

Large-scale DFT simulations with a linear-scaling DFT code CONQUEST on K-computer

Michiaki Arita^{1,2}, Sergiu Arapan¹, David R. Bowler^{3,4}, Tsuyoshi Miyazaki^{1,2,*}

¹Computational Materials Science Unit, National Institute for Materials Science

²Department of Physics, Science and Technology, Tokyo University of Science

³London Centre for Nanotechnology, University College London

⁴MANA, National Institute for Materials Science

*MIYAZAKI.Tsuyoshi@nims.go.jp

Received: April 8, 2014; Accepted: June 10, 2014; Published: October 31, 2014

Abstract. In order to realize the DFT simulations on large-scale complex systems, we have been developing a linear-scaling DFT code CONQUEST. In this paper, we report the parallel efficiency of the code on K-computer and show that it has almost ideal parallel efficiency even when we use more than 200,000 cores. Using the code on such large-scale parallel computers, we are now ready to do actual DFT study on million-atom systems. By showing our current study on the nucleation of Ge dimers on three-dimensional Ge nano-islands on Si(001), we demonstrate that accurate, efficient and robust structure relaxation based on the DFT is possible in the actual scientific research on complex nano-structured materials.

Keywords: first-principles calculations, density functional theory (DFT), linear-scaling or $O(N)$ methods

1. Introduction

First-principles calculation methods can provide reliable information of structures and properties of materials through electronic structure calculations based on the density functional theory (DFT). They have been playing important roles in various fields as a powerful research tool which can provide reliable data independently from experimental results. However, compared with the calculations using classical force fields, this method requires a large amount of computation time. In addition, the computational cost of the conventional first-principles methods increases rapidly when the size of target systems becomes larger, in proportion to the cube of the number of atoms N contained in the system. It has been very difficult to treat the systems containing more than one thousand atoms and the targets of the DFT study have been limited only for small parts of the real materials or idealized model systems.

On the other hand, the speed of supercomputers keeps growing rapidly. Roughly speaking, the peak performance of the supercomputers have been going up tenfold every four years. In June and November 2011, Japanese supercomputer named "K-computer" was ranked as the world's fastest computer in the TOP500 list [1]. (It is now ranked as the 4th.) The computer, whose performance exceeds 10 PFLOPS, was open to public in September 2012, bringing about a dramatic change in the computational materials science. By using such a huge and powerful computer, it is expected to realize first-principles study of large-scale systems having the size of complex materials or real devices. However, this would require high-efficiency in ultra-large-scale parallel calculations. The K-computer has 88,128 CPUs. As each CPU has 8 cores (128 gigaflops), it has more than 700,000 cores in total. It is not easy to achieve high parallel efficiency using such a large number of cores. Especially the efficiency of the planewave DFT methods, which are most widely used in the DFT calculations at present, strongly depends on the efficiency of the FFT libraries and it is usually difficult for FFT libraries to achieve high efficiency on massively parallel computers. In this respect, development of a new calculation method and advanced calculation technique is essential.

We have been developing a DFT code CONQUEST, which is designed for massively parallel calculations. In addition, the code uses a linear scaling (or $O(N)$) method [2] where the computation time and memory requirement are both only proportional to N . We have recently optimized the code on K-computer to realize the DFT study on very large systems. In this paper, we report the performance of the code on K-computer and show that the first-principles study on million-atom systems is now possible using such a big computer, with the $O(N)$ technique.

2. Linear-scaling DFT code CONQUEST

CONQUEST is a linear-scaling DFT code which enables us to perform the first-principles calculations on very large-scale systems. Since the details of the calculation method are explained in a series of previous papers [3, 4, 5], we only summarize main points here.

In the conventional DFT method, we calculate Kohn-Sham orbitals $\psi_n(\mathbf{r})$, which are the eigenfunctions of the Kohn-Sham equations. Here, n is the index of a band or an electronic state, and the number of n is proportional to N . In CONQUEST, instead of calculating the set of Kohn-Sham orbitals, we work on the Kohn-Sham density matrix ρ defined by summation of the occupied Kohn-Sham orbitals

$$\rho(\mathbf{r}, \mathbf{r}') = \sum_n f_n \psi_n(\mathbf{r}) \psi_n^*(\mathbf{r}'). \quad (1)$$

Here, f_n is the occupation number for the state n . The DFT total energy (Kohn-Sham total energy) can be calculated from the density matrix. In CONQUEST, the density matrix is represented in terms of localized orbitals, called "support functions":

$$\rho(\mathbf{r}, \mathbf{r}') = \sum_{i\alpha, j\beta} \phi_{i\alpha}(\mathbf{r}) K_{i\alpha, j\beta} \phi_{j\beta}(\mathbf{r}'). \quad (2)$$

Here, the support functions $\phi_{i\alpha}(\mathbf{r})$ are functions that are non-zero only inside “support regions” centered on the atoms, where i labels the atom and α runs over the support functions on a given atom; the coefficients $K_{i\alpha,j\beta}$ are the matrix elements of the density matrix in the non-orthogonal “basis” of support functions. The support functions $\phi_{i\alpha}(\mathbf{r})$ themselves are expressed as linear combinations of localized basis functions associated with each atom i . CONQUEST provides two types of basis functions, one being B-splines on regular grids [6] as a systematically improvable basis set and the other being numerical pseudo-atomic orbitals (PAOs) [7, 8] for efficient calculations.

In order to calculate the density matrix, we use the density matrix minimization (DMM) method proposed by Li, Nunes and Vanderbilt (LNV) [9]. CONQUEST can also employ conventional diagonalization method, but the CPU time is proportional to N^3 in this case. In the LNV method, we optimize ρ to minimize the total energy with the constraints that ρ is weakly idempotent (all occupation numbers f_n lie between 0 and 1) and ρ gives the correct total valence electron number. In order to satisfy the constraint of idempotency, K is expressed in terms of an auxiliary density matrix $L_{i\alpha,j\beta}$ by the matrix relation:

$$K = 3LSL - 2LSLSL, \quad (3)$$

with $S_{i\alpha,j\beta} \equiv \langle \phi_{i\alpha} | \phi_{j\beta} \rangle$ the overlap matrix of support functions. To use the locality of the density matrix, LNV method imposes a spatial cut-off R_L on the L -matrix: $L_{i\alpha,j\beta} = 0$ for $|\mathbf{R}_i - \mathbf{R}_j| > R_L$, where \mathbf{R}_i are the atomic positions. The method is variational; the increase of R_L results in decrease of the total energy, and an infinite R_L should give us the exact result. Thus, once support function is given, the accuracy of the method can be controlled only by one parameter R_L . The decaying behavior of the density matrix depends on the energy gap of the system, and the method is more efficient for systems having large energy gaps.

3. Parallel Efficiency of CONQUEST on K-computer

As explained in the last section, CONQUEST uses the locality of the electronic structure and thus has an advantage also in parallelization efficiency. In our previous paper [10], we have shown that CONQUEST is efficient on parallel computers, but the number of cores of K-computer is much larger than those used in our previous works. It is not clear whether the code can still show good performance on K-computer. Through the program of the trial use of K-computer, we have done the optimization of CONQUEST from April in 2011. We have done OpenMP parallelization, optimization of computation and communication parts in the sparse matrix multiplications, which is the most expensive part in CONQUEST. As a result, we have obtained a parallel scaling efficiency on K-computer shown in Fig. 1.

In the parallel sparse matrix multiplications, CONQUEST is designed to overlap some of the communication and computation parts by using a series of non-blocking MPI sends and receives. The details of the scheme is explained in our previous paper [11]. On K-computer, we have found that the time for the communication part can be reduced simply by calling MPI_test regularly. This is probably because calling MPI_test issues the controlling signals for MPI communications and invokes the actual MPI sends and receives. Figure 1 (a) shows the timing of the sparse matrix multiplications in the calculations of the 32768-atom bulk

silicon system using different number of CPUs. Here, we use the double zeta polarization function (DZP) basis set as support functions, and employ non-selfconsistent calculations using the Harris-Foulkes energy functional with the local density approximation. The range R_L is 14 bohr¹, and the cutoff energy for the charge density grid is 36 Ry. The timing of the computation and communication parts is shown, with the total elapsed time spent for the matrix multiplications. The figure shows the data for two cases, with and without calling MPI_test. We can see that the timing of the communication part is reduced by calling MPI_test, and the reduction of the time obtained in the communication part is also observed in the total time.

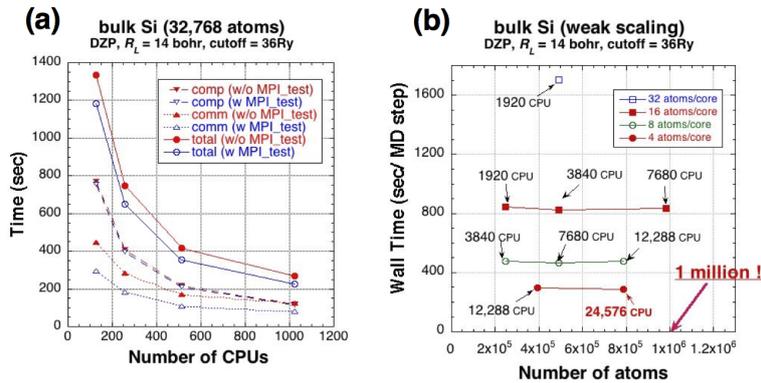


Figure 1: (a) Timing spent for parallel matrix multiplications in the calculations of bulk Si system having 32,768 atoms. Timing of two subroutines mainly for computation part and that of a subroutine for communication part are plotted for two cases with or without calling MPI_test. (b) Timing spent for optimizing density matrix for bulk Si systems having various number of atoms, using different number of CPUs.

Figure 1(b) shows the time spent in the calculations of bulk silicon systems having various number of atoms. These calculations use the same calculation conditions explained above. With the present calculation condition used in the density matrix minimization, we can consider the measured time as an approximate time needed for each step in structure relaxation or molecular dynamics. The lines in the figure show the measured time when the number of atoms per core is fixed. We can see that the weak scaling, shown by each line, is ideal even when the number of CPUs is about 25,000 and the number of cores is about 200,000. On the other hand, we can extract a strong scaling behavior by comparing the timing of different lines. The parallel efficiency in the strong scaling is also good at least if the number of atoms per core is larger than 4. (The efficiency evaluated by decreasing the number of atoms per core from 8 to 4 is 83 % of the ideal performance.) It should be noted that these data were measured in the trial use of K computer and they may have been improved or will be improved in the future. But, the present data are already encouraging. The wall time for 786,432-atom system using 24,576 CPUs (4 atoms/core) is 289.0 sec. Considering

¹We expect that the parallel efficiency would be probably better if the cutoff radius of L matrix (R_L) is larger than 14 bohr, but the actual efficiency may strongly depend on each platform used. For more details of the parallel implementation of the code, see our previous papers [4, 5, 11, 12, 13].

that the weak-scaling is almost ideal, we expect that one step of MD or structure relaxation in the DFT calculation on a million-atom system is shorter than five minutes if the number of CPUs is larger than about 32,000 CPUs. It means that the simulation time of MD can reach almost 10 pico seconds if we run the calculation for fifteen days. Accurate and efficient calculations of MD simulations with the $O(N)$ -DFT method is a challenging task and we will report our scheme with CONQUEST in our forthcoming paper [14].

4. Example: Linear-scaling DFT study of the nucleation of Ge dimers on three-dimensional Ge nano-islands on Si(001)

Using the high parallel efficiency of CONQUEST on the K-computer, we have been recently conducting a large-scale DFT study on the growth of three-dimensional(3D) nanostructures on a semiconductor surface. Here, we show that such large nanostructures can be studied by using a linear-scaling DFT technique and demonstrate the performance of the structure relaxation employed in this study.

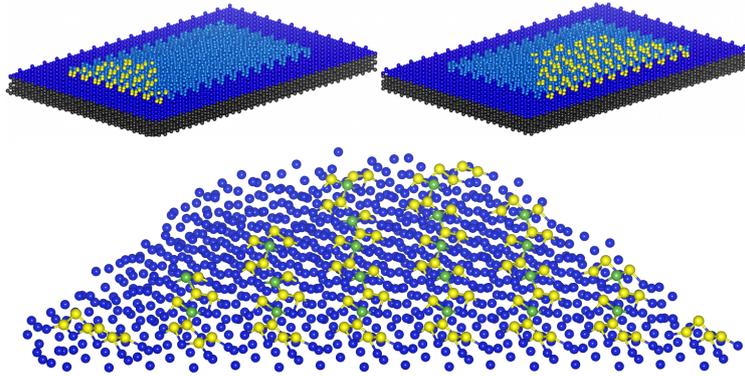


Figure 2: (Top) The model of a 16×26 Ge hut (light blue) on a 24×36 SiGe (001) film (blue and grey) with additional {105} facets (yellow). (Bottom) Additional Ge dimers on the "small" {105} facet: "unstable" dimers have one Ge atom (green) with an extra dangling bond.

The system we have been working is the Ge 3D structures, called 'hut clusters', which are formed when Ge is deposited on the Si (001) surface [15]. The Ge/Si (001) system has been extensively studied, because it is a prototypical example of hetero-epitaxial Stranski-Krastanov growth [16, 17]. When Ge atoms are deposited on Si (001), growth initially occurs layer by layer, up to a critical thickness of about three mono-layers (ML). Strain due to the lattice mismatch is relieved by the formation of regularly spaced rows of dimer vacancies in the two-dimensional (2D) structure, resulting in the $2 \times N$ structure. Deposition of further Ge leads to another strain-relief structure, 3D island structures made by four {105} facets, called 'hut cluster'. The electronic structure of strained Ge (105) surface states has a wide energy gap [18], and thus is suitable for $O(N)$ DFT study. In our previous paper [19], we confirmed that the surface energy of Ge (105) converges quickly with respect to the

cutoff range R_L . Furthermore, we also showed that a cross-over of the total energy from 2D to 3D structures occurs when the coverage of Ge is close to three, being consistent with the experimental observations [20].

To utilize the novel properties of the nano-structured materials, it is often necessary to control their structures at the atomic scale. For the Ge/Si systems, there are many experimental studies aiming to control the structures by changing the growth conditions [21, 22]. There is a report that using a growth condition where the Ostwald ripening can be suppressed, hut clusters are stable and the density of hut clusters is almost constant [22]. In this case, hut clusters, when they grow, become larger only by increasing the length of the longer side while keeping the width constant. Discovering such a specific growth condition is encouraging for controlling the morphology of the nanostructures in the future, but unfortunately it is unclear how hut clusters would grow in this condition because there are no observations of imperfect Ge (105) facets. (It is proposed that the growth speed observed in experiments can be reproduced using a theoretical model assuming that an additional layer of (105) facet is formed from the apex of the smaller facet of hut cluster.)

In order to obtain the detailed atomic-scale information how the 3D nano-structure would grow, we first need to know how the stability of the adsorbed Ge atoms or dimers would vary depending on the adsorbed positions. Hence, we have first calculated an energy map of single Ge dimers adsorbed on two different facets of an elongated hut, in order to identify stable adsorption sites. The system we have been mainly working on is shown in Fig. 2. The typical number of atoms in this system is about 20,000. The dimers shown in yellow are the candidates of the positions for a single Ge dimer adsorbed on the facet. Although the size of the hut cluster in the figure is still a little smaller than typical hut clusters, we still need to calculate more than 100 cases to study the position dependent energies for the adsorbed single Ge dimer. In addition, structure relaxations by DFT are essential in this study since the system includes various irregular sites, like ridges or boundaries between the facets, and there are no experimental information for the atomic positions at these sites. It should be also noted that the adsorption energy strongly depends on the relaxation of the atoms near the adsorbed Ge dimer. Such theoretical studies had been impossible, but they are not difficult to carry out now by using CONQUEST on K-computer.

For the structure relaxation, we mainly used the Fast Inertia Relaxation Engine (FIRE) method [23]. Within this method one performs a quenched MD run with variable time step, as well as other adjustable parameters. Employing a quenched MD relaxation became efficient due to recent implementations that identify the transformation between the two lists of neighbor atoms before and after the movement of atoms. As explained before, we work with the sparse matrix $L_{i\alpha,j\beta}$ in the $O(N)$ method. When the atoms move during the structure relaxation, we need to update the list of the pairs of neighbor atoms $\{i, j\}$ whose distance is within the cutoff range R_L . The transformation between lists of current and previous atom positions allows to use the optimized density matrix in the last step of a MD run as an initial guess for the density matrix at the current MD step. This scheme can reduce the number of steps largely in the DMM.

In the structure optimization, as we need to relax a large number of atoms, we have adopted an approach where we relax groups of atoms step by step; we first relax the atoms having large forces with the initial positions, then increase the number of target atoms whose

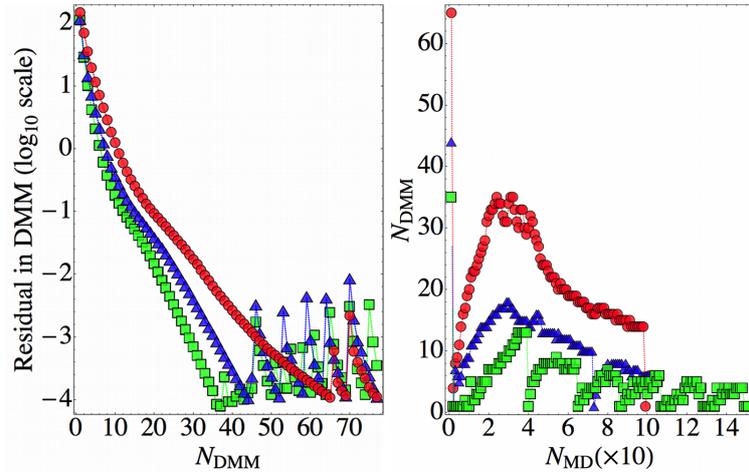


Figure 3: Change in the residual of the density matrix during first steps of the quenched MD (left), as well as the number of DMM iterations during the structure optimization run (right) at different stages of structure relaxation: circle - relaxation of surface atoms only, triangle - relaxation of all Ge atoms, square - relaxation of the entire system. For the same number of DMM iterations 3, 6 and 18 MD steps are accomplished for each relaxation stage, respectively.

positions will be optimized. In the optimization of hut clusters without adsorbed Ge dimers, we performed calculations in three steps: i) optimizing first only surface atoms (atoms of the {105} facets and their nearest neighbors, and surface atoms of the Ge wetting layer) and keeping the rest of the system frozen; ii) relaxing all Ge atoms and keeping the Si substrate fixed; iii) allowing Si atoms to relax too. In the calculation of a single Ge dimer adsorbed on a hut cluster, we have done two-step relaxation using the optimized atomic positions of a hut cluster; i) optimization of the dimer, and its neighbor atoms within the four hops (Ge-Ge bonds) from the dimer atoms, then ii) relaxation of the all atoms except the three bottom layers of Si substrate.

The change of the residual in the DMM during different stages of the structure relaxation of a hut cluster is shown in Fig. 3. The graph on the left-hand side of the figure shows the change in the density matrix residual during several MD steps for different stages of structure relaxation. Number of iterations required to minimize the density matrix at starting configuration are 65, 44, and, 35 for each stage of structure optimization respectively. Using density matrix minimized at the previous step as starting point for the next one reduces the number of DMM iterations to 4, 1, and 1 in the second MD step, and to 8, 7, and 1 in the third one for three different quenched MD runs. The graph on the right-hand side of Fig. 3 shows the number of iterations to achieve the given tolerance, 10^{-4} , of the density matrix residual during structure optimization. In the beginning of each MD run, when the time step and its change are small, the number of DMM iterations is considerably reduced. However, as the MD simulation proceeds, the velocity increases and two consecutive configurations may differ at greater extent, the number of iterations to minimize the density matrix will

increase, but still is smaller by, at least, a factor of two even in the worst case. As the system gets closer to its relaxed configuration, the number of DMM decreases again, so the new implementation considerably speeds up the MD run.

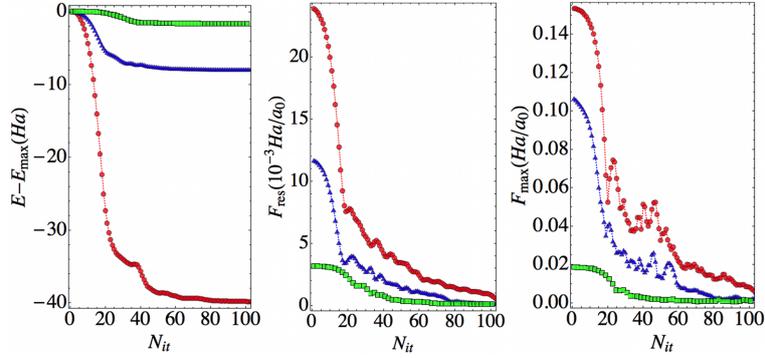


Figure 4: Evolution of total energy (left), force residual (middle), and maximum force component (right), during different stages of optimization: circle - relaxation of surface atoms only, triangle - relaxation of all Ge atoms, square - relaxation of the entire system.

Figure 4 shows the changes in total energy E , force residual (average force) F_{res} , and maximum force component on an atom F_{max} at different stages of relaxation. The FIRE method employs an efficient structure optimization by increasing the time step when the scalar product P of velocity and force is positive (downhill direction), and by decreasing the time step with zeroing the velocity when P is negative. The method also has a control parameter Δt^{max} , maximum of the time step, which prevents the time step being too large. We have confirmed that the FIRE optimization is quite efficient at the first and second stages. However, at the third stage, we sometimes observed a very slow reduction or even the increase of F_{res} , even when the scalar product P is positive. In such cases, we have found that Δt^{max} needs to be very small. But, since this small Δt^{max} would cause a slowdown of the optimization even when the method works, we have introduced another resetting condition; if P keeps increasing and F_{res} does not decrease for several steps, we decrease the time step and set the velocity zero. We have observed that this resetting would stabilize the structure optimization and would decrease the number of MD steps to reach the fully relaxed structure. The green lines in Fig. 4 show the results with this resetting. We also note here that a periodical decrease of N_{DMM} in Fig. 3 for the third stage of relaxation is due to this new condition of resetting. This resetting also helps to reduce the number of total DMM steps.

We have shown that even in such large and complex systems, we can perform robust and accurate structure relaxation with CONQUEST using the $O(N)$ DFT technique. As a result, we have succeeded to obtain an energy map of single Ge dimers, showing that ridges and edges between two facets are stable and that the upper positions are usually more stable than lower positions by 0.2eV/atom approximately. We have been also investigating the completion of $\{105\}$ facets by calculating the stability of "imperfect" facets. The results together with the energy map of single Ge dimers on the facets will be reported in our forthcoming paper [24]. We also point out that we have done structure relaxations of bigger systems, having more

than 100,000 atoms in total.

5. Conclusion

In this paper, we have demonstrated that our linear-scaling DFT code CONQUEST has good parallel efficiency on K-computer. It was shown that parallel scaling is almost ideal even when about 200,000 cores are used, and the time needed to do one MD or one structure relaxation step is small enough to realize the actual DFT study on very large, for example million-atom systems. For the DFT study on the adsorption of Ge dimers on Ge hut cluster on Si(001) substrate, we show that the accurate, efficient and robust structure relaxation in the actual scientific research on complex systems is now possible.

Acknowledgement

The authors thank the useful discussions with Lianheng Tong, Mike Gillan, Takao Otsuka and Ayako Nakata. Most of the calculations in this work were done using the K computer, and the support from the AICS-RIKEN to optimize the CONQUEST code on K computer is gratefully acknowledged. This work was partly supported by the Strategic Programs for Innovative Research (SPIRE), the Computational Materials Science Initiative (CMSI) and Grant-in-Aid for Scientific Research on Innovative Areas (No. 22104005) by MEXT, Japan.

References

- [1] TOP500 Supercomputer Sites: <http://www.top500.org>, 24, February, 2014.
- [2] D. R. Bowler and T. Miyazaki: $O(N)$ methods in electronic structure calculations, *Reports on Progress in Physics*, 75 (2012), 036503.
- [3] E. Hernández, M. J. Gillan, and C. M. Goringe: Linear-scaling density-functional technique : The density matrix approach, *Phys. Rev. B*, 53 (1996), 7147–7157.
- [4] D. R. Bowler, T. Miyazaki, and M. J. Gillan: Recent progress in linear scaling *ab initio* electronic structure techniques, *J. Phys.:Condens. Matter*, 14 (2002), 2781–2798.
- [5] D. R. Bowler, R. Choudhury, M. J. Gillan, and T. Miyazaki: Recent progress with large-scale *ab initio* calculations: the CONQUEST code, *Phys. Stat. Sol. (b)*, 243 (2006), 989–1000.
- [6] E. Hernández, M. J. Gillan, and C. M. Goringe: Basis functions for linear-scaling first-principles calculations, *Phys. Rev. B*, 55 (1997), 13485–13493.
- [7] J. M. Soler, E. Artacho, J. D. Gale, A. García, J. Junquera, P. Ordejón, and D. Sánchez-Portal: The SIESTA method for *ab initio* order-N materials simulation, *J. Phys.:Condens. Matter*, 14 (2002), 2475–2779.

- [8] A. S. Torralba, M. Todorović, V. Brázdová, R. Choudhury, T. Miyazaki, M. J. Gillan, and D. R. Bowler: Pseudo-atomic orbitals as basis sets for the O(N) DFT code CONQUEST, *J. Phys.:Condens. Matter*, 20 (2008), 294206.
- [9] X.-P. Li, R. W. Nunes, and D. Vanderbilt: Density-matrix electronic-structure method with linear system-size scaling, *Phys. Rev. B*, 47 (1993), 10891–10894.
- [10] D. R. Bowler and T. Miyazaki: Calculations on millions of atoms with density functional theory: linear scaling shows its potential, *J. Phys.: Condens. Matter*, 22 (2010), 074207.
- [11] D. R. Bowler, T. Miyazaki, and M. J. Gillan: Parallel sparse matrix multiplication for linear scaling electronic structure calculations, *Comp. Phys. Commun.*, 137 (2001), 255–273.
- [12] C. M. Goringe, E. Hernandez, M. J. Gillan, and I. J. Bush: Linear-scaling DFT-pseudopotential calculations on parallel computers, *Comp. Phys. Commun.*, 102 (1997), 1–16.
- [13] V. Brázdová and D. R. Bowler: Automatic data distribution and load balancing with space-filling curves: implementation in CONQUEST, *J. Phys.:Condens. Matter*, 20 (2008), 275223.
- [14] M. Arita, D. R. Bowler, and T. Miyazaki: Stable and Efficient Linear Scaling First-Principles Molecular Dynamics for 10,000+ atoms, *J. Chem. Theory Comput.*, submitted.
- [15] Y. W. Mo, D. E. Savage, B. G. Swartzentruber, and M. G. Lagally: Kinetic pathway in stranski-krastanov growth of ge on si(001), *Phys. Rev. Lett.*, 65 (1990), 1020–1023.
- [16] J. N. Aqua, I. Berbezier, L. Favre, T. Frisch, and A. Ronda: Growth and self-organization of SiGe nanostructures, *Physics Reports*, 522 (2013), 59–189.
- [17] J. Stangl, V. Holý, and G. Bauer: Structural properties of self-organized semiconductor nanostructures, *Reviews of Modern Physics*, 76 (2004), 725–783.
- [18] T. Hashimoto, Y. Morikawa, and K. Terakura: Stability and electroic structure of Ge (105) 1x2: a first-principles theoretical study, *Surface Science*, 576 (2005), 61–66.
- [19] T. Miyazaki, D. R. Bowler, R. Choudhury, and M. J. Gillan: Density functional calculations for Ge(105): Local basis sets and O(N) methods, *Phys. Rev. B*, 76 (2007), 115327.
- [20] T. Miyazaki, D. R. Bowler, M. J. Gillan, and T. Ohno: The energetics of hut-cluster self-assembly in Ge/Si(001) from linear-scaling DFT calculations, *J. Phys. Soc. Jpn.*, 77 (2008), 123706.
- [21] M. Kästner and B. Voigtländer: Kinetically self-limiting growth of Ge islands on Si(001), *Physical Review Letters*, 82 (1999), 2745–2748.

- [22] M. R. McKay, J. A. Venables, and J. Drucker: Kinetically suppressed ostwald ripening of Ge/Si hut clusters, *Physical Review Letters*, 101 (2008), 216104.
- [23] E. Bitzek, P. Koskinen, F. Gähler, M. Moseler, and P. Gumbsch: Structural relaxation made simple, *Physical Review Letters*, 97 (2006), 170201.
- [24] S. Arapan, D. R. Bowler and T. Miyazaki: A linear scaling DFT study of the growth of a new 105 facet layer on a Ge hut cluster, in preparation.