Computation of Strain Distributions in Quantum Dot Nanostructures by Means of Atomistic Simulations

Franz Gähler, Christopher Kohler, Johannes Roth, and Hans-Rainer Trebin

Institut für Theoretische und Angewandte Physik, Universität Stuttgart,
D-70550 Stuttgart, Germany

Abstract. Strain distributions around Ge quantum dots embedded in a Si matrix are computed by means of classical molecular dynamics simulations using the molecular dynamics code IMD. The Tersoff potential is employed in order to model covalent bonds. Two crystal lattice structures are considered, the cubic and hexagonal diamond structure. The distributions of the planar strain are studied for a large number of system sizes and lattice misorientations. In a second part, the scaling of IMD is analyzed for different parallelization schemes and machine architectures.

1 Introduction

Quantum dots (QDs) in semiconductors are of great technological importance because of their peculiar optical and electronic properties. In the growth of several layers of planar arrays of QDs, a vertical correlation is observed which leads to regular three dimensional superstructures of QDs. Several forms of such vertical self-orderings have been observed: Perfect vertical alignment [1], vertical anticorrelation with different forms of stackings [2, 3], and vertical correlation with lateral shift for the case of misoriented samples [4].

The reason for these vertical correlations of QDs is usually attributed to the strain arising from the lattice mismatch between the material of the QDs and the embedding material. The tensile strain in the vicinity of the QDs favours the nucleation of QDs in succesive layers at positions where the strain is enlarged. In order to control the growth of QD superlattices, the knowledge of the strain distributions around the QDs is thus of importance. Several theoretical methods have been applied so far to study the correlation of QDs driven by the strain: Continuum elasticity theory [5–7], the finite element method [8, 9], and classical molecular dynamics (MD) simulations [5, 10–15]. MD simulations have the advantage of providing the strain distributions on the atomic level and are the most realistic methods. Quantum mechanical simulations are not feasible in this case due to the large dimensions of the systems amounting to at least several nanometers.

In this article, we report on large-scale MD simulations of strain distributions in QD structures. Similar simulations have been performed previously focusing on different aspects of the strain and stress due to the QDs [5,10–15].

The present article deals particularly with the influence of the lattice orientation as well as of the elastic anisotropy on the strain distributions. In order to gain insight into the dependence on the anisotropy, we have performed simulations of Ge/Si QDs with two different crystal structures: cubic diamond structure and hexagonal diamond structure. These structures can be considered as model systems for many semiconductors.

Most of the results presented in this article have been obtained on the CRAY-T3E. There is, however, a trend to new hardware architectures, like large clusters with cheap commodity hardware, or clusters with large multiprocessor nodes. Moreover, on machines with shared memory multiprocessor nodes new parallelization options become available. For this reason, we have undertaken a comparison of the performance and scaling of our MD code IMD for several different machine architectures and parallelization schemes. These results are expected to be relevant also for other MD codes using short range interactions.

This article is organized as follows. In section 2, we describe the methods of the simulations. The results of the simulations are presented in section 3. In section 4, we present an analysis of the parallel performance of our MD program IMD for different parallelization schemes and machine architectures. Finally, in section 5, we summarize our work.

2 Simulation Methods

Typical semiconductors used in the fabrication of QDs are group IV elements (e.g. Si, Ge) or compounds of group III-V elements (e.g. GaAs, InGaN, InP). These are predominantly covalent materials. For their modeling in classical MD simulations, many-body potentials have to be used which favour specific bond angles. In the present simulations, the Tersoff potential has been employed. This potential depends on a number of parameters which have to be determined by fitting to experimental data. Parameterizations for several semiconductors have been given in the literature. In the present work, we have performed simulations of Ge QDs in a Si matrix. The corresponding parameters are taken from Reference [16].

We have used two different crystal lattice structures which are representative for many semiconductors: The cubic diamond structure and the hexagonal diamond structure. The former has a cubic elastic anisotropy similar to that of the zincblende structure while the latter is similar to the wurtzite structure with a hexagonal elastic anisotropy.

For the simulations, we have employed a starting configuration that is analogous to the experimental setup in the epitaxial growth of QDs. The simulation box consists of a substrate of Si atoms above which a thin wetting layer of Ge atoms is placed. The QD of Ge atoms has the form of a truncated square pyramid and is situated above the wetting layer. A spacer layer (SL) of Si atoms covers the wetting layer and the QD. A part of this configuration can

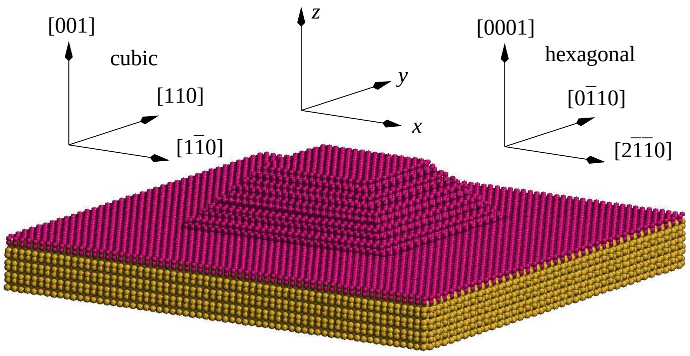


Fig. 1. Starting configuration of a QD simulation. The Si atoms (yellow) belong to the substrate. The wetting layer and the QD consist of Ge atoms (red). The capping spacer layer of Si atoms is not shown in this figure.

be seen in Figure 1 where the upper part of the substrate and the uncovered QD is shown. At the sides of the simulation box, periodic boundary conditions are imposed which means that the system effectively consists of an infinite two dimensional horizontal superlattice of QDs. At the bottom of the simulation box, some layers of the substrate are held fixed during the simulation while at the top of the SL, free boundary condition are used. We have performed simulations with different sizes of the simulation box varying the width from 18 nm to 50 nm and the height from 14 nm to 40 nm. The simulation box for the largest system contained about four milion atoms. The QD always has a fixed size of 12 nm width and 3 nm height. We have also used simulation boxes with different misorientations of the lattices. The crystallographic directions indicated in Figure 1 for the cubic and hexagonal diamond lattice correspond to the systems without misorientation. In the case of misoriented systems, the lattice is tilted about the [110] and the [$2\overline{1}\overline{1}0$] directions for the cubic and hexagonal structures, respectively.

The relaxed states of the QD systems are obtained by MD simulations where during the dynamic evolution of the systems in the NVE ensemble, the velocities of all particles are set to zero whenever the velocity in the configuration space of the system is in the direction opposite to the force vector. The MD program package IMD [17] has been used for these simulations. Depending on the system size, an MD run took 4-6 hours on the CRAY-T3E where usually 128 processors have been used.

From the displacements of the atoms of the relaxed configurations relative to the starting configuration , the strain tensor has been computed using a utility program that is also part of IMD. In the following we discuss only the planar strain $\epsilon_{xx} + \epsilon_{yy}$ since this is the most important part of the strain for the explanation of the vertical correlations in QD systems.

3 Results

3.1 Quantum Dots in the Cubic Diamond Lattice

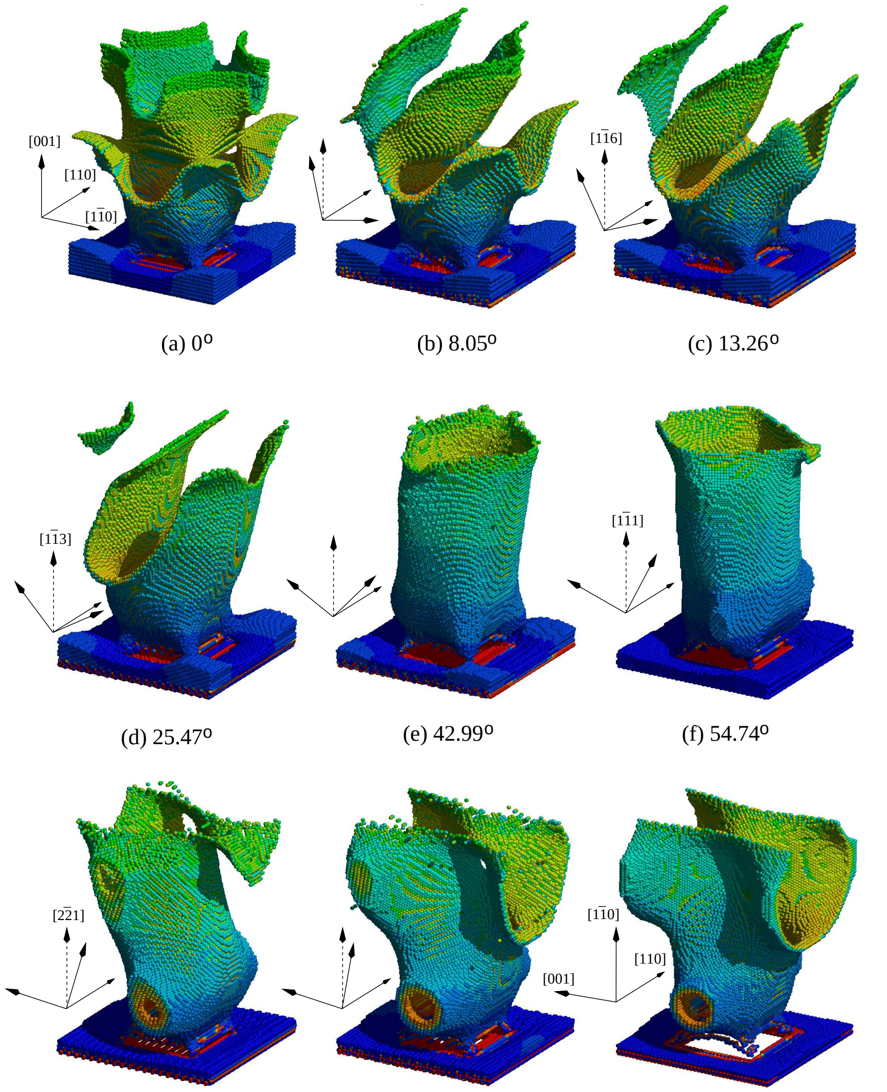
Due to the elastic anisotropy of the matrix material, the strain distributions of the relaxed QD systems display a complex behaviour with regions of positive and negative strain. In order to gain a qualitative insight into the distributions of planar strain, it is convenient to visualize the surfaces of vanishing planar strain which form the boundaries of the regions of different signs of strain. Figure 2 shows these surfaces for QD systems with diamond lattice for angles θ of misorientation in the range $0° \leq \theta \leq 90°$. The system widths are about 23 nm and the SL thicknesses in all cases are 28 nm. Simulations of systems with smaller SL thicknesses have shown that the strain distributions can be obtained by truncating the system of 28 nm SL thickness (apart from small boundary effects at the SL surface). Thus, from the shape of the surfaces in Figure 2, the sign of the planar strain on the SL surfaces of systems of different height can be read off.

Generally, the planar strain has a strong decay and a characteristic change of sign as a function of increasing distance from the QD. In the case of the misorientation angle 0°, Figure 2 (a) shows that the strain is oscillating with the distance frome the QD. Directly above the QD (in between the curved surfaces) the strain is tensile. Above the upper surface and below the lower surface, the strain is compressive. For misorientation angles greater than zero, the two surfaces merge to a tube that is tilted in the direction opposite to the misorientation (Figures 2 (b)-(d)). For misorientation angles $\approx 50°$, the positive planar strain is confined within a nearly vertical tube. In these cases, there is no oscillation of the strain (Figures 2 (e),(f)). Increasing the misorientation angle further, the tube is bent in the direction of the misorientation where at different heights of the tube, necks between the tubes of different QDs in the QD superlattice form (Figures 2 (g)-(i)).

Using simulation boxes of different heights, that is, different SL thicknesses, the correlation behaviour of QDs in several vertical layers can be predicted by determining the positions of maximal planar strain on the top of the SL. Depending on the misorientation angle and the SL thickness, we thus obtain a variety of possible correlations which partly have been predicted by continuum elasticity calculations and also found experimentally.

The oscillatory behaviour of the strain above the QDs is caused by the finite horizontal extent of the simulation box. For the case of a QD system without misorientation, the profile of the planar strain along a vertical line through the middle of the QD is shown in Figure 3. In order to characterize this profile quantitatively, we have fitted the following function to the profile:

$$\epsilon(z) = \epsilon_0 \frac{\cos\left[2\pi\left(\frac{z-z_0}{p}\right)\right]}{1+\left(\frac{z-z_0}{h}\right)^{\alpha}}. \tag{1}$$



(a) 0°　(b) 8.05°　(c) 13.26°

(d) 25.47°　(e) 42.99°　(f) 54.74°

(g) 70.53°　(h) 82.33°　(i) 90°

Fig. 2. Surfaces of vanishing planar strain in a Ge/Si quantum dot system with cubic diamond lattice for different misorientation angles. The size of the simulation box is 23 nm × 23 nm × 35 nm. The tilt axis is the [110] direction and the growth direction is indicated by a dashed arrow.

This function describes an oscillation of the strain and a decay by a power law. The exponent α characterizes the strength of the decay of the strain and p is the period of the oscillation. The paramter h corresponds to the height

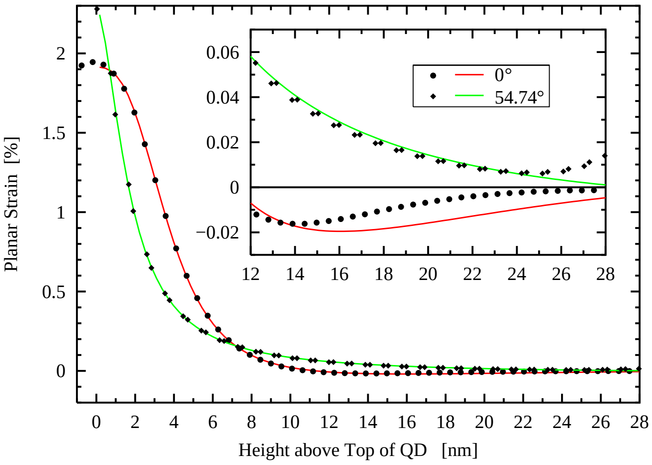


Fig. 3. Planar strain along a vertical line through the middle of a QD in systems with misorientation angles 0° and 54.74°. The size of the systems is 23 nm × 23 nm × 40 nm with SL thicknesses of 28 nm. The dots correspond to the simulated strain and the curves correspond to the fitting function (1). The inset shows the enlarged profiles for larger heights above the dot.

where the strain has fallen off to half its maximal value ϵ_0 at the height z_0 which corresponds to the height of the top of the QD. For the QD system the profile of which is shown in Figure 3, the fitted values of the parameters are $\alpha = 2.9$, $p = 45$ nm, $h = 4$ nm, and $\epsilon_0 = 1.9$ %. The fitted curve is also plotted in Figure 3. It can be seen that the simple function (1) provides a good description of the strain profile. The deviations for larger heights can partly be attributed to the boundary effects of the free surface.

In order to gain insight into the dependence of the parameters α, p, h, and ϵ_0 on the lattice constant of the QD superlattice, we have performed simulations with different widths of the simulation boxes from 18 nm to 50 nm. We have found that the period p of the oscillation is nearly proportional to the superlattice constant while α, h, and ϵ_0 are unchanged for a wide range of superlattice constants.

3.2 Quantum Dots in the Hexagonal Diamond Lattice

The surfaces of vanishing planar strain of QD systems with hexagonal diamond structure for several misorientation angles θ between 0° and 90° are shown in Figure 4. It can be seen that the strain distributions are different from the ones of the QD systems with cubic diamond lattice given in Figure 2. In the case of misorientation angle 0° (Figure 4 (a)), the tensile planar strain above the QD is confined to a narrow tube. There is no oscillation of

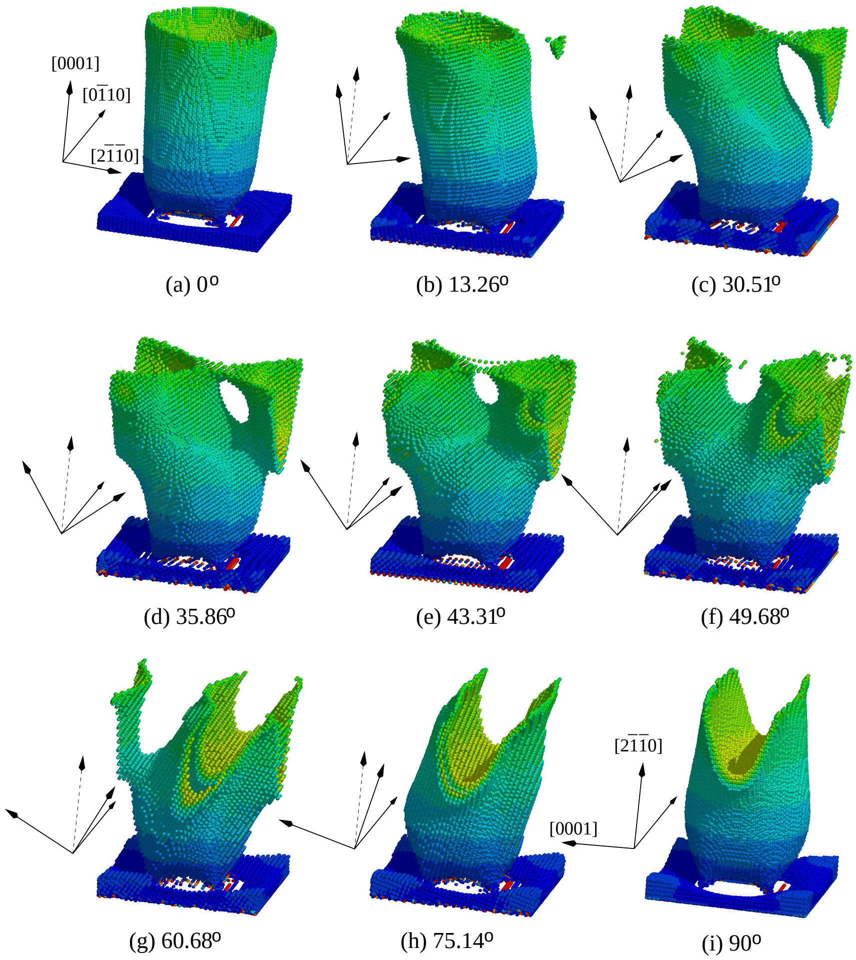


Fig. 4. Surfaces of vanishing planar strain in a Ge/Si quantum dot system with hexagonal diamond lattice for different misorientation angles. The size of the simulation box is 25 nm × 25 nm × 40 nm. The tilt axis is the [0$\bar{1}$10] direction and the growth direction is indicated by a dashed arrow.

the strain along the vertical direction. For finite misorientation angles, the tube is bent in the direction of the misorientation (Figures 4 (b), (c)). For larger misorientations, the tubes of neighbouring QDs of the QD superlattice merge. For intermediate misorientation angles, the tubes of QDs in the direction perpendicular to the tilt axis are connected (Figures 4 (d)-(f)) whereas for large angles the tubes in the direction of the tilt axis are connected (Figures 4 (g)-(i)). The double sheets of surfaces in Figures 4 (e)-(g) are due to internal relaxations of the hexagonal diamond lattice which cause an internal displacement of sublattices. As in the case of the cubic diamond lattice, we expect a variety of possible vertical correlations of QDs as a function of the misorientation angle.

Comparing Figures 2 and 4, it can be seen that there is a certain similarity of the shapes of the surfaces for the angular range $54.74° < \theta < 90°$ (Figures 2 (f)-(i)) in the cubic diamond case and for the angular range $0° < \theta < 43.31°$ (Figures 4 (a)-(e)) in the hexagonal diamond case. This can be explained by the similarity of the elastic properties of the cubic diamond lattice along the <111> directions and the hexagonal diamond lattice along the [0001] direction. The reason for this is that the orientation of the coordination tetrahedra in both structures are similar along the respective lattice directions. The vertical decay of the strain is also very similar in the two types of lattices for the considered ranges of misorientations. We have fitted the function (1) to the profiles of the planar strain along a vertical line through the middle of the QDs for the systems in Figures 2 (f) and 4 (a). In the former case we obtain $\alpha = 1.7$ and $h = 1.7$ nm while in the latter case $\alpha = 1.6$ and $h = 1.8$ nm. For the system with cubic diamond lattice, the profile of the planar strain along with the fitted curve is included in Figure 3. Since in both systems, the planar strain on the SL surface is strongly concentrated above the position of the QD, a perfect vertical correlation can be expected.

4 Performance and Scaling of IMD

The performance and scaling of IMD had last been analyzed in [17]. Since then, IMD has evolved considerably. Large parts of the communication routines have been rewritten, and new parallelization methods have been implemented. Besides the standard Message Passing Interface (MPI), which is still the main parallelization workhorse, it is now also possible to use OpenMP [18], either instead of MPI or in combination with MPI. OpenMP is a standard set of compiler directives, which instruct the compiler to set up several threads to share the work. OpenMP is suitable for multiprocessor machines with a *shared* memory. A combination of OpenMP and MPI can be used on clusters of several such multiprocessor nodes; inside the nodes the parallelization is via OpenMP, whereas the inter-node communication is via MPI. Of course, in all these cases it is also possible to use just plain MPI, also for intra-node communication. It is a priori not clear which parallelization scheme is more efficient. OpenMP might have an advantage in that the work sharing is dynamic inside a node, whereas in a pure MPI scheme it is static. This might alleviate load balancing problems, which are hard to resolve in a pure MPI scheme (IMD uses a geometric domain decomposition). Also, with OpenMP it is not necessary to make copies of certain data; the same data can be read by all threads in a node. On the other hand, if OpenMP is combined with MPI, all the communication of a node is done by a single thread, whereas with MPI every process does its own communication. This might well turn out to be a bottleneck.

There is currently a trend towards building massively parallel clusters with cheap commodity hardware, typically running under Linux. A crucial point for such machines is the communication network between the cluster nodes. It would be interesting to know how an MD code like IMD performs with different types of networks, such as 100 Mbit Ethernet, or the more expensive, but faster Myrinet, and how this compares to real supercomputers. A high performance network makes up for a considerable fraction of the total system cost, and the question is whether this is a good investment.

The performance tests were run on the T3E at HLRS, the Hitachi SR8000 at HLRS, with both pure MPI communication and with a combination of OpenMP and MPI, the Kepler cluster [19] of the Sonderforschungbereich 382 in Tübingen, and the Volvox cluster at HLRS. Both the Kepler cluster and the Volvox cluster are similarly configured, with dual processor Pentium III nodes of 650 MHz and 1000 MHz, respectively. While the Kepler cluster has more nodes than the Volvox cluster, on the latter we could compare the communication performance for different networks (Ethernet and Myrinet 2000). On Kepler, only Myrinet was available. On the other hand, on Kepler only the gcc compiler was installed, whereas on Volvox we could use the Intel compiler, which produces significantly fast code (almost by a factor of 2).

For the performance tests we used a simple fcc crystal with tabulated pair interactions. We should emphasize, however, that the performance characteristics and scaling behaviour is not significantly different for different types of *short range* interactions. Therefore, the general trends are relevant also for Tersoff interactions, or EAM interactions. We determined two kinds of performance scaling, with samples of a fixed size, and with samples of a size proportional to the number of processors. In the first case, a given problem of fixed size is solved with the help of more and more processors, whereas in the latter case more and more processors are used to solve larger and larger problems. In both cases, two sample sizes were used. One should keep in mind that the communication overhead is proportional to the surface of the block of material a processor has to deal with. Also, the processes have to be synchronized after each time step. The more work per time step a processor has to do, the lower are the synchronization costs. Therefore, for a given sample size there is a maximum number of processors which is feasible. If one goes beyond this number, the parallelization overhead becomes overwhelming.

The results of the performance tests are shown in Figures 5 (samples of constant size) and 6 (samples of constant size per processor). In an ideal world (perfect scaling), all curves would be horizontal. The curves for the T3E come close to this ideal, which shows that the T3E still has an excellent network. Except for the big sample of constant size, the scaling of the Hitachi SR8000 is less perfect. Also, the pure MPI mode on the SR8000 is 10-20% more efficient than the OpenMP+MPI mode. In addition to the reasons discussed above, this may also be due to the still fairly new OpenMP implementation (both in the compiler and in IMD). So far, we could not make comparisons for other machines where both MPI and OpenMP are available.

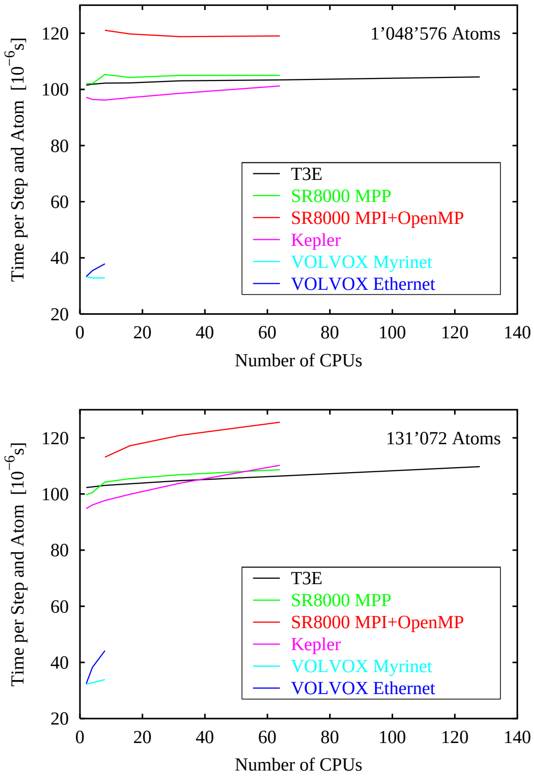


Fig. 5. Performance and scaling of IMD on different machines and architectures for a fixed total number of atoms. These atoms are distributed on a varying number of CPUs. The results for two sample sizes are shown.

So, for the moment a pure MPI scheme seems preferable. This may be different, however, if the atom distribution is very inhomogeneous in the sample. One might think of the simulation of an indentor, which is pushed into the surface of a block of atoms, and which consists itself of atoms whose movements have to be simulated. In such a situation, it is very difficult to divide the system into many blocks of the same geometric size and the same number of atoms. One can easily make such a division into four blocks, but not more. If the different blocks contain largely different numbers of atoms, there will be severe load balancing problems in a pure MPI scheme. If one divides the sample into four blocks only, and uses for each a multiprocessor node with 16 or even 32 processors, one can use a large number of processors without load balancing problems, because with OpenMP the work sharing between the different threads in a node is dynamic. The results for the clusters show that with the faster Myrinet a scaling performance similar to that of the SR8000 can be achieved. With Ethernet, however, the performance is much

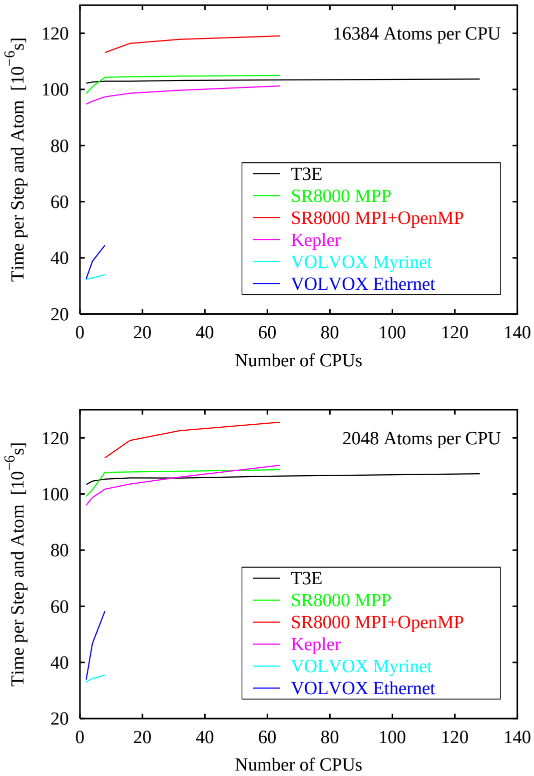


Fig. 6. Performance and scaling of IMD on different machines and architectures for a fixed number of atoms per CPU. The results for two sample sizes are shown.

worse already for moderate CPU numbers. Therefore, communication over Ethernet is clearly not suitable for massively parallel MD simulations, where the different MPI processes have to be synchronized after every time step.

5 Conclusions

In this article, we have reported on large-scale MD simulations of QD structures performed on the CRAY-T3E using the MD program package IMD. We have performed simulations of strain distributions in QD structures for systems of realistic size where the dependence of the strain on the form of the simulation box has been investigated for a larger variety of sizes than have been done so far. The results of these simulations are in agreement with experiments and theoretical strain computations using continuum elasticity theory. Furthermore, we have predicted new vertical correlations in QD superlattices for semiconductor systems of cubic (diamond, zincblende) and hexagonal (wurtzite) lattices.

In a second part, the performance and scaling of IMD has been analyzed for different parallelization schemes and machine architectures. A pure MPI parallelization scheme is usually preferable over a combination of OpenMP+MPI, even though the latter can be more efficient in particular cases with load balancing problems. A massively parallel cluster built from commodity hardware is suitable for IMD (and presumably other MD codes), *provided* it is equipped with a fast network such as Myrinet.

Acknowledgements

This work was supported in part by the Deutsche Forschungsgemeinschaft (DFG) through the Graduiertenkolleg "Grenzflächen in kristallinen Materialien" (GRK 285) and the Sonderforschungsbereich 382 "Verfahren und Algorithmen zur Simulation physikalischer Prozesse auf Höchstleistungsrechnern".

References

1. Q. Xie, A. Madhukar, P. Chen, and N. P. Kobayashi, Phys. Rev. Lett. 75, 2542 (1995)
2. M. Strassburg, V. Kutzer, U. W. Pohl, A. Hoffmann, I. Broser, N. N. Ledentsov, D. Bimberg, A. Rosenauer, U. Fischer, D. Gerthsen, I. L. Krestnikov, M. V. Maximov, P. S. Kop'ev, and Zh. I. Alferov, Appl. Phys. Lett. **72**, 942 (1998)
3. G. Springholz, V. Holý, M. Pinczolits, and G. Bauer, Science 282, 734 (1998)
4. A. Fantini, F. Phillipp, C. Kohler, J. Porsche, and F. Scholz, to appear in J. Crystal Growth
5. C. Pryor, J. Kim, L. W. Wang, A. J. Williamson, and A. Zunger, J. Appl. Phys. **83**, 2548 (1998)
6. V. A. Shchukin, D. Bimberg, V. G. Malyshkin, and N. N. Ledentsov, Phys. Rev. B **57**, 12262 (1998)
7. V. Holý, G. Springholz, M. Pinczolits, and G. Bauer, Phys. Rev. Lett. **83**, 356 (1999)
8. T. Benabbas, P. Francois, Y. Androussi, and A. Levebvre, J. Appl. Phys. **80**, 2763 (1996)
9. T. Benabbas, Y. Androussi, and A. Levebvre, J. Appl. Phys. **86**, 1945 (1999)
10. Y. Kikuchi, H. Sugii, and K. Shintani, J. Appl. Phys. **89**, 1191 (2001)
11. M. A. Makeev and A. Madhukar, Phys. Rev. Lett. **86**, 5542 (2001)
12. I. Daruka, A.-L. Barabási, S. J. Zhou, T. C. Germann, P. S. Lomdahl, and A. R. Bishop, Phys. Rev. B **60**, R2150 (1999)
13. W. Yu and A. Madhukar, Phys. Rev. Lett. **79**, 905 (1997); **79** 4939(E) (1997)
14. K. Scheerschmidt, D. Conrad, H. Kirmse, R. Schneider, and W. Neumann, Ultramicroscopy 81, 289 (2000)
15. C. Kohler (preprint 2002)
16. J. Tersoff, Phys. Rev. B **39**, 5566 (1989)
17. J. Stadler, R. Mikulla, and H.-R. Trebin, Int. J. Mod. Phys. **C8**, 1131 (1997)
18. `http://www.openmp.org`
19. `http://kepler.sfb382-zdv.uni-tuebingen.de/kepler/index.shtml`