Learning more about soft matter

Dr Axel Arnold has developed software that aids the description of soft matter interactions. Here, he discusses algorithmic efficiency, technological advances and some of the applications of his work

In broad terms, have any scientific advances of particular significance been yielded from molecular dynamics (MD) simulations in recent years?

Strictly speaking, scientific advances can only be made through experiments. However, computer simulations have become an important tool in interpreting experiments. For example, the Schulten group at the University of Illinois at Urbana-Champaign (UI-UC) could determine the structure of the HIV capsid through large scale MD simulations based on data from non-magnetic resonance (NMR) spectroscopy and X-ray scattering. MD simulations are often used to test predictions such as the like-charge attraction of DNA or the electrophoretic mobility of colloidal particles.

Can you give an insight into developments around algorithms, methods and high-performance computing (HPC) that have facilitated these experiments? Has your own lab reaped success in developing novel methods or unearthing exciting discoveries to date?

The basic MD algorithms date back to the 1960s. However, only 10 years ago simulation codes including GROMACS, NAMD or ESPResSo became robust enough to be used by non-experts. At the same time, the improved performance of computers has significantly pushed the limits of MD. Nanotechnology and biochemistry have also shifted the focus to smaller systems, which are harder to study experimentally, but easier to simulate.

Our lab is involved in the development of electrostatics algorithms, in particular for systems that are not fully periodic. These algorithms were used, for example, to demonstrate that DNA strands can exhibit electrostatic attraction despite the fact that they are like-charged. This attraction is mediated by the counter-ions necessary to neutralise the DNA.

How can the exploitation of the entire capabilities of the computational architectures that will appear in 8-10 years be ensured?

Future exascale computers will very likely provide most of their computational power in the form of accelerators. At present, NVIDIA graphics processing units (GPUs) are particularly successful, and all major MD codes can make use of them. From my experience this success is due to the availability of easy-to-use compilers, the relatively simple hardware hierarchy, and the important fact that one can develop and test efficiently on a desktop computer. For example, the ESPResSo LB code was developed by a PhD student of mine on his laptop. New accelerators will only be successful if they provide a comparably low entry barrier.

With scientists only recently becoming aware of the need for energy-efficient algorithms, do you think enough resources are devoted to this kind of development and to the field of molecular simulations at large?

At the moment the field of GPU computing is moving very fast, with an active community behind it. In 2011 and 2013 we organised two symposia on GPU computing. It is interesting to see the way in which fast GPU computing has become a mainstream technique. While in 2011 many reports were still on proof-of-concepts, this year the majority reported simulation results that simply would not have been possible without the use of GPUs.

Most MD GPU code is developed by physicists and therefore application-driven. I believe that future exascale computers can only be exploited effectively if computer scientists are included more closely in this development.

In what direction will your research move in the next five to 10 years?

For me, the greatest benefit of GPU acceleration is that the computational power of desktop computers has suddenly grown by an order of magnitude. In my field of soft matter, simulations often do not need to run on hundreds of processors, but must be repeated many times to gather statistics. For these applications, a farm of desktop computers with GPUs can be as effective as a much more expensive parallel supercomputer. I think it is time for scientists to harvest this new computational power. Our group will focus on the effect of long range interactions such as hydrodynamic and electrostatic interactions, whose role in electrophoresis, translocation and nucleation is far from understood.

Efficient modelling

At the University of Stuttgart, research is ensuring computer simulations run efficiently. **ESPResSo** is a recently developed software package with potential impact in a wide range of scientific fields

COMPUTER SIMULATIONS ARE a widely applied method for understanding more about a given system from a theoretical perspective. They take into account given properties of the components of a system and use these to predict its behaviour. The study of soft matter aims to describe the class of materials that includes polymers, colloids, liquid crystals and most biological materials such as DNA, membranes and proteins. Advances in technology, biophysics and nanomaterials have sparked great interest in these materials due to their potential applications in each respective field.

Simulation methods *in silico* are one avenue through which these materials can be studied, and so there is a corresponding need for software that can accurately describe the important features of such materials. Most computer simulations in this field work from the ground up: they aim to model the materials at the atomic level. While these systems are accurate, in some cases they rapidly become computationally expensive and even unfeasible when considering large time intervals or a large number of atoms. This is particularly the case for the study of soft matter molecular dynamics (MD), where computational hurdles can prevent realistic simulations from being devised.

THE ESPRESSO SOFTWARE

Dr Axel Arnold, Junior Professor for simulation algorithms in soft matter at the University of Stuttgart, has developed an extensible simulation package for research on soft matter systems (ESPResSo) that aims to reproduce soft matter properties and interactions by MD, ie. moving particles according to Newton's equations of motion. It simplifies the problems posed by atomic-level resolution models by considering groups of atoms together, and applies this to systems that can be sufficiently described in this way. For example, polymers such as polyelectrolytes or rubber can often be modelled using simple bead-spring models, ie. charged spheres connected by springs, where each of the spheres represents a whole group of atoms, sometimes a complete monomer or even larger compounds. This method is called coarse graining, as it removes degrees of freedom from a system in order to achieve a simpler, more computationally feasible model. Coarse grained 'superatoms' can be used to represent hundred or even billions of atoms. Naturally, these superatoms behave differently from atoms and cannot be described by the normal force fields. Instead ESPResSo provides these interactions.

A particular strength of ESPResSo is charged systems, which require Coulomb and dipolar algorithms such as mesh-based Ewald summation and electrostatic layer correction (ELC), that can also handle partially or fully periodic geometries. ESPResSo's methods can be applied to a wide range of systems, including the study of non-conventional ionic liquid supercapacitors, DNA electrophoresis, crystallisation in suspensions and the flame synthesis of nanoparticles. It has also been used to examine DNA translocation through nanopores, a prospective method to sequence DNA.

The current ESPResSo 3.1 release introduces new features and extended key capabilities, including an alternative local electrostatic solver based upon the electrodynamic equations and new methods for dealing with dielectric contrasts common to coarse grained implicit water models of charged systems. It now also allows the study of agglomeration by allowing the addition of bonds during integration.

ESPResSo utilises a mesh-based Ewald method for bulk system simulations, and also provides the efficient, but less common ELC and MMM2D algorithms. In particular, ESPResSo can deal with varying dielectric constants, a requirement in coarse-grained modelling. One particularly noteworthy addition is a thermalised lattice Boltzmann (LB) solver, which allows the consideration of hydrodynamic interactions between solutes

CODES OPEN!

MD research is an open community characterised by strong collaboration. The most popular codes, for example ESPResSo, GROMACS or LAMMPS, are open source, meaning other scientists can use parts of them for their own projects. Each of these codes is a collaborative project comprising dozens to hundreds of active contributors.

ESPRESSO - USING GPUS TO ACCELERATE COARSE GRAINED MOLECULAR DYNAMICS SIMULATIONS

OBJECTIVES

Computer simulation is an important tool to gain insight into microscopic systems limited by the efficiency of available algorithms and hardware. The goal is to develop improved algorithms for state-of-the-art hardware to study systems that were previously inaccessible to simulations, and to improve our theoretical understanding of biological and soft matter.

KEY COLLABORATORS

Tanja Schilling, University of Luxembourg, Luxembourg • **Dirk Reith**, Fraunhofer Institute for Algorithms and Scientific Computing SCAI, St Augustin, Germany • **Daan Frenkel**, University of Cambridge, UK • **Kurt Kremer**, Max-Planck-Institute for Polymer Research, Mainz, Germany

FUNDING

This work is funded by the German Science Foundation through the research programme SFB 716 'Dynamic simulation of systems with large particle numbers' and the cluster of excellence 'Simulation Technologies' of the University of Stuttgart.

CONTACT

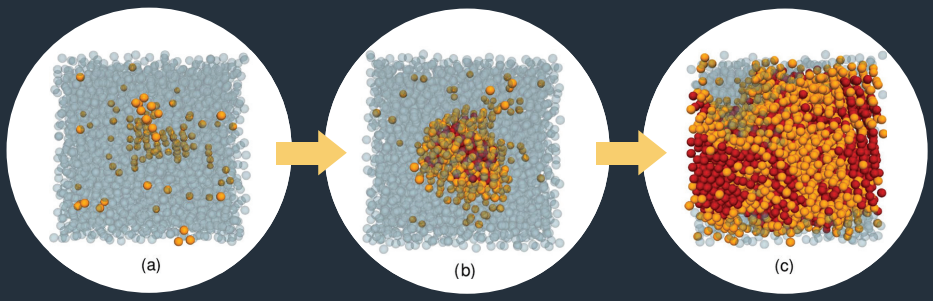
Dr Axel Arnold

Universität Stuttgart
Institute for Computational Physics
Allmandring 3, 70569 Stuttgart, Germany

T +49 711 685 67609
E arnolda@icp.uni-stuttgart.de

http://espressomd.org

AXEL ARNOLD obtained his degree in mathematics in 2001 at the University of Mainz, Germany. He defended his PhD in physics (2004) at the Max-Planck-Institute for Polymer Research, Mainz, working on the simulation package ESPResSo for charge soft matter. He then moved as a postdoc to the Frankfurt Institute for Advanced Studies, and later to the FOM-Institute for atomic and molecular physics in Amsterdam, where he investigated bacterial DNA segregation. In 2008 Axel returned to Germany for a two-year postdoc at the Fraunhofer Institute for Scientific Computing and Algorithms, and in 2010 joined the University of Stuttgart as Junior Professor for simulation algorithms in soft matter.



Simulated crystallisation in a colloidal suspension. Coloured particles are in a crystalline environment, where red marks the final, stable structure.

and their solvent: "Hydrodynamic interactions are mediated by a liquid, which we model either by dissipative particle dynamics or a LB fluid that represents the fluid as particle densities on a lattice," explains Axel. Hydrodynamic interactions have important implications in simulations of MD since they affect the movement of solute particles within a solvent and have thus far remained absent from models due to computational difficulties. Crucially, this Boltzmann solver, coupled with the MD modelling capability of ESPResSo, allows the evaluation of hydrodynamic effects in various experimental and theoretical paradigms. Since this process is computationally expensive, ESPResSo can offload LB computation to graphics processing units (GPUs), which has been implemented in order to speed up the computation and make it feasible for real-world application. Axel is currently involved in the development and testing of novel GPU setups, for example to accelerate the Coulomb algorithms.

HOW DOES THE GPU COMPARE?

The performance of ESPResSo's GPU-LB has been tested against the standard CPU-based MPI-parallelised computation – both are part of the ESPResSo package. Incredibly, the GPU combination showed performance up to 50 times better than the eight core CPU alternative. In terms of its actual application to typical simulations involving particles in solution the GPU is also superior, adding only up to 30 per cent additional computation time on top of MD simulation. This is in contrast to the CPU, which takes roughly four times longer to compute than the MD itself. On more complex simulations, where the MD computation takes longer, this additional computational time diminishes, because the two processes run in parallel. Axel has also shown that even when an entire cluster of CPUs is used it does not reach the computational power of a single GPU – because of communication overhead when the number of halo nodes is comparable to the number of real nodes present. The increased speed of the GPU even exceeds their initially calculated expectations based on Gigaflop number comparison.

EXPECTED OUTCOMES OF ESPRESSO APPLICATION

The wider impact of this change is clear; once computationally expensive and unfeasible processes such as the hydrodynamic behaviour of solvents can now be accounted for in simulations. It is thought that this may have interesting implications within a variety of fields since all biological and many nanotechnological processes happen in suspension: "This makes our GPU-accelerated LB solver so important – it practically removes the cost of computing hydrodynamic interactions," notes Axel. He has already had success in the application of this form of MD simulation to the study of soft colloidal crystallisation. His findings challenge the assumption that crystallisation is a quasi-static process, unaffected by details of particle transport other than the bulk diffusion coefficient. This challenge is significant since colloid solutions are used as a model for more experimentally complex processes including metal melts. In essence, this change in computational ability will have a tangible effect both theoretically and experimentally.

Looking forward, Axel plans to extend this to also deal with larger meshes. This will be dealt with primarily through multiple GPUs that are utilised in the same way as the CPU alternative – following the same halo code and domain decomposition as the CPU code. In addition, other ESPResSo processes that are computationally expensive will be transferred to GPU, particularly mesh-based electrostatics algorithms like P3M, ELC or MEMD.

ESPResSo

INSTITUTE FOR COMPUTATIONAL PHYSICS