

Implementation of a Divide and Conquer Strategy for the Materials Modelling Code CRYSTAL

Daniel R. Jones

Numerical Algorithms Group, Oxford.

December 6, 2011

Abstract

CRYSTAL is an *ab initio* electronic structure and materials properties code, in which the electron density is represented as a linear combination of Bloch functions constructed from localised Gaussian type atomic orbitals. This dCSE project introduces a new algorithm, the divide and conquer algorithm for computing the electronic structure that scales linearly with system size. This report outlines the algorithm used and shows that it produces similar results to the conventional CRYSTAL algorithms for a weakly interacting system. Continuations of this work are identified and outlined, these are necessary for its more general application.

Contents

1	Introduction	3
1.1	CRYSTAL Materials Modelling Code	3
1.2	Linear Scaling in DFT	3
1.3	Original dCSE Application Milestones	4
2	Divide and Conquer Algorithm	4
3	Proof of Principle Example	6
3.1	Liquid Neon	6
4	Performance Tests	7
5	Additional Useful Outcomes from this dCSE Project	8
5.1	Task Level Parallelism	8
5.2	Improved Initial Guess of the Density Matrix	9
6	Limitations	10
7	Conclusions	10
8	Acknowledgements	11

1 Introduction

1.1 CRYSTAL Materials Modelling Code

The CRYSTAL materials modelling code [1–3] is a program developed by the Computational Science and Engineering Department at STFC, and the theoretical chemistry group at Turin University (Italy). Its function is to compute, *ab initio*, the electronic structure of a material and from this many material properties. The one electron wavefunctions in CRYSTAL are stored as a linear combination of localised Gaussian functions, which enables CRYSTAL to evaluate the electronic structure for aperiodic systems, and systems that are periodic in one, two or three dimensions efficiently. The most popular alternative, using a plane wave basis set, requires a vacuum gap between periodic images which is computationally wasteful for systems that are not periodic in three dimensions.

A significant advantage of using a local basis set, is that it enables relatively inexpensive use of the Hartree-Fock (HF) method and hybrid density functional theory (DFT) which incorporates a fraction of the “exact” HF exchange. These hybrid density functionals are becoming increasingly widely used, as their ability to more accurately model material properties that are poorly represented by either HF or local and semi-local DFT functionals is repeatedly reported [4–6]. The other significant advantage of using a local basis set, is that all electron calculations can be performed routinely. Pseudopotentials to represent core electrons are not required as is almost always the case when using a plane wave basis set.

The CRYSTAL code is widely used internationally by quantum chemists, solid state physicists and materials scientists, and has been for many years. It is particularly important for UK materials chemists. The functionality and performance of CRYSTAL continues to be improved. As the size and complexity of the systems of interest increase, it is necessary to continue to develop CRYSTAL to make best use of high performance computing (HPC) resources, such as HECToR.

Currently CRYSTAL can be run in parallel using either a replicated data strategy using the program called PCRYSTAL, or alternatively using a distributed data strategy for large ($O(N^2)$ or greater) objects, using the program called MPP CRYSTAL, developed by Dr. I. J. Bush [2]. The systems for which the divide and conquer algorithm is most appropriate are large, therefore the memory requirement means that a replicated data strategy cannot be used. As such, the work described in this report is an extension to MPP CRYSTAL.

1.2 Linear Scaling in DFT

The most commonly used algorithms to self-consistently find the electron density of a material, including the standard algorithms in the CRYSTAL code, scale as the third power of the system size, i.e. $O(N^3)$. In CRYSTAL, for large systems, the bottle neck is the explicit diagonalisation of the Hamiltonian (Fock or Kohn-Sham) matrix, an $O(N^3)$ operation. For codes that do not diagonalise the Hamiltonian, the cubic scaling occurs because N states need to be found which are $O(N)$ size and also must be orthogonal to the other N states. This unfavourable scaling limits traditional DFT computations to systems with a maximum size of a few thousand of atoms. Computations for systems larger

than a few hundred atoms are uncommon.

To enable the study of large systems, the $O(N^3)$ scaling must be reduced. Fortunately, the interaction between electrons in a many electron system is almost always short ranged as the density matrix decays exponentially with distance [7]. There have been several strategies developed to exploit this “near-sightedness” of electrons in a variety of codes including ONETEP [8], CONQUEST [9] and SIESTA [10]. The divide and conquer algorithm chosen for this project is described later in Section 2.

1.3 Original dCSE Application Milestones

The aim of this dCSE project is to implement an $O(N)$ divide and conquer algorithm for use within the CRYSTAL code. The details of this algorithm are given in Section 2. The major milestones outlined in the dCSE proposal are as follows:

1. Validation of MPP code working within a task farming harness.
2. Automatic decomposition of a system into subsystems and distribution of work so that subsystems are run independently within taskfarmed instances of MPPCRYSTAL.
3. Implement the communication of eigenvectors and eigenvalues to determine a global Fermi energy and from this reconstruct a global density matrix.
4. Decompose the density matrix as a multipole expansion and embed the subsystems within this multipole expansion in order to include long range electrostatics.

The original dCSE application was for two man years of effort on this project. The project was awarded 18 man months of effort of which 15 man months have been used.

2 Divide and Conquer Algorithm

The divide and conquer algorithm (D&C), is a relatively old idea [11] that is ideally suited to current HPC resources as there is significantly more computation than communication. Conceptually, it is a relatively simple scheme, using the “near-sightedness” of electrons, one can separate a system into subsystems comprising a *core* region and a *halo* region, as shown in Figure 1. The subsystems are decoupled, and the electronic structure of each subsystem is determined independently. The systems are then reconnected by identifying a common Fermi energy and a new global density matrix, from which the system’s total energy can be easily computed, is constructed by adding the contributions from the core regions. The electronic structure of the halo is discarded. The presence of the halo region is just to ensure the electronic structure of the core region is as close as possible to that of the reconstructed system.

More formally, a computation to calculate the single point energy of a system Ω , whose single particle wavefunctions are written as a linear combination of a

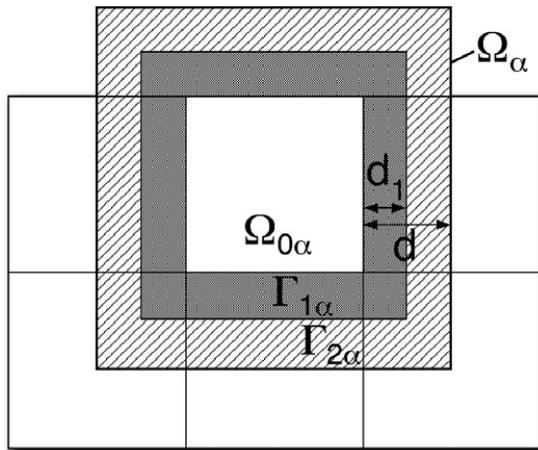


Figure 1: The system is partitioned into a set of subsystems, $\{\alpha\}$, each subsystem contains a core region $\Omega_{0\alpha}$ and a halo region $\Gamma_{2\alpha}$ and, optionally, a region which contributes partially, $\Gamma_{1\alpha}$. The electronic structure of the subsystems is determined independently and from these computations the global electronic structure is reconstructed. Figure from ref. [12].

localised basis set, $\{i\}$, can be approximately solved by dividing Ω into a set of subsystems, $\{\alpha\}$, and defining a partition matrix

$$\mathbf{P}_{ij} = \sum_{\alpha} \mathbf{P}_{ij}^{\alpha} = 1. \quad (1)$$

Once the electronic structure of the independent subsystems has been determined, using a traditional DFT or HF algorithm, a global Fermi energy must be identified that returns the (known) correct number of electrons, N_{el} . In the case of a non-magnetic system, where all electrons are paired,

$$N_{el} = \sum_{ij} \rho_{ij} \mathbf{S}_{ij} = \sum_{ij} \left(2 \sum_{\alpha} \mathbf{P}_{ij}^{\alpha} \sum_m f_{\beta}(\varepsilon_F - \varepsilon_m) \mathbf{C}_{im}^{\alpha} \mathbf{C}_{jm}^{\alpha\dagger} \right) \mathbf{S}_{ij}, \quad (2)$$

where ρ_{ij} is the (global) density matrix, $\mathbf{S}_{ij} = \langle i|j \rangle$ is the overlap matrix, f_{β} is the Fermi function at an arbitrary electronic temperature β and \mathbf{C}_{im}^{α} is the matrix representation of the eigenfunctions of subsystem α . The Fermi energy ε_F that satisfies (2) is found iteratively. Once the Fermi energy has been identified, the density matrix for the system can be computed by summing the contributions from the subsystems, $\{\alpha\}$,

$$\rho_{ij} = 2 \sum_{\alpha} \mathbf{P}_{ij}^{\alpha} \sum_m f_{\beta}(\varepsilon_F - \varepsilon_m) \mathbf{C}_{im}^{\alpha} \mathbf{C}_{jm}^{\alpha\dagger}. \quad (3)$$

The accuracy of the electronic structure computed using this method varies depending on the partitioning of the system into subsystems. During this dCSE project, two methods for partitioning the system have been implemented. The simplest, most conservative method is for each subsystem to contain a single core

atom. This method has been implemented in SIESTA [10] and is equivalent to a Mulliken population analysis. This method has a single parameter, the radius of the halo around the core atom. The appropriate size for the halo depends on the material being studied and can be defined in the CRYSTAL input file. The partition matrix for the system is defined as follows:

$$\mathbf{P}_{ij}^{\alpha} = \begin{cases} 1, & \text{if } i \in \Omega_{\alpha} \wedge j \in \Omega_{\alpha} \\ 0.5, & \text{if } (i \in \Omega_{\alpha} \wedge j \ni \Omega_{\alpha}) \vee (i \ni \Omega_{\alpha} \wedge j \in \Omega_{\alpha}) \\ 0, & \text{if } i \ni \Omega_{\alpha} \wedge j \ni \Omega_{\alpha}, \end{cases} \quad (4)$$

where Ω_{α} is the set of basis functions centred on the core atom. This very simple partitioning of the system is likely to be quite inefficient, it is simple to improve performance by including several atoms in the core region of each subsystem. However, it has been noted that having multiple core atoms per subsystem may increase the likelihood of having discontinuities in the potential energy surface [10]. One core atom per subsystem is the safest default option.

The alternative solution implemented in CRYSTAL (during this dCSE project) is to leave the definition to the user and explicitly define each subsystem by selecting a set of atoms (or ghost atoms) and defining the partition matrix in atom-by-atom terms by hand. The way it has been implemented, \mathbf{P}_{ij}^{α} must be identical for all basis functions localised on the same atom. This allows users to have the flexibility to use the D&C algorithm in the way that is most suited to their problem. The code which partitions the systems into subsystems has been designed so that new heuristics for defining the partition matrix can be added at a later date without altering the structure of the code. The newly implemented D&C algorithm simply needs a mapping array, which associates atoms in the subsystems with atoms in the reconstructed system, and a partition matrix for each subsystem. The \mathbf{P}_{ij}^{α} matrices defined must satisfy Equation (1).

In the time available, point 4 on the list of milestones has not yet been addressed. The long range electrostatics of the system have been discarded, for systems that are largely non-polar, this should make no difference and the electrostatics included by the interaction of the core with the halo region should be sufficient, as will be shown in the next section.

3 Proof of Principle Example

3.1 Liquid Neon

To initially test that the D&C algorithm returns an acceptable approximation of the density matrix, a simple test case was used. The D&C algorithm is most appropriate for weakly interacting systems. Long range Coulombic interactions are not yet included in the CRYSTAL implementation and periodic covalently bonded systems cause inconsistencies in the subsystems' electronic structure, see Section 6. A 10 \AA^3 box of neon atoms created using a short classical molecular dynamics simulation¹ was used as a test case, see Figure 2.

The system was partitioned into subsystems using one core-atom subsystems, and a set of partition matrices as defined in Equation (4). The convergence of

¹Molecular dynamics simulation was run using the Tinker molecular modelling package [13] with the MM3 forcefield using an NVT ensemble at 50K for 50 ps

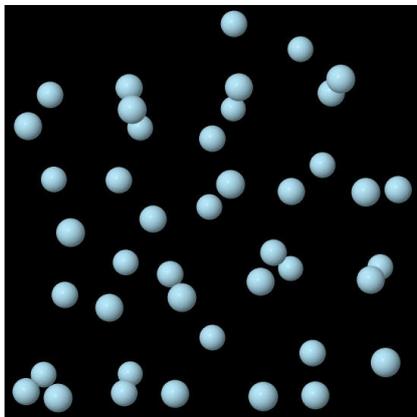


Figure 2: The system used to test that the density matrix partitioning and reconstruction functioned correctly. 10 \AA^3 of neon atoms in positions determined by a short classical molecular dynamics computation.

the system energy as a function of halo radius is reported in Table 1. It can be seen that as the size of the halo is increased the energy converges on the result of the traditional Hamiltonian diagonalisation used in standard CRYSTAL calculations. This is unsurprising considering the weak interatomic interactions in this system, but indicates that the divide and conquer implementation correctly partitions the system and reconstructs the density matrix based on the subsystem computations.

4 Performance Tests

To measure the performance of the D&C algorithm with increasing system size, the density matrix was computed for various supercells of crystalline neon with

Table 1: Convergence on the DFT energy for the D&C algorithm as the radius of the halo increases. $E_{\text{D\&C}}$ is the energy computed using the D&C algorithm, E_{trad} is the energy computed using Hamiltonian Matrix diagonalisation — the standard CRYSTAL algorithm.

Radius of Halo / \AA	$E_{\text{D\&C}} - E_{\text{trad}} /$ millihartree cell $^{-1}$
0.1	17.3
3.0	5.1
4.5	2.6

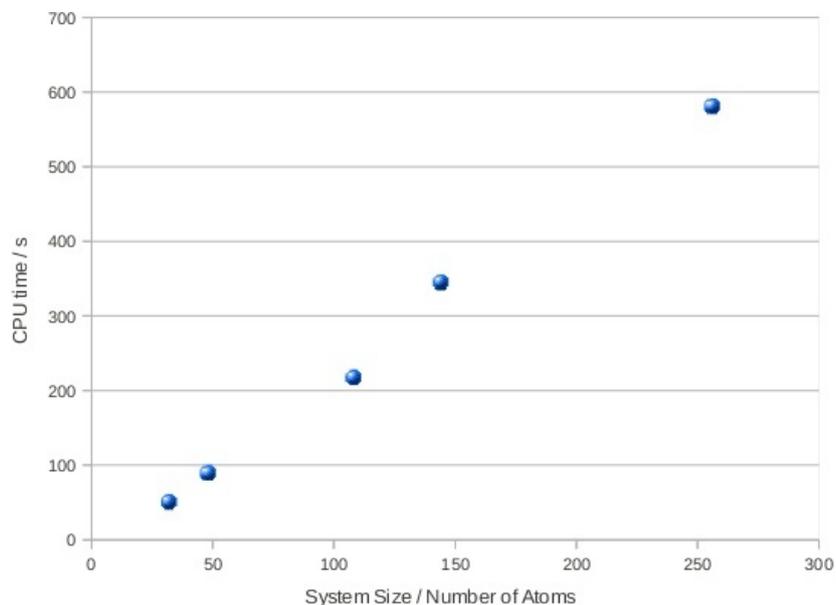


Figure 3: The time taken to compute the density matrix using the divide and conquer algorithm as a function of system size. The algorithm displays near linear scaling.

a face centred cubic crystal structure (from the Inorganic Crystal Structure Database [14]). The time taken to compute the density matrix with varying systems size is shown in Figure 3. The scaling is nearly linear. The non-linear part of the scaling comes from computing the Fermi energy for the system. Figure 4 shows the time taken to set up and compute the electronic structure of the subsystems. This shows almost perfect linear scaling, and also shows that a substantial part of the run time is spent computing the Fermi energy. Improvements to the algorithm to compute the Fermi energy should be a priority in the continuation of this work.

5 Additional Useful Outcomes from this dCSE Project

5.1 Task Level Parallelism

Taskfarming was introduced to the CRYSTAL code as part of this dCSE project. This functionality is essential for the use of the D&C algorithm. Finding the electronic structure of subsystems sequentially is comparatively highly inefficient. There are several other problems that will benefit from taskfarming being available in CRYSTAL. Notably, CRYSTAL computes the second derivative of the potential energy surface with respect to atomic positions using numerical finite differences in order to compute vibrational frequencies and phonon dispersion. This task is ideally suited to a taskfarming strategy to achieve maximal

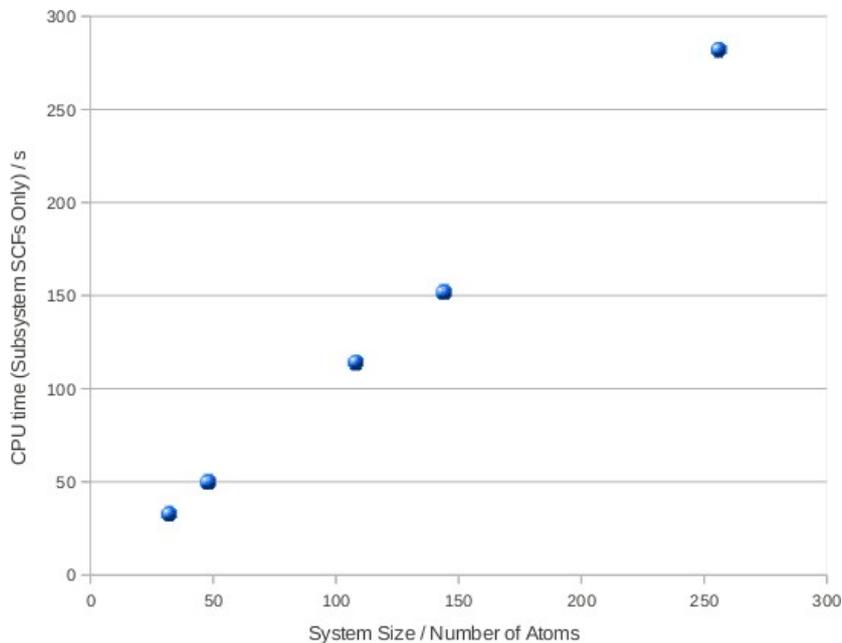


Figure 4: The CPU time taken to converge the SCF for all subsystems using the D&C algorithm. This shows near perfect linear scaling. Further investigation is needed to reduce the time taken to calculate the Fermi energy for the system.

speed up on HPC resources. The utilisation of the taskfarming code introduced as part of this dCSE project is being investigated by Matteo Ferabone at Torino University, Italy.

Many scientific problems are studied by finding and comparing the electronic structure of many similar arrangements of atoms. For example, collaborators at Imperial College London have produced phase diagrams for chalcopyrites as the chemical potential of the constituent elements is varied. This sort of computation, where many similar but independent computations need to be performed may benefit from the task level parallelism introduced as part of this dCSE project.

5.2 Improved Initial Guess of the Density Matrix

One of the challenges associated with using a Gaussian basis set electronic structure code is that the convergence of the SCF depends on the quality of the initial guess. During this project, code has been written which allows the density matrix to be constructed based on smaller (usually simpler and less expensive) computations. There are several instances where this code may be useful, for example:

- Molecular crystals — the initial guess electron density can be constructed based on the converged electronic structure of the constituent molecules.
- Macromolecules, e.g. DNA, proteins, polymers — the initial guess elec-

tron density can be constructed from converged electronic structure of monomer units.

- Supercells, e.g. for studying defect science — while this is a more complex use of the code requiring a little more thought, the bulk electron density can be used to construct the guess for the “bulk like” region of the supercell.

6 Limitations

The taskfarming in CRYSTAL, as implemented as part of this dCSE project, requires the number of taskfarms to be an exact divisor of the number of processors. That is, an equal number of processors are used for each task. It is possible a more sophisticated division of the processors may be desirable. In addition, the round robin implementation may need to be improved if load balancing for the taskfarms proves to be a bottleneck.

The current implementation of D&C in CRYSTAL does not include long range Coulombic interactions. This could be corrected by performing a multipole expansion of the charge distribution and embedding the cluster subsystems in a field created by static multipoles. An Ewald summation could be used to correctly represent the long range electrostatic potential of the periodic system. The machinery for much of this work exists in the CRYSTAL code base, time constraints meant that it was not possible to apply these to this algorithm during this project.

For periodic covalently bonded systems, e.g. silicon, the D&C algorithm requires modification to prevent the division of the system into subsystems causing undercoordinated terminal atoms. Simply terminating the subsystem at a defined halo radius leads to unphysical “dangling bonds.” which cause the subsystem’s electronic structure to differ significantly from the global system, the subsystems are unphysically spin polarised. The “dangling bonds” can be capped, usually with hydrogen, however a method to add these hydrogen atoms automatically has not been implemented yet. It is also unknown how the presence of capping hydrogen affects the convergence of the D&C algorithm with respect to the size of the halo region. This procedure, of predicting the location of hydrogen atoms in covalent systems, is similar to problems encountered when determining the crystal structure of biological systems by X-ray crystallography. It may be possible to apply similar heuristic approaches to locating hydrogen atoms to cap the clusters used in the D&C algorithm.

Further work is required to make the D&C algorithm generally useful, currently only weakly interacting systems such as molecular crystals can be studied using this method. However, there are several examples of *ab initio* studies of molecular crystals in the literature, e.g. [15, 16]. So even with these limitations this algorithm may be a useful alternative to the traditional Hamiltonian matrix diagonalisation approach for certain systems.

7 Conclusions

A divide and conquer algorithm to compute the energy of a system using density functional theory has been implemented in the CRYSTAL code. While more

work is required to apply this algorithm to extended covalent networks and ionic systems, the method developed can be used for the *ab initio* study of weakly interacting systems. As a result of this project task farming has been added to the CRYSTAL code and a method to improve the initial guess of the density matrix has been developed. Once methods for dealing with covalent systems and long range electrostatics have been added, optimisations for the memory requirements of this algorithm may be needed. Unfortunately, there has not been time to address this during this dCSE project.

8 Acknowledgements

This project was funded under the HECToR Distributed Computational Science and Engineering Service operated by NAG Ltd. HECToR: A Research Councils UK High End Computing Service — is the UK's national supercomputing service, managed by EPSRC on behalf of the participating Research Councils. Its mission to support capability science and engineering in UK academia. The HECToR supercomputers are managed by UoE HPCx Ltd. and the CSE Support Service is provided by NAG Ltd. <http://www.hector.ac.uk>. The author would very much like to thank the following people for their contribution to this work:

- Dr. Ian J. Bush
- Prof. Nicholas Harrison, Dr. Barbara Montanari and Dr. Leonardo Bernasconi.

References

- [1] R. Dovesi, V. R. Saunders, C. Roetti, R. Orlando, M. Zicovic-Wilson, F. Pascale, B. Civalleri, K. Doll, M. N. Harrison, I. J. Bush, P. D'Arco, and M. Llunell, *CRYSTAL09 User's Manual*. University of Torino, Torino, 2009.
- [2] I. J. Bush, S. Tomić, B. G. Searle, G. Mallia, C. L. Bailey, L. Bernasconi, B. Montanari, J. M. Carr, and N. M. Harrison, "Parallel implementation of the *ab initio* CRYSTAL program: Electronic structure calculations for periodic systems," *Proc. R. Soc. A*, vol. 467, pp. 2112–2126, 2011.
- [3] R. Dovesi, B. Civalleri, R. Orlando, C. Roetti, and V. R. Saunders, *Reviews in Computational Chemistry*, vol. 21, ch. 1, pp. 1–125. John Wiley & Sons, 2005.
- [4] F. Corà, "The performance of hybrid density functionals in solid state chemistry: the case of BaTiO₃," *Mol. Phys.*, vol. 103, pp. 2483–2496, 2005.
- [5] N. Martsinovich, D. R. Jones, and A. Troisi, "Electronic structure of TiO₂ surfaces and effect of molecular adsorbates using different DFT implementations," *J. Phys. Chem. C*, vol. 114, pp. 22659–22670, 2010.
- [6] J. Graziani, A. M. Marquez, J. J. Plata, Y. Ortega, N. C. Hernandez, A. Meyer, C. M. Zicovich-Wilson, and J. F. Sanz, "Comparitive study on

- the performance of hybrid DFT functionals in highly correlated oxides: The case of CeO_2 and Ce_2O_3 ,” *J. Chem. Theory Comput.*, vol. 7, pp. 56–65, 2011.
- [7] W. Kohn, “Density functional and density matrix method scaling linearly with the number of atoms,” *Phys. Rev. Lett.*, vol. 76, p. 3168, 1996.
- [8] C.-K. Skylaris, P. D. Haynes, A. A. Mostofi, and M. C. Payne, “Introducing ONETEP: Linear-scaling density functional simulations on parallel computers,” *J. Chem. Phys.*, vol. 122, p. 084119, 2005.
- [9] D. R. Bowler, T. Miyazaki, and M. J. Gillan, “Recent progress in linear scaling *ab initio* electronic structure techniques,” *J. Phys.: Condens. Matter*, vol. 14, p. 2781, 2002.
- [10] B. O. Cankurtaran, J. D. Gale, and M. J. Ford, “First principles calculations using density matrix divide-and-conquer within the SIESTA methodology,” *J. Phys.: Condens. Matter*, vol. 20, p. 294208, 2008.
- [11] W. Yang, “Direct calculation of electron density in density-functional theory: Implementation for benzene and a tetrapeptide,” *Phys. Rev. A*, vol. 44, pp. 7823–7826, 1991.
- [12] F. Shimojo, R. K. Kalia, A. Nakano, and P. Vashishta, “Embedded divide-and-conquer algorithm on hierarchical real-space grids: parallel molecular dynamics simulation based on linear-scaling density functional theory,” *Comp. Phys. Comm.*, vol. 167, pp. 151–164, 2005.
- [13] <http://dasher.wustl.edu/ffe/>.
- [14] J. de Smedt, W. H. Keesom, and H. H. Mooy, “On the crystal structure of neon,” *Koninklijke Nederlandse Akademie van Wetenschappen, Series B*, vol. 33, pp. 255–257, 1932.
- [15] K. Hannewald and P. A. Bobbert, “*Ab initio* theory of charge-carrier conduction in ultrapure organic crystals,” *Appl. Phys. Lett.*, vol. 85, p. 1535, 2004.
- [16] B. Civalleri, C. M. Zicovich-Wilson, L. Valenzano, and P. Ugliengo, “B3LYP augmented with an empirical dispersion term (B3LYP-D*) as applied to molecular crystals,” *CrystEngComm*, vol. 10, pp. 405–410, 2008.