-state, even when the correct minimum is located. The result is a method with a computational cost that is linear in the number of atoms but with a very large prefactor so that the linear-scaling calculation only becomes faster than conventional schemes for extremely large systems.

In order to address the problems mentioned above, we are attempting to formulate a linear-scaling scheme with the following features:

- Single minimum in the search space.
- Fast, guaranteed convergence to this minimum from any starting configuration.
- Computational cost and memory as small as possible making the technique competitive with conventional techniques even for moderately sized systems.

Ultimately, we wish to have a linear-scaling code with the same reliability and efficiency as present wavefunction-based *ab initio* codes.

As mentioned above, the cubic scaling of conventional techniques results from the cost of solving Schrödinger's equation for the single-particle wavefunctions which are the eigenfunctions of the Kohn-Sham Hamiltonian. Within a self-consistent *ab initio* scheme it is also necessary to find the ground-state electronic density, but since this stage of the calculation already scales linearly with system-size, we will only address the solution of Schrödinger's equation for a fixed electronic density (and thus Hamiltonian). In a conventional scheme this may be achieved by diagonalizing the Hamiltonian (directly or iteratively) in the representation of some set of basis functions. Plane-waves and Gaussians have been popular choices for this purpose. We have chosen to use truncated spherical-waves [9] for this purpose for the following reasons:

- Truncated spherical-waves naturally obey the boundary conditions imposed to truncate the density-matrix (or localize the Wannier functions).
- The completeness of the basis set may be controlled by means of a kinetic energy cut-off, as for plane-waves.
- Meaningful comparison may thus be made to the results of plane-wave calculations.
- Quantities such as the overlap matrix and matrix elements of the kinetic energy and non-local pseudopotential operators may be calculated analytically.

Instead of solving for the extended wavefunctions, we choose to solve for the single-particle density-matrix which contains the same information but is not subject to the constraint of orthogonality and is short-ranged so that it may be truncated yielding a linear-scaling technique. In conventional methods, the lowest energy wavefunctions are occupied according to Pauli's exclusion principle, but in density-matrix-based schemes, the occupation numbers of these wavefunctions become variable parameters and it is necessary to impose the following constraints:

- Normalization: the occupation numbers must sum to the correct number of electrons in the system.
- Idempotency: the occupation numbers should be zero or unity, in order to obey Pauli's exclusion principle.

The first constraint is straightforward to impose, but the second, namely that the density-matrix and its square are identical, is non-linear and hence rather problematic. Some methods [6] have used the purifying transformation [10] to impose this constraint implicitly, but it has been argued [11] that this scheme may suffer from a rate of convergence that is less than optimal. Kohn [7] first proposed the use of a penalty-functional for imposing the idempotency constraint, but his use of a non-analytic functional to obtain a variational principle rules out the application of current efficient algorithms to this scheme [12]. We have proposed the use of an analytic penalty-functional [8] to approximately impose the idempotency constraint. The advantages of this method are an improved rate of convergence and a single minimum in the search space. However, the disadvantage of the scheme is that this minimum is not obtained for the idempotent ground-state density-matrix, but a nearly-idempotent approximation to it, which approaches the correct limit as the strength of the penalty-functional is increased. The solution to this problem has been the development of an analytic correction to the estimated total energy which gives accurate values for the true ground-state energy from minimizing density-matrices which need not be idempotent. The method has so far been applied to crystalline silicon and yields results in good agreement with experiment and conventional plane-wave calculations. In Figure 2 we plot the

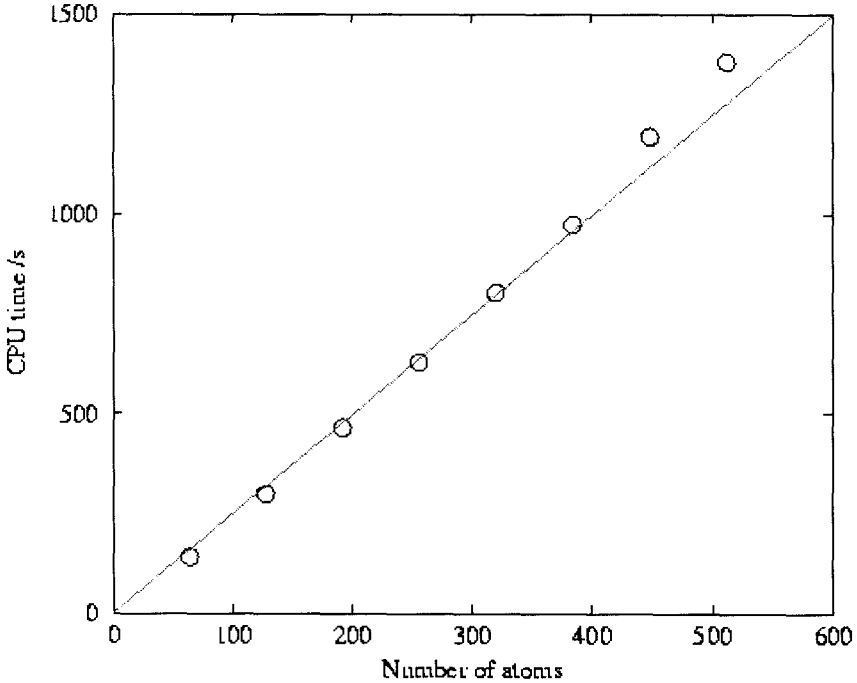

FIGURE 2 Scaling of the penalty-functional method with system-size.

CPU time per iteration *versus* the system-size for the penalty-functional method running on a DEC 500 au computer with 96 Mb of memory, and thus demonstrate the linear scaling of the method.

CONCLUSIONS

In conclusion, we have shown that although linear-scaling techniques are a significant step forward in the development of *ab initio* quantum-mechanical calculations with respect to the range of applications to which such methods may be applied, they do not in themselves overcome all of the limitations of conventional techniques. We have argued that the combination of *ab initio* linear-scaling techniques with hybrid modelling schemes is essential in order to efficiently and accurately model large-scale processes in real materials which are of interest to scientists and engineers today. We have outlined some of our work towards achieving this goal, namely the development of a penalty-functional-based *ab initio* linear-scaling scheme.

Acknowledgment

This work was supported by grant GR/L55346 from the Engineering and Physical Sciences Research Council, UK.

References

[1] Car, R. and Parrinello, M. (1985). *Phys. Rev. Lett.*, **55**, 2471.

[2] Hohenberg, P. and Kohn, W. (1964). *Phys. Rev.*, **136**, 864; Kohn, W. and Sham, L. J. (1965). *ibid.*, **140**, 1133; for a review see, for instance, Jones, R. O. and Gunnarsson, O. (1989). *Rev. Mod. Phys.*, **61**, 689.

[3] Payne, M. C., Teter, M. P., Allan, D. C., Arias, T. A. and Joannopoulos, J. D. (1992). *Rev. Mod. Phys.*, **64**, 1045.

[4] See, for instance, Allen, M. P. and Tildesley, D. J. (1987). *Computer Simulation of Liquids*, pp. 270ff. (Oxford University Press).

[5] Yang, W. (1991). *Phys. Rev. Lett.*, **66**, 1438; Yang, W. and Lee, T.-S. (1995). *J. Chem. Phys.*, **103**, 5674; Baroni, S. and Giannozzi, P. (1992). *Europhys. Lett.*, **17**, 547; Drabold, D. A. and Sankey, O. F. (1993). *Phys. Rev. Lett.*, **70**, 3631; Wang, L.-W. (1994). *Phys. Rev. B*, **49**, 10154; Sankey, O. F., Drabold, D. A. and Gibson, A. (1994). *ibid.*, **50**, 1376; Silver, R. N. and Röder, H. (1997). *Phys. Rev. E*, **56**, 4822; Qiu, S.-Y., Wang, C. Z., Ho, K. M. and Chan, C. T. (1994). *J. Phys.: Condens. Matter*, **6**, 9153; Canning, A., Galli, G., Mauri, F., de Vita, A. and Car, R. (1996). *Comp. Phys. Comm.*, **94**, 89; Horsfield, A. P., Bratkovsky, A. M., Pettifor, D. G. and Aoki, M. (1996). *Phys. Rev. B*, **53**, 1656; Wang, Y. *et al.* (1995). *Phys. Rev. Lett.*, **75**, 2867; Abrikosov, I. A. *et al.* (1996). *Phys. Rev. Lett.*, **76**, 4203; Abrikosov, I. A. *et al.* (1997). *Phys. Rev. B*, **56**, 9319; Smargiassi, E. and Madden, P. A. (1994). *ibid.*, **49**, 5220; Foley, M. and Madden, P. A. (1996). *ibid.*, **53**, 10589; Goedecker, S. and Colombo, L. (1994). *Phys. Rev. Lett.*, **73**, 122; Goedecker, S. and Teter, M. (1995). *Phys. Rev. B*, **51**, 9455; Baer, R. and Head-Gordon, M. (1997). *J. Chem. Phys.*, **107**, 10003; Stephan, U. and

Drabold, D. A. (1998). *Phys. Rev. B*, **57**, 6391; Nicholson, D. M. C. and Zhang, X.-G. (1997). *ibid.*, **56**, 12805; Gagel, F. (1998). *J. Comp. Phys.*, **139**, 399; Silver, R. N., Röder, H., Voter, A. F. and Kress, J. D. (1996). *ibid.*, **124**, 115; Voter, A. F., Kress, J. D. and Silver, R. N. (1996). *Phys. Rev. B*, **53**, 12733; Galli, G. and Parrinello, M. (1992). *Phys. Rev. Lett.*, **69**, 3547; Mauri, F., Galli, G. and Car, R. (1993). *Phys. Rev. B*, **47**, 9973; Mauri, F. and Galli, G. (1994). *ibid.*, **50**, 4316; Ordejón, P., Drabold, D. A., Grumbach, M. P. and Martin, R. M. (1993). *ibid.*, **48**, 14646; Ordejón, P., Drabold, D. A., Martin, R. M. and Grumbach, M. P. (1995). *ibid.*, **51**, 1456; Kim, J., Mauri, F. and Galli, G. (1995). *ibid.*, **52**, 1640; Pandey, K. C., Williams, A. R. and Janak, J. F. (1995). *ibid.*, **52**, 14415; Ordejón, P., Artacho, E. and Soler, J. M. (1996). *ibid.*, **53**, 10441; for reviews see, for instance, Bowler, D. R. *et al.* (1997). *Modelling Simul. Mater. Sci. Eng.*, **5**, 199; Galli, G. (1996). *Current Opinion in Solid State and Materials Science*, **1**, 864.

[6] Li, X.-P., Nunes, R. W. and Vanderbilt, D. (1993). *Phys. Rev. B*, **47**, 10891; Daw, M. S. (1993). *ibid.*, **47**, 10895; Stechel, E. B., Williams, A. R. and Feibelman, P. J. (1994). *ibid.*, **49**, 10088; Hierse, W. and Stechel, E. B. (1994). *ibid.*, **50**, 17811; Hernández, E. and Gillan, M. J. (1995). *ibid.*, **51**, 10157; Hernández, E., Gillan, M. J. and Goringe, C. M. (1996). *ibid.*, **53**, 7147.

[7] Kohn, W. (1996). *Phys. Rev. Lett.*, **76**, 3168.

[8] Haynes, P. D. and Payne, M. C. (1999). *Phys. Rev. B*, **59**, 12173.

[8a] Shenoy, V. B., Miller, R., Tadmor, E. B., Phillips, R. and Ortiz, M. (1998). *Phys. Rev. Lett.*, **80**, 742; Miller, R., Tadmor, E. B., Phillips, R. and Ortiz, M. (1998). *Modelling Simul. Mater. Sci. Eng.*, **6**, 607; Shenoy, V. B., Miller, R., Tadmor, E. B., Rodney, D., Phillips, R. and Ortiz, M. (1999). *J. Mech. Phys. Solids*, **47**, 611; Tadmor, E. B., Smith, G. S., Bernstein, N. and Kaxiras, E. (1999). *Phys. Rev. B*, **59**, 235; Abraham, F. F., Broughton, J. Q., Bernstein, N. and Kaxiras, E. (1998). *Europhys. Lett.*, **44**, 783; Rudd, R. E. and Broughton, J. Q. (1998). *Phys. Rev. B*, **58**, 5893.

[8b] Eichler, U., Kolmel, C. M. and Sauer, J. (1997). *J. Comput. Chem.*, **18**, 463; Eichler, U., Brandle, M. and Sauer, J. (1997). *J. Phys. Chem. B*, **101**, 10035; Brandle, M., Sauer, J., Dovesi, R. and Harrison, N. M. (1998). *J. Chem. Phys.*, **109**, 10379; Brandle, M. and Sauer, J. (1998). *J. Am. Chem. Soc.*, **120**, 1556; Svensson, M., Humbel, S., Froese, R. D. J., Matsubara, T., Sieber, S. and Morokuma, K. (1996). *J. Phys. Chem.*, **100**, 19357; Matsubara, T., Sieber, S. and Morokuma, K. (1996). *Int. J. Quantum Chem.*, **60**, 1101; Froese, R. D. J., Musaev, D. G. and Morokuma, K. (1998). *J. Am. Chem. Soc.*, **120**, 1581.

[9] Haynes, P. D. and Payne, M. C. (1997). *Comp. Phys. Comm.*, **102**, 17.

[10] McWeeny, R. (1960). *Rev. Mod. Phys.*, **32**, 335.

[11] Goedecker, S. (1999). *Rev. Mod. Phys.*, **71**, 1085.

[12] Haynes, P. D. and Payne, M. C. (1998). *Solid State Communications*, **108**, 737.