LA-UR-13-25808
Approved for public release; distribution is unlimited.

Title: Recent and future developments in the VOTCA package

Author(s): Junghans, Christoph

Intended for: Simulating Soft Matter with ESPResSo, 2013-10-07/2013-10-11 (Stuttgart, Germany)

Issued: 2013-10-16 (rev.1)

Disclaimer:
Los Alamos National Laboratory, an affirmative action/equal opportunity employer,is operated by the Los Alamos National Security, LLC for the National NuclearSecurity Administration of the U.S. Department of Energy under contract DE-AC52-06NA25396. By approving this article, the publisher recognizes that the U.S. Government retains nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes. Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Departmentof Energy. Los Alamos National Laboratory strongly supports academic freedom and a researcher's right to publish; as an institution, however, the Laboratory does not endorse the viewpoint of a publication or guarantee its technical correctness.

Recent and future developments in the VOTCA package

Christoph Junghans

July 24, 2013

Coarse-graining is a systematic way of reducing the number of degrees of freedom used to represent a system of interest. The Versatile Object-oriented Toolkit for Coarse-graining Applications (VOTCA) provides a uniform interface to commonly used coarse-graining techniques such as iterative Boltzmann inversion, force-matching, and inverse Monte Carlo. Further, it provides a flexible modular platform for the further development of new coarse-graining techniques.

Recently two new methods for coarse-graining have been added to the package and were tested on SPC/E water and methanol-water mixtures. We will discuss these results in comparison to earlier structure-based studies, but also talk about the development of a non-structure-based model.

Finally, we will discuss features for the upcoming release including interfaces to more simulation packages Additionally, we will debate how to make the development process more adapted to the distributed developer team and, at the same time, allow for better testing to guarantee code quality.

Recent and future developments in the VOTCA package

Christoph Junghans

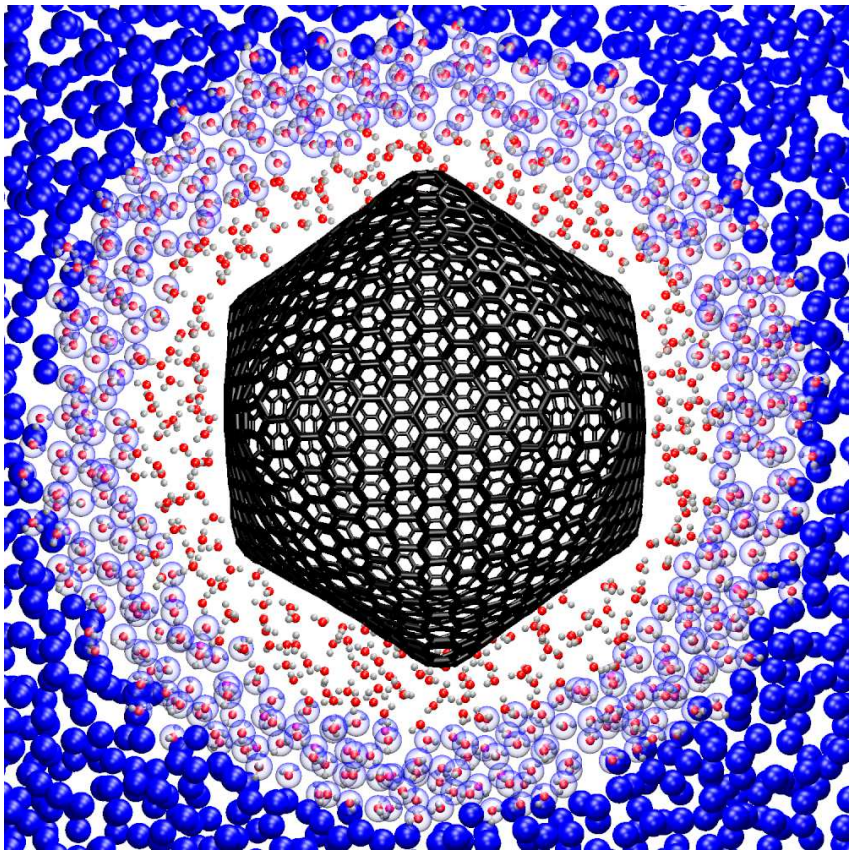
Los Alamos National Laboratory
NM, USA

Oct 9, 2013

Introduction

Coarse-graining is an essential part of multi-scale simulations!

- Reduces number of degrees of freedom
- Enhances accessible range of time- and length-scales
- Links atomistic and coarse-grained representations



Introduction

Systematic Coarse-Graining

Is there a force-field for the coarse-grained model which reproduces a certain property?

- Structure (e.g. bond distribution or two-body correlations):
  - (Iterative) Boltzmann inversion
  - Inverse "Monte Carlo"
  - *Relative Entropy Method*
- Forces $\rightarrow$ Force matching (multi-body PMF)
- Free energy (MARTINI force-field)
- Further properties:
  - General $\rightarrow$ *Optimization*
  - Pressure $\rightarrow$ Pressure correction
  - Diffusion $\rightarrow$ Thermostat (friction constant fitting)
  - Kirkwood-Buff Integrals $\rightarrow$ *Ramp correction*[1]

[1] Ganguly et al., JCTC 8, 1802 (2012)

Introduction
VOTCA Framework

- Consistent implementation of most of these methods
  $\rightarrow$ Allow for direct comparison
- Platform for the implementation of new methods
- Integrate existing sampling programs (e.g. MD codes)

Parts of VOTCA[2]- www.votca.org

- Mapping engine
- Parallel analysis framework
- Automated iterative coarse-graining
- Charge transport modules

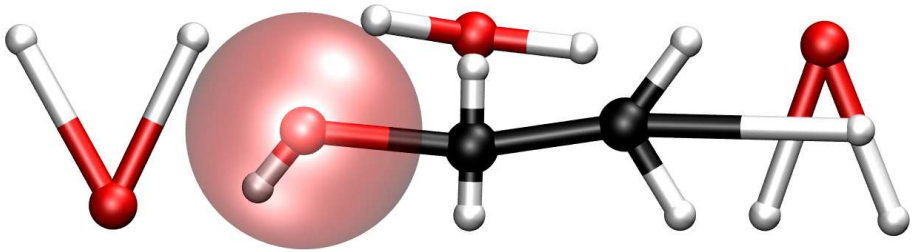
- Ohloh: 12 Person Years / 49.5k Lines / $ 647.5K
- 15 Developers
- Packages in Fedora, OpenSuse, Gentoo

[2] JCTC 5, 3211 (2009) & Macromol. Theo. Simul. 20, 472 (2011)

Introduction
VOTCA Team

Core developers

Christoph Junghans
Victor Rühle



Versatile **O**bject-oriented **T**oolkit for **C**oarse-graining **A**pplications

Modular C++ kernel
Scripting for iterative workflow
Simple integration of other simulation packages

Iterative Boltzmann inversion
Inverse Monte Carlo
Force matching

Implementations

Sebastian Fritsch	interface to AdResS
Mara Jochum	SIMPLEX algorithm
Konstantin Koschke	parallel analysis engine
Alexander Lukyanov	force-matching
Sikandar Mashayak	relative entropy method
Tristan Bereau	orig. interface to ESPResSo
Dominic Röhm	interface to ESPResSo
Louis Vernon	interface to LAMMPS

Los Alamos NATIONAL LABORATORY EST. 1943

Kirkwood-Buff Models

Introduction

Find a coarse-grained model that reproduces the Kirkwood-Buff Integrals:

$$G_{ij} = 4\pi \int_0^\infty [g_{ij}^{\mu VT}(r) - 1] r^2 \, \mathrm{d}r$$

Motivation

Describes salting-in/salting-out of Biomolecules on a coarse-grained level:

$$f_{cc} = \left(\frac{\partial \ln \gamma_c}{\partial \ln \rho_c} \right)_{p,T} = -\frac{\rho_c \left(\mathrm{G}_{cc} - \mathrm{G}_{cw} \right)}{1 + \rho_c \left(\mathrm{G}_{cc} - \mathrm{G}_{cw} \right)},$$

$k_B T \ln \gamma_c$: co-solvent solvation free energy
γ_c: co-solvent molar scale activity coefficient
ρ_c: co-solvent number density

Assumption: large systems ($g^{\mu VT} \approx g^{NVT}$)

Los Alamos NATIONAL LABORATORY EST. 1943

Kirkwood-Buff Models

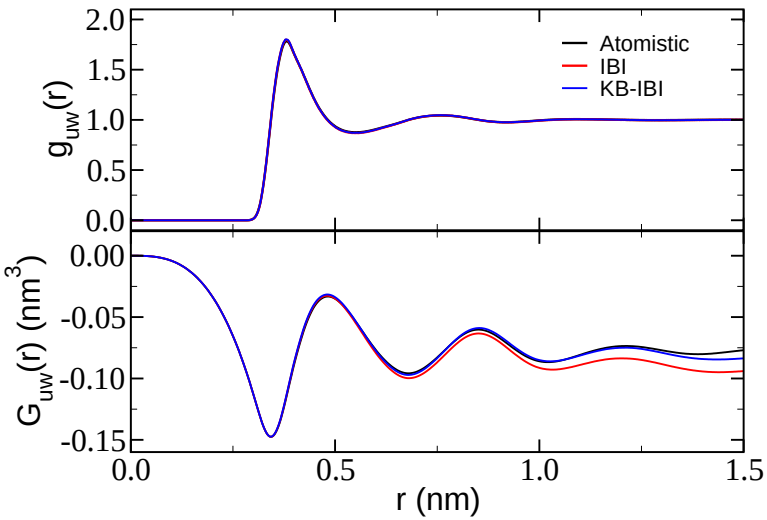
Aqueous Urea Mixture

Algorithms

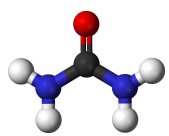
- Infinitely long iterative Boltzmann inversion $\rightarrow$ fails

$$\Delta U_{ij}(r) = \mathrm{k_B}\, T \ln \frac{g_{ij}(r)}{g^{\mathrm{Ref}}(r)}$$

- Ramp correction [3]



(aqueous urea mixture at 4.7M)



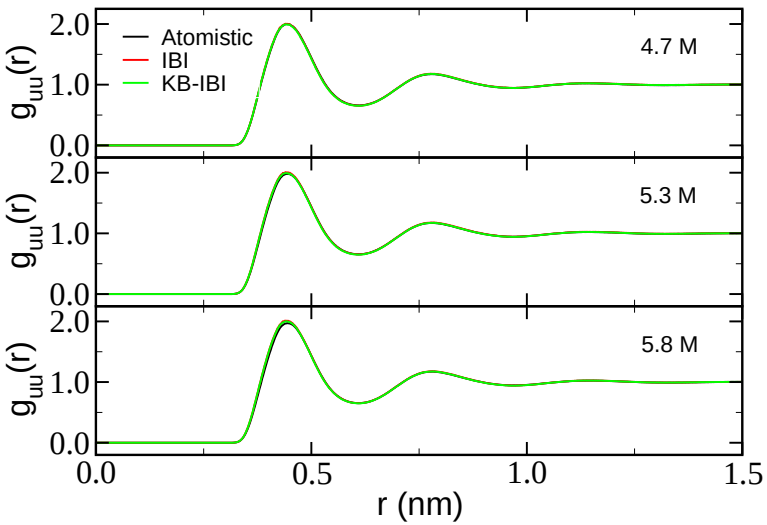
$$\Delta U_{ij}(r) = A(\mathrm{G}_{ij}^{(n)} - \mathrm{G}_{ij}^{(ref)}) \left(1 - \frac{r}{r_{\mathrm{cut}}}\right)$$

Problem: A is difficult to determine

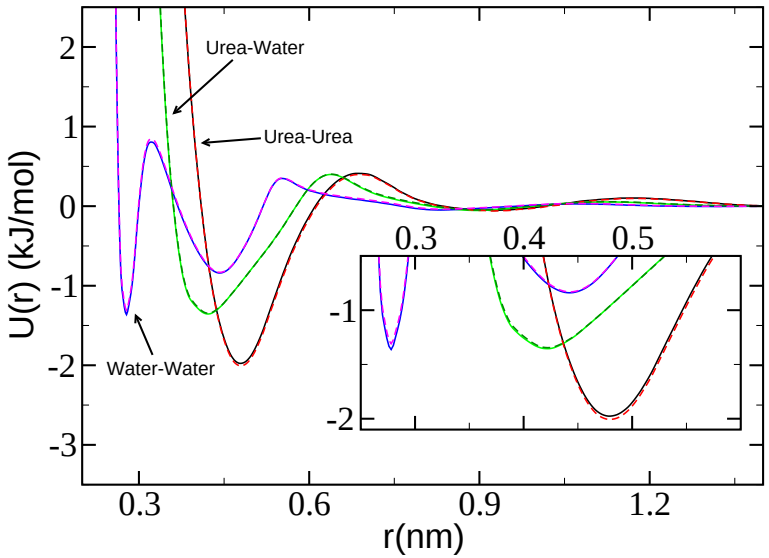
[3] Ganguly et al., JCTC 8, 1802 (2012)

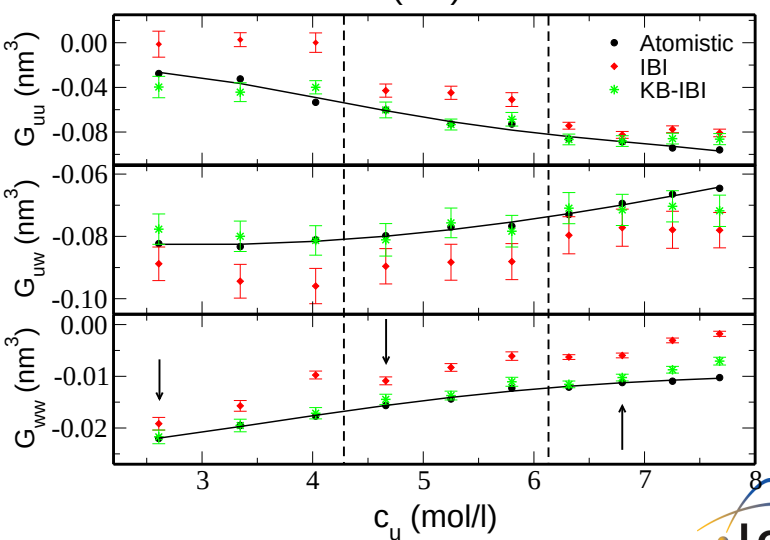
Kirkwood-Buff Models

Aqueous Urea Mixture



- Minimal differences in the potential
- Potentials are transferable in a small concentration interval

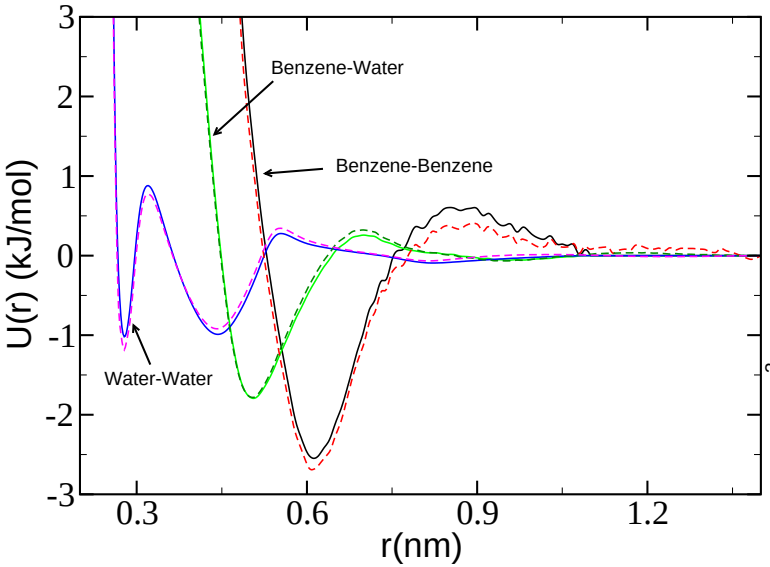


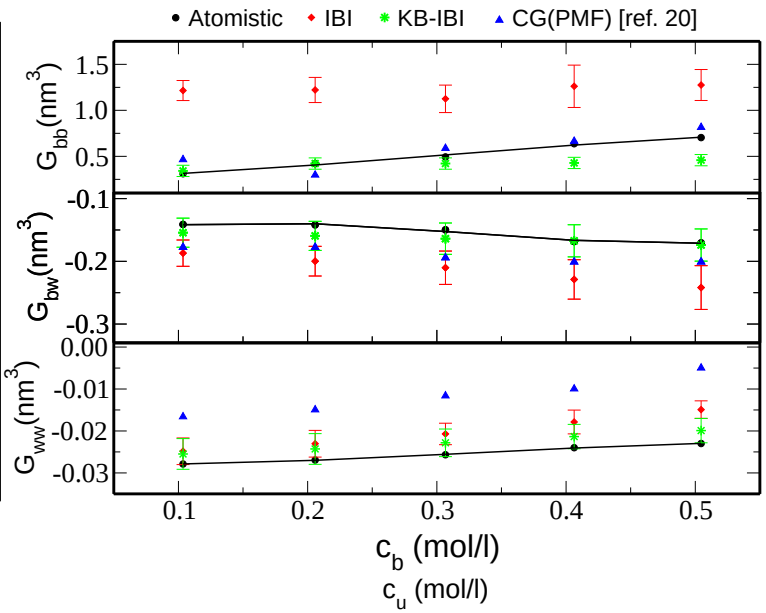


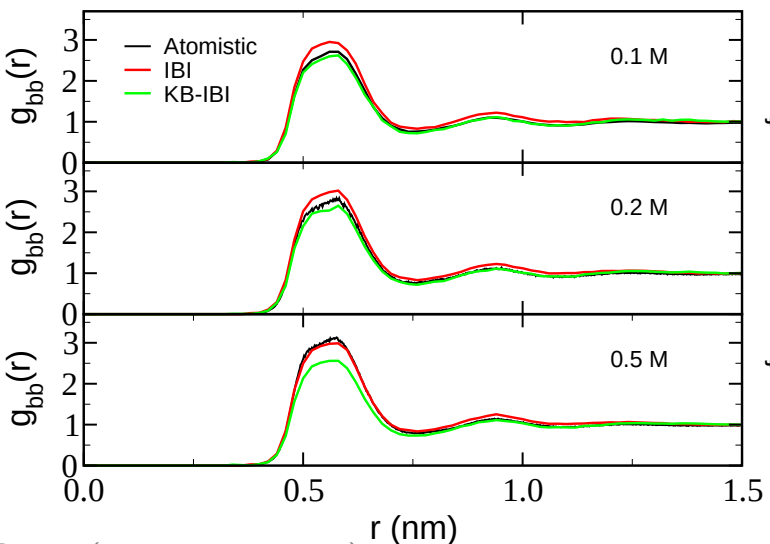
Los Alamos NATIONAL LABORATORY EST. 1943

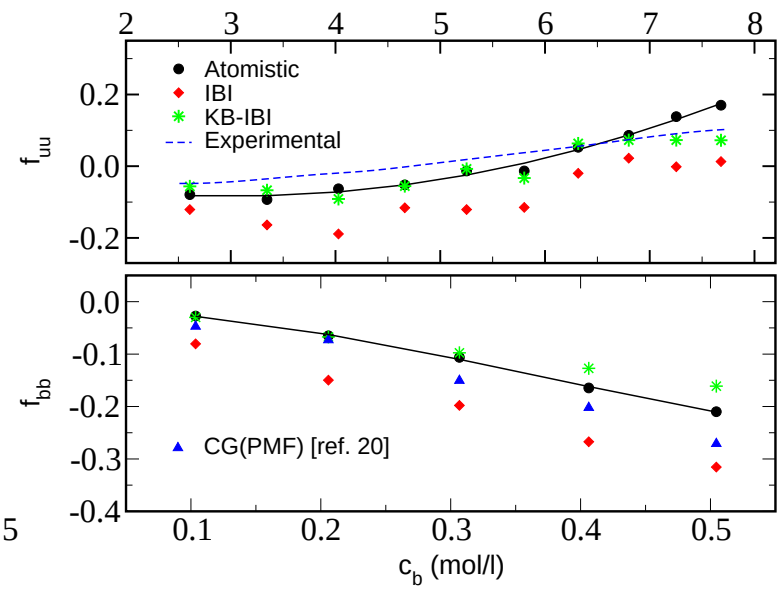
Kirkwood-Buff Models

Benzene in Water









Kirkwood-Buff Models

Conclusion

What did we learn?

- Iterative Boltzmann inversion alone is not enough
- Transferable potentials over different concentrations
- Useful method to develop models to study salting-in and salting-out

Open questions:

- Are there less arbitrary ways of correcting?
- Is it possible to incorporate the correction in an inversion scheme?

Targeted Coarse-Graining

Introduction

Find a coarse-grained model, which reproduces other non-structural related property.

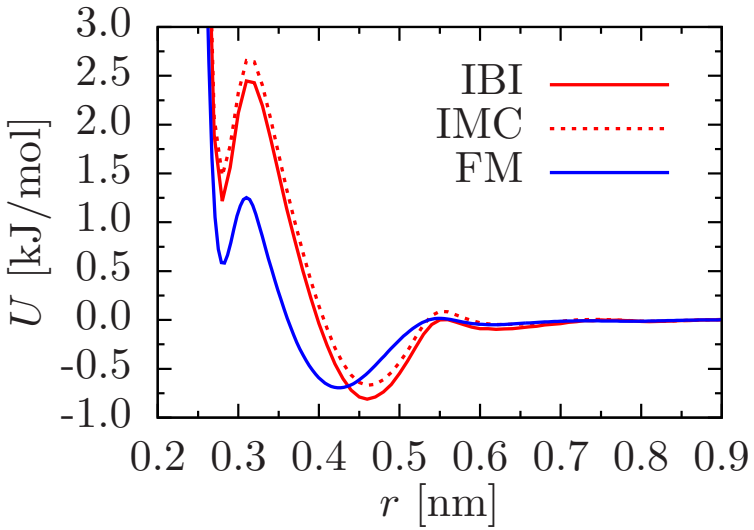
Reformulation

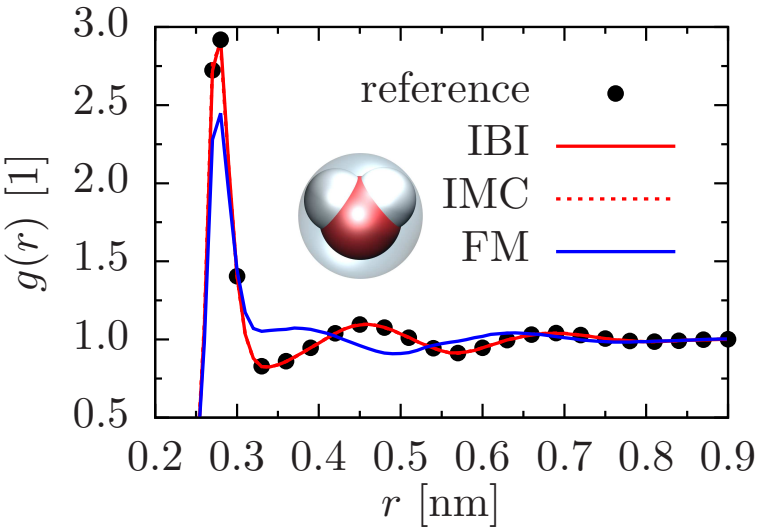
Use n ($\sim$5000) input parameters (potential tables) to generate m output parameters (properties measured in the MD simulation) and rank their quality.

- The problem is overdetermined $\rightarrow$ use $\sim$10 essential parameters
- Equivalent to a standard optimization problem
- Minimization would be possible if all ∂input/∂output exist

Targeted Coarse-Graining

Example: Water



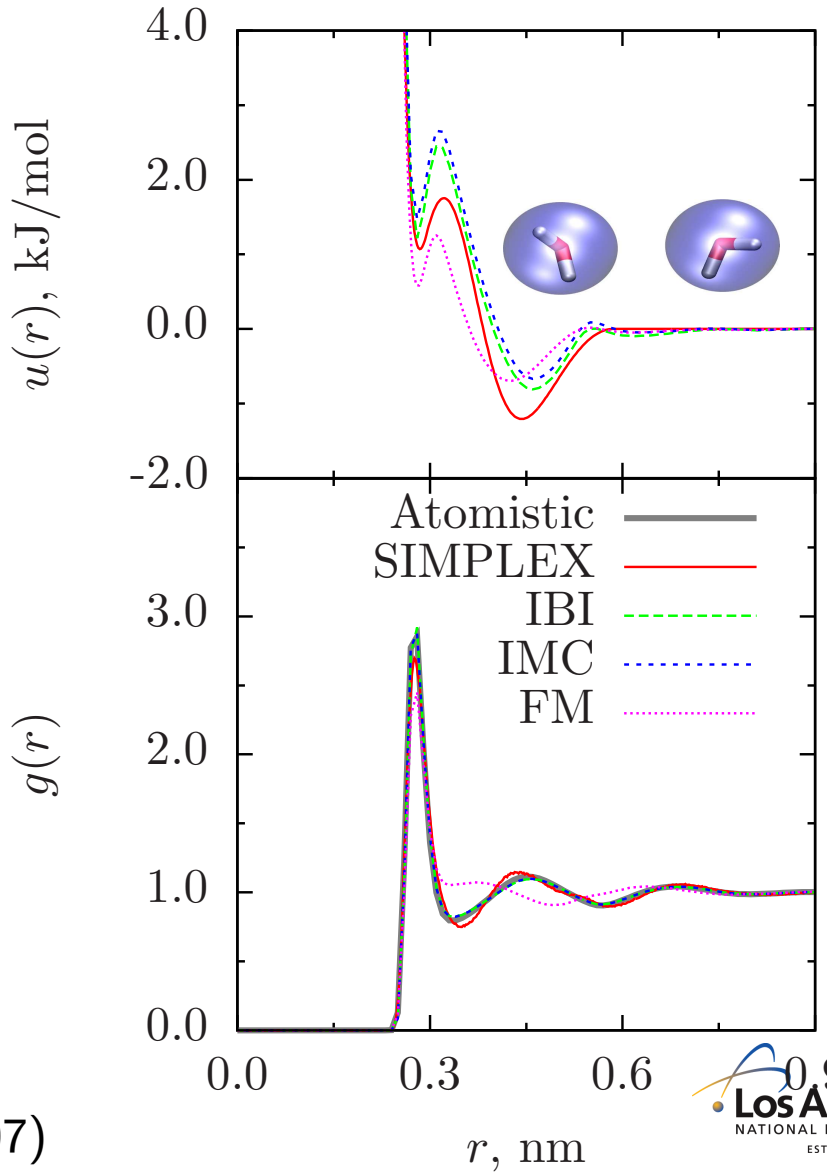


Potentials should have 2 minima.

Targeted Coarse-Graining

Example: Water

- Center of mass mapping
- CKD (= WCA + $\cos^2$ attraction) + Gaussian (6 parameters)[4]
- Optimize parameters with Nelder-Mead method (Simplex)[5]



[3] Idea: M. Jochum, Phd Thesis

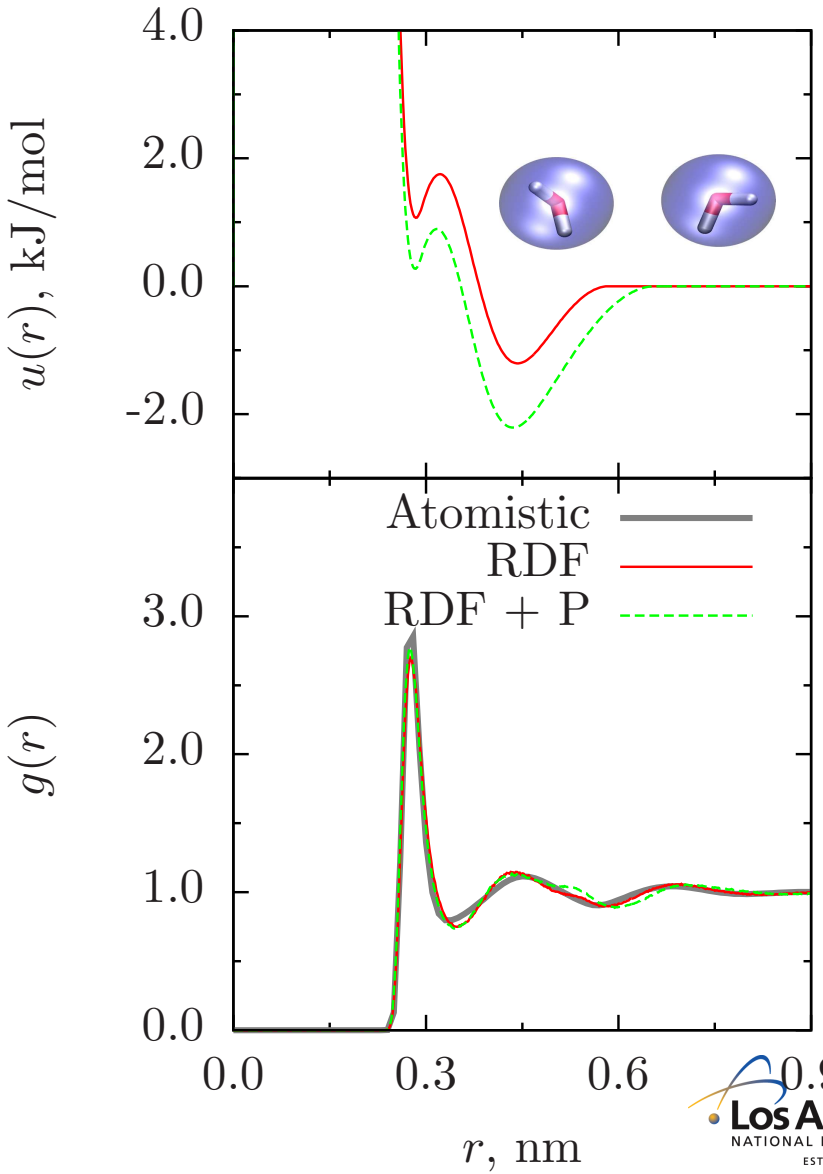
[4] Shinoda et al., Mol. Sim. 33, 27 (2007)

Targeted Coarse-Graining

Example: Water

What about the pressure?

- Can easily be incorporated
- Objective (penalty) function needs modification



Targeted Coarse-Graining

Example: Water

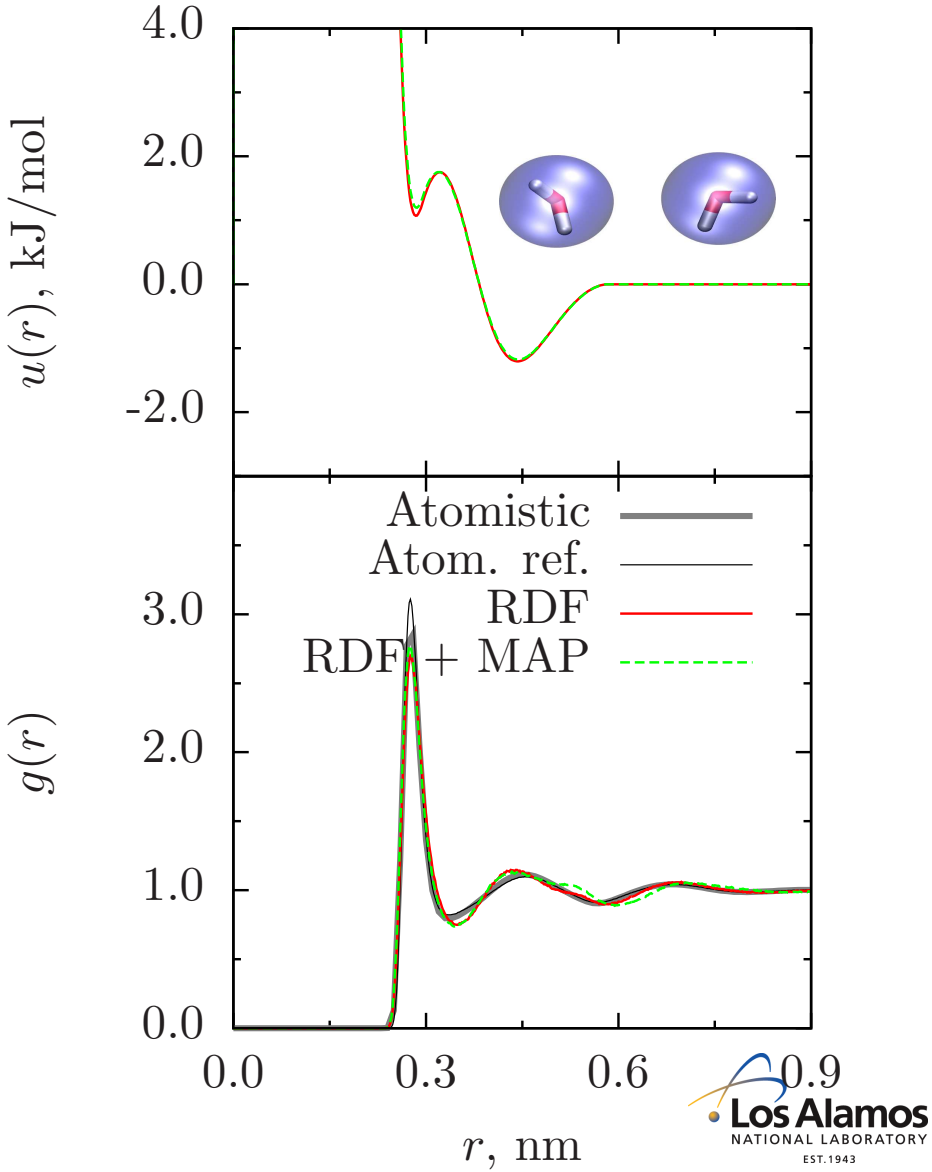
What about the mapping?

$$\vec{R} = \sum_i \lambda_i \vec{r}_i$$

with

$$\sum_i \lambda_i = 1$$

- Can easily be incorporated
- Adds 1 extra parameter for symmetric mappings
- Objective (penalty) function needs no modification
- Reference rdf changes



Targeted Coarse-Graining

Conclusion

What did we learn?

- 6 parameters are enough, but simple LJ (2) is not
- Potential is short ranged
- Other target properties can be incorporated
- Simplex is fast, but can be trapped, inefficient for $\geq$10 parameters
- Use of learning optimizers (e.g. CMA Evolution Strategy or genetic algorithms) possible
- Functional potential can speed up the simulations
- Mapping can be optimized as well

The optimization view provides a framework to aim for a broader class of coarse-grained models.

Relative Entropy Method

Introduction

Find a coarse-grained model, which minimizes the relative entropy[6]:

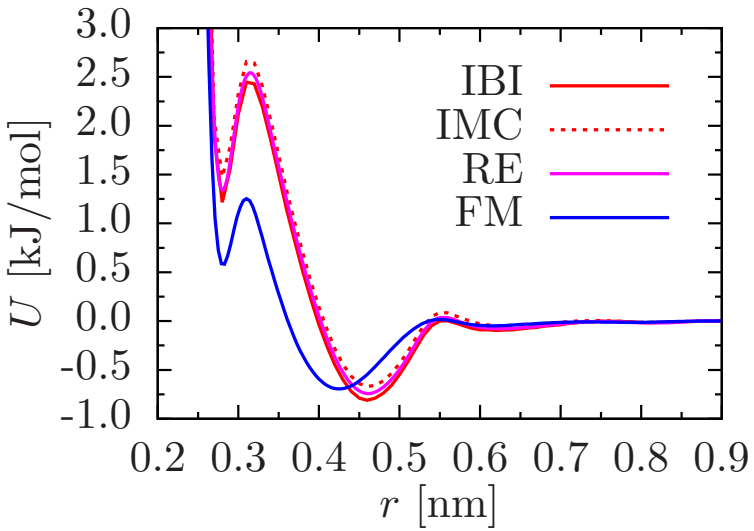
$$S_{\text{rel}} = \sum_i p_{\text{AA}}(r_i) \ln \left(\frac{p_{\text{AA}}(r_i)}{p_{\text{CG}}\left(M(r_i)\right)} \right) + \langle S_{\text{map}} \rangle_{\text{AA}}$$

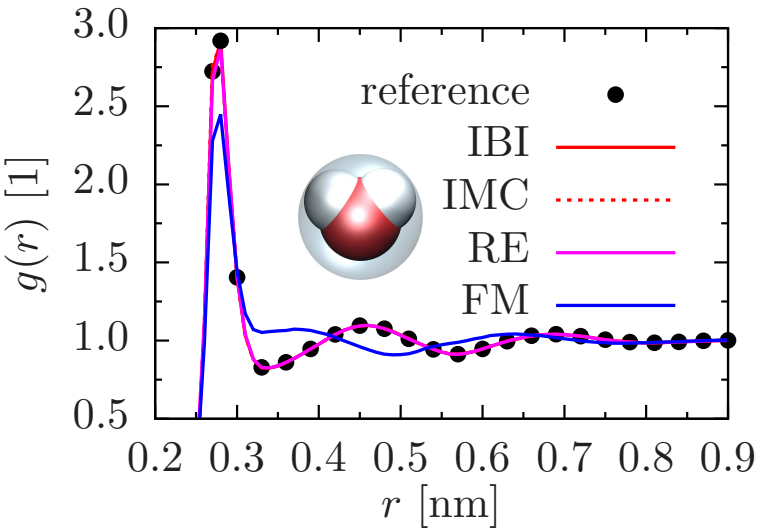
- Minimal S_{rel} = maximum likelihood that configuration of the model CG is representative of the target AA ensemble
- Minimizing S_{rel} optimizes the CG model
- Closely related to the inverse Monte Carlo method

[6] M. S. Shell, JCP **129**, 1441 (2008)

Relative Entropy Method

Example: Water





Similar to previous results, but method can be used for non-bulk cases as well[7].

[7] S. Y. Mashayak and N. R. Aluru, JCP **137**, 214707 (2012)

More Simulation Backends

Example: Water

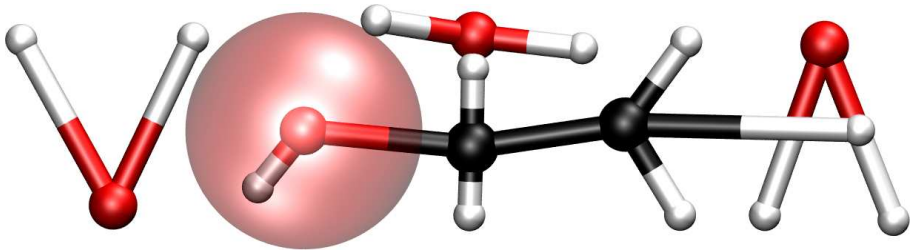
VOTCA was developed around GROMACS, but other packages might have other special feature and more advanced technique

- In 2010 (together with Tristan) first ESPResSo interface
- Interface very restricted, because too similar to GROMACS interface
- Support for ESPReSo, LAMMPS, dl_poly and ESPReSo++ was added through a very thin interface
- Let the user write the simulation script and VOTCA calls it.
- dl_poly interface is similar to GROMACS interface as both are made for atomistic simulations

Conclusion
VOTCA Team

Core developers

Christoph Junghans
Victor Rühle



Versatile **O**bject-oriented **T**oolkit for **C**oarse-graining **A**pplications

Modular C++ kernel
Scripting for iterative workflow
Simple integration of other simulation packages

Iterative Boltzmann inversion
Inverse Monte Carlo
Force matching

Implementations

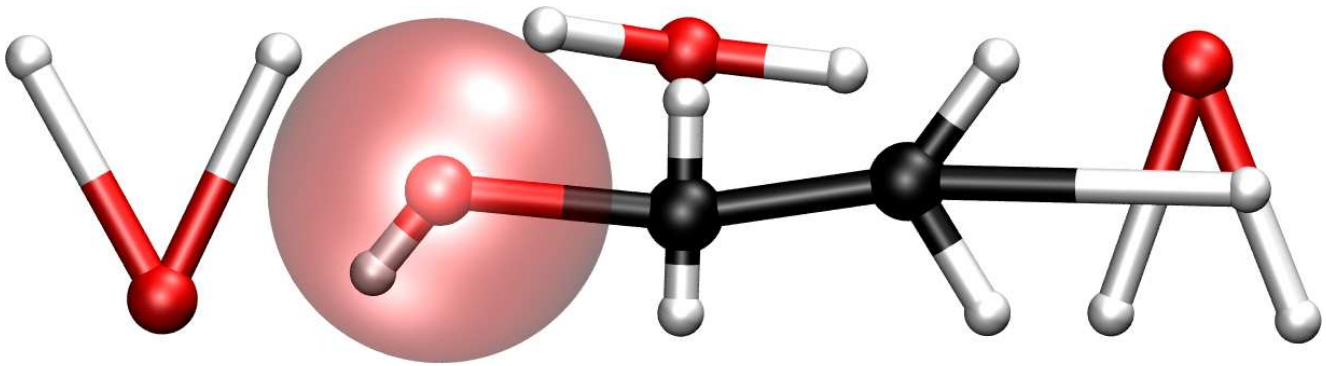
Sebastian Fritsch	interface to AdResS
Mara Jochum	SIMPLEX algorithm
Konstantin Koschke	parallel analysis engine
Alexander Lukyanov	force-matching
Sikandar Mashayak	relative entropy method
Tristan Bereau	orig. interface to ESPResSo
Dominic Röhm	interface to ESPResSo
Louis Vernon	interface to LAMMPS

Los Alamos NATIONAL LABORATORY EST. 1943

Conclusion
VOTCA Package

Linus Torvalds:

Talk is cheap, show me the code.



Versatile **O**bject-oriented **T**oolkit for **C**oarse-graining **A**pplications

Modular C++ kernel
Scripting for iterative workflow
Simple integration of other simulation packages

Iterative Boltzmann inversion
Inverse Monte Carlo
Force matching

- It's free
- All examples are in the tutorial
- It's flexible and expandable

Visit us at www.votca.org

Los Alamos NATIONAL LABORATORY EST. 1943

Conclusion
Acknowledgments

$$$

- Max Planck Society
- SFB 625 – “From Single Molecules to Nanoscopically Structured Materials”
- Department of Energy, Office of Science

Los Alamos NATIONAL LABORATORY EST. 1943

The End

Thank you for your attention !