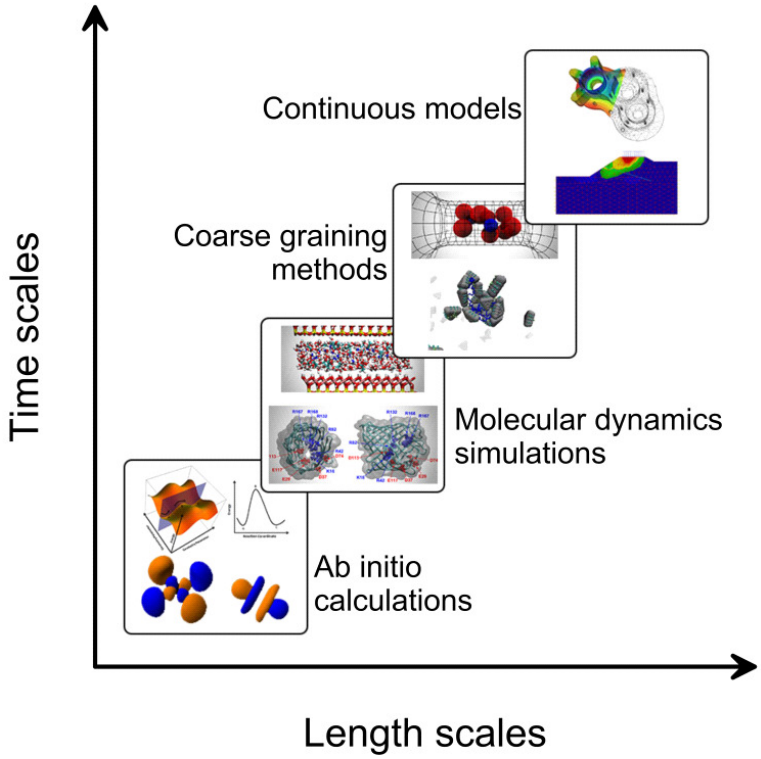
Research plan

Between two worlds; bridging natural and social sciences

At the dawn of supercomputers, people in the natural science scene faced new challenges. The instruments they used produced more and more data. In some cases the production rate was faster than the rate of doing the analysis. In the fields of astronomy, chemistry, physics, mathematics, molecular biology, etc., data are often passed to the public to do some steps of the research. This contribution can be human resource (for example classifying galaxies) or computational time (donating computational resources for distributed computing projects). Data sets were growing rapidly and traditional ways of collecting, storing, analyzing, transferring (as a whole, managing) of these big chunks of information became more and more difficult. The era of Big Data has begun.

This challenge also holds for the social sciences. Venturini said (Venturini et al., 2015), "In the last decade, the spread of digital technologies has flooded the study of social phenomena with more data than ever dreamed of." There are many distributed computing projects in the field of natural sciences. In the field of social sciences one example is Ideologias@Home. As the project description says: "The aim of the project Ideologias@Home is to study how people in a certain region evolve ideologically over time with respect to an idea. When an idea is introduced in a society, the population is divided naturally into four groups: Extremists: those who defend the idea extremely. Moderates: those who defend the idea moderately. Opponents: those who are against the idea. Abstentionists: those who do not care, abstain or have no opinion. People change their minds because of peer pressure, the influence of mass media or because they come to a different conclusion of their own accord. Under this assumption, we propose dynamic models, determine the parameters, predict trends and analyse results." [http://infernonet.com/en/project-browser/ideologiashome/ Accessed June 22, 2016.] To my best knowledge, high performance computing are not used as widespread as in natural sciences.

So far I have studied systems from the perspective of natural sciences. I am mainly interested in molecular simulations of materials. I (together with coworkers) use microscopic and mesoscopic models to describe the behavior of ordered, disordered, homogenous and inhomogeneous, gaseous and condensed materials in physical, chemical, and biological systems. We try to get macroscopic properties of the system from the employed microscopic input model. We can also measure some of these macroscopic physical quantities and compare our results from the simulations with them. Although simulations at microscopic level, typically producing results in very small spatial and time dimensions, tend to larger scales, their results do not always meet with the data obtained from common experimental measurements.



If we would like to achieve deeper understanding of a phenomena (for example, What is the reason that one type of membrane is more selective than the other? or, Why does a biological ion channel rectify?) we should use modeling and, today, computer simulations. The questions we ask determine the tools that are conceivable for handling the length and time scales of the processes we are interested in. If we are interested in molecular orbit of an electron we should not employ large-scale continuum mechanics models, because they omit the corresponding level of information. But vice versa, if we would like to get some information about flow patterns in a pipe system, the main tool should not be quantum mechanics because of the spatial size of the system. It is very important to choose the right model and the right technique for the simulations. However, deciding what is the optimal model and the most suitable method is far from being trivial.

It is an inverse problem; we choose a model and a simulation method to produce some results. If these results make sense and (in the ideal case) reproduce experimental findings, we can look into our simulation and find what else can we say about our system (either on micro-scale or macro-scale). It is important that there are overlaps between the techniques and, more importantly, they are needed to be overlapped, because that is the way we can connect the different levels of modelling together.

I am a member of the Computer Modeling Research Group at the Department of Physical Chemistry at the University of Veszprém (http://almos.uni-pannon.hu/~boda/test/). I also expressed my intention to participate in the Multiscale Modelling Research network developed by Dr. Boda in the Complexity and Big Data Centre of iASK (http://www.iask.hu/hu/s/3533/multiscale-modelling-of-complex-molecular-systems). My main profile is conducting molecular simulations with the molecular dynamics method. In my proposed research, I would like to continue my work and advance towards larger systems (in length and also time scales).

I intend to continue our work on the molecular simulation of layer silicate intercalation phenomena. The layer silicate we studied recently in detail is kaolinite. This clay mineral exhibits an excellent performance when used as a nanofiller in the production of clay/polymer nanocomposites due to its natural occurrence with a relatively high purity and almost perfect layer structure. Classical molecular simulation is a suitable tool to study the adsorption and intercalation of molecules in clays, though finite size effects limit the number of observable properties and phenomena. Recently, we simulated kaolinite intercalation with realistic layer size [Ható et al. 2014]. In ordinary atomic simulations using 10^3-10^4 atoms, the size of the model system (and/or the time step) basically determines the properties that can be sampled. Today, it is not unusual to consider large systems (10^5-10^{10} atoms) in detailed all-atom MD simulations (e.g. modeling nucleation effects, complex biological systems, or material failures at the nanoscale). Despite the apparent advantage of being able to extend the scope of investigated phenomena, even today, large-scale all-atom simulations involve a substantial computational effort and were therefore rarely utilized in the field of clay minerals.

I would also like to do simulations on nanopores constructed from PET (Polyethylene terephthalate). These artificial channels are made by bombarding a polymer sheet with high energy particles resulting in penetration and a “hole” in the fabric of the material. With chemical methods, we can fine-tune the properties of the inner surface of the tube, allowing us to engineer artificial pores with desired properties (for example ion selectivity or rectification). My main contribution can be dealing with the molecular dynamics simulations and providing anchor points to other levels of modeling. Parameters of models on a more coarse grained modeling level as used by Dr. Boda and his coworkers [Boda et al. 2012] can be determined by fitting the results of their calculations to the results of my molecular dynamics simulations.

I am also familiar with Monte Carlo simulation methods. Currently these are our primary tools for simulations [Ható et al. 2012] in the Computer modeling research group at the Department of Physical Chemistry. We are constantly seeking new ways of simulating diffusion processes in porous media and incorporating these into existing simulation tools. [Ható et al. 2015]

Monte Carlo is a stochastic computational method, so my experience in it would allow me to participate in the work of the Complexity and Big Data Centre in stochastic modelling of various processes. With this scholarship I could allocate more human time for the multi-scale

modelling project. I am very curious whether the above described modeling techniques are applicable for social science researches. Certainly, I know, this is quite a different field, but maybe with the help of good models we can handle masses of people as we describe millions of atoms in our simulations. The time-independency of the models we use currently are the weak points in this analogy. If we are dealing with interacting particles, the repulsion or attraction between the particles are the same today, the next year, and was the same in the last century. This is not the case if we treat individuals as interacting centrums. People can change their minds, and they do it often as they got some input information from the outside world. So the relationships between individuals or between a person and an idea (or meme) change over time. It is not the same as it was hundred years ago. So in social science outdated input data –in the sense I described above–is a kind of problem.

Personally, I think that we should not seek individuals, rather we should use the tools of statistics as we use them in statistical mechanics. Maybe we could find adequate models for well-chosen smaller or larger groups of people and with it we will be able to answer some well-posed questions. Posing good questions is the decisive step in the scientific process, in my view.

A starting point can be for example studying the model of coordinated animal motion such as bird flocks and fish schools. [C. W. Reynolds 1987] With these so called Boids one can model the very complex motion behavior of a group of animals. And the most beautiful thing is, that it is governed by fairly simple rules, and also that it is closely related to particle systems modeling smoke, clouds or ocean waves. Yet it is different, from the model of the interacting particles. The difference is that Boid behavior is dependent not only on internal state but also on external state [C. W. Reynolds 1987]. Also there are epidemic models describing the transmission of diseases from person to person. In my opinion these can be used as a starting point to model spreading of ideas or memes (or as they are often called “viruses of the minds”).

It is clear for the reader that I am not an expert in social sciences, therefore during my scholarship I would like to broaden my horizon and learn from social scientists. I will try to find connections between these two worlds and together with the colleagues we should build bridges (on the foundations made by modelling) that are strong enough to connect these different perspectives of science. I offer my programming skills to help in the projects of the Complexity and Big Data Centre beyond the Multiscale Modelling group of Dr. Boda. I am open to learn new things.

References:

D. Boda, D. Gillespie: Steady-State Electrodiffusion from the Nernst–Planck Equation Coupled to Local Equilibrium Monte Carlo Simulations Journal of Chemical Theory and Computation 8, 824 (2012)

Z Ható, D Boda, T Kristóf: Simulation of steady-state diffusion: Driving force ensured by dual control volumes or local equilibrium Monte Carlo The Journal of Chemical Physics 137 (5), 054109 (2012)

Z Ható, G Rutkai, J Vrabec, T Kristóf: Communication: Molecular simulation study of kaolinite intercalation with realistic layer size The Journal of Chemical Physics 141 (9), 091102;(2014)

Z Ható, Á Kaviczki, T Kristóf: A simple method for the simulation of steady-state diffusion through membranes: pressure-tuned, boundary driven molecular dynamics Molecular Simulation, 1-10 (2015)

C. W. Reynolds: Flocks, Herds, and Schools: A Distributed Behavioral Model SIGGRAPH 87 (Computer Graphics 21(4), July 1987, edited by Maureen C. Stone, pages 25-34).

T. Venturini, P. Jensen, B. Latour : Fill in the Gap. A New Alliance for Social and Natural Sciences Journal of Artificial Societies and Social Simulation 18 (2) 11;(2015)