Coarse-graining using the VOTCA package

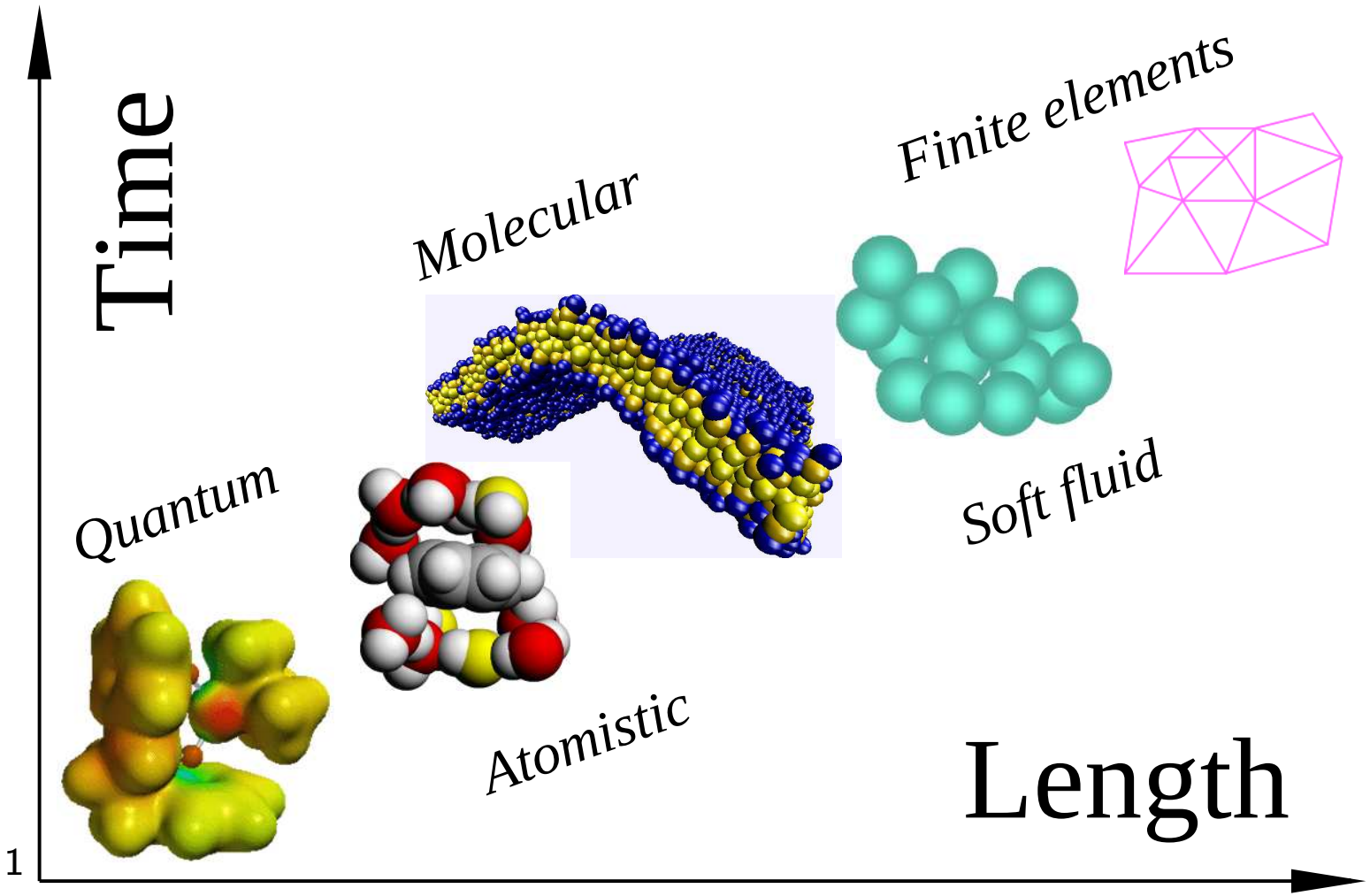
Christoph Junghans

Los Alamos National Laboratory
NM, USA

May 21, 2012

Introduction

Time- and length-scales



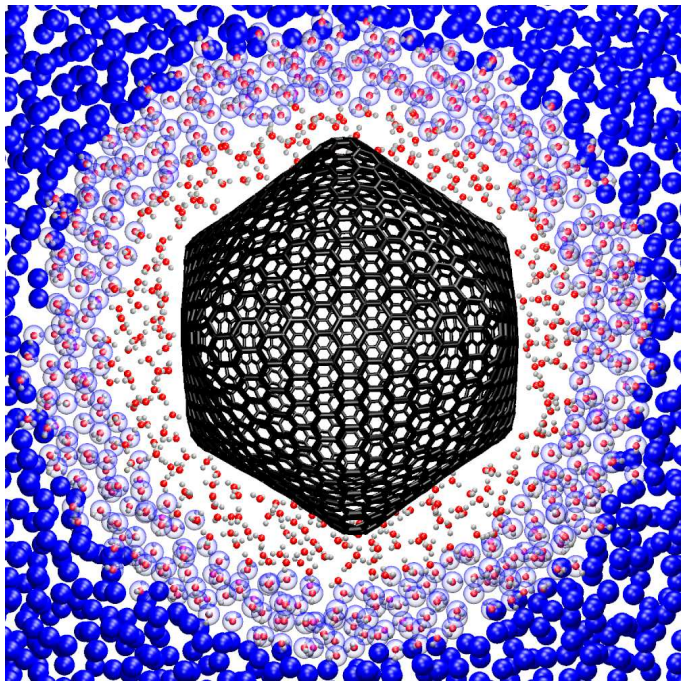
[1] D. Andrienko and K. Kremer, 2007

Introduction
Systematic Coarse-Graining

Systematic Coarse-Graining

Find a simplified model, which reproduces selected properties of a given (atomistic) reference system.

- Bottom-up approach
- Reduce number of degrees of freedom
- Enhance accessible range of time- and length-scales
- Link atomistic and coarse-grained representations

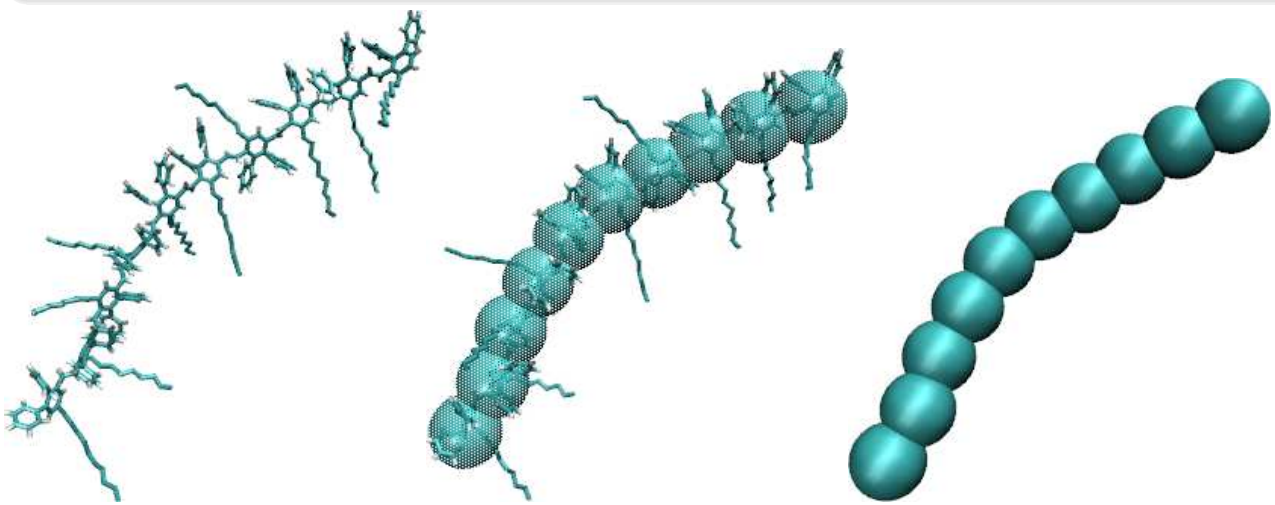


Los Alamos NATIONAL LABORATORY EST.1943

Introduction
Mapping

Mapping matrix

$$\underbrace{\vec{R}}_{\text{Vector of all coarse-grained coordinates}} = \underbrace{\hat{M}}_{\text{Mapping matrix (block diagonal)}} \underbrace{\vec{r}}_{\text{Vector of all reference coordinates}}$$



Methods
Coarse-Graining

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):
  - Boltzmann inversion
  - Iterative Boltzmann inversion
  - Inverse "Monte Carlo"
- Forces $\rightarrow$ Force matching (multi-body PMF)
- Free energy $\rightarrow$ Simplex (MARTINI force-field)
- Further properties:
  - Pressure $\rightarrow$ Pressure correction
  - Diffusion $\rightarrow$ Thermostat (friction constant fitting)
  - Kirkwood-Buff Integral $\rightarrow$ Ramp correction

Incomplete list!

Methods

Boltzmann Inversion

- Canonical sampling $\rightarrow$ Boltzmann distribution

$$P(\{q\}) = \exp\left[-\frac{U(\{q\})}{k_B T}\right]$$

- Assumption: Independent degrees of freedom

$$P(\{q\}) = P(q_1)\,P(q_2)\cdots P(q_N)$$

Invert the Boltzmann Distribution

$$U(q) = -\mathrm{k_B}\,T\,\ln P(q)$$

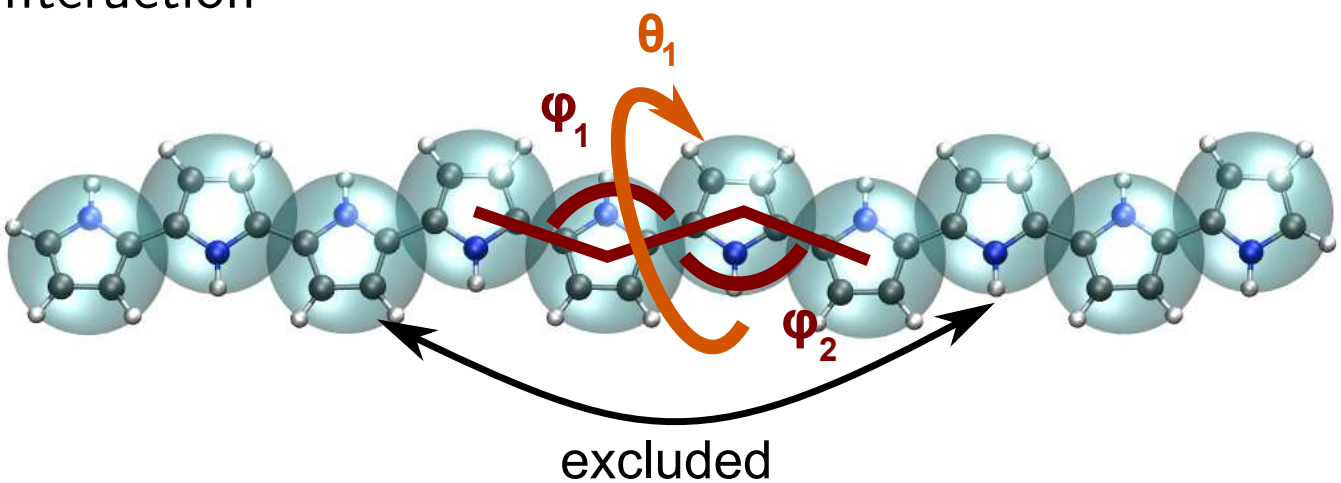
- Used for bonded interactions in e.g. polymers

[1] W. Tschöp et al., Acta Polymerica **49**, 61-74 (1998)

Methods

Separation ansatz

- "Special trajectory" to decouple bonded and non-bonded interactions
- Create exclusion list for all atoms that do not share an bonded interaction



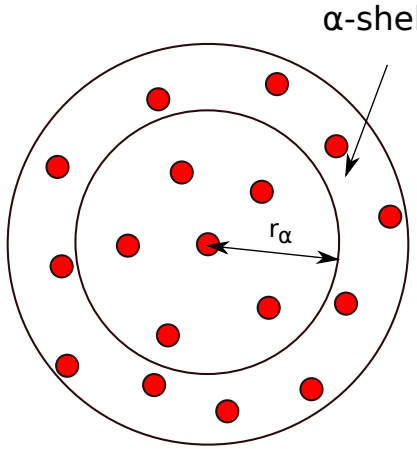
- Derive non-bonded interactions via a different method

[1] W. Tschöp et al., Acta Polymerica **49**, 61-74 (1998)

Methods

Iterative procedures for non-bonded interactions

- Try to fit target distribution function (e.g. $g(r)$)



$g(r)$: radial distribution function

S_α : number of of particles in alpha-shell

$$S_\alpha = \frac{N(N-1)}{2} \frac{4\pi r_\alpha^2 \Delta r}{V} g(r_\alpha)$$

- Initial guess for iterative process

$$U_0 = -k_b T \ln g(r)$$

- Run coarse-grained simulations, adjust potential to fit distribution function

$$U_{n+1} = U_n + \Delta U_n$$

Methods

Iterative Boltzmann inversion

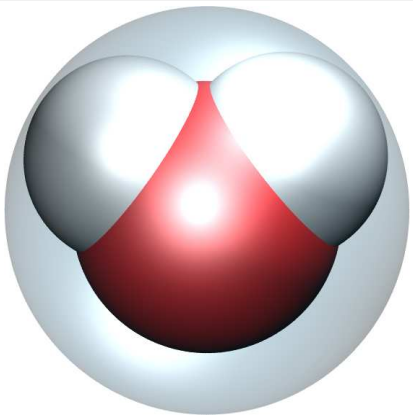
Potential Update[2]

$$\Delta U_n(r) = \mathrm{k_B}\, T \ln \frac{g_n(r)}{g^{\mathrm{Ref}}(r)} = U^{\mathrm{Ref}}_{\mathrm{PMF}}(r) - U_{\mathrm{PMF},n}(r)$$

where the potential of Mean Force is:

$$g(r) = \mathrm{e}^{-U_{\mathrm{PMF}}(r)/\mathrm{k_B}\, T}$$

- Fits selected distribution functions
- Cross-correlations neglected in update



[2] D. Reith et al., J. Comp. Chem. **24**, 1624-1636 (2003).

Methods

Inverse Monte Carlo

Potential Update[3]

Expanding number of interacting particle pairs S_α at distance α up to first order in the potential U:

$$\begin{aligned}
\mathcal{H} &= \sum_\alpha S_\alpha U_\alpha \rightarrow < S_\alpha > \cong \int S_\alpha \mathrm{e}^{-\mathcal{H}/\mathrm{k_B}T} \\
\Delta < S_\alpha > &= \sum_\gamma \frac{\partial < S_\alpha >}{\partial U_\gamma} \Delta U_\gamma \\
&= \sum_\gamma \frac{< S_\alpha >< S_\gamma > - < S_\alpha S_\gamma >}{\mathrm{k_B}T} \Delta U_\gamma \\
&\overset{\text{Def}}{=} < S_\alpha > - S_\alpha^{\text{Ref}}
\end{aligned}$$

- Fits selected distribution functions
- Update accounts for cross-correlations

[3] A. Lyubartsev and A. Laaksonen, Phys. Rev. E **52**, 3730-3737 (1995).

Methods

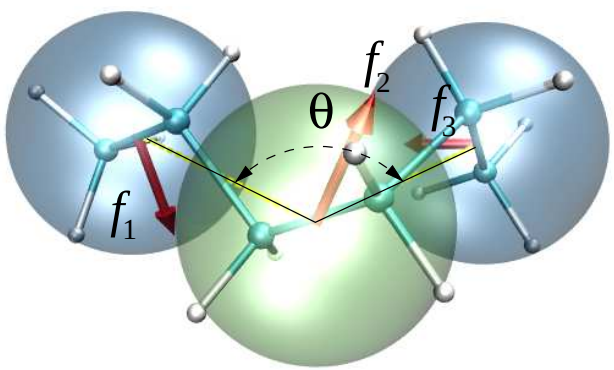
Force matching (Multiscale coarse-graining)

Basic Idea[4]

Minimize:

$$\chi^2 = \left\langle \sum_{I \in CG} |\vec{F}_I^{CG} - \sum_{i \in I} w_i f_i^{Ref}|^2 \right\rangle_{\text{Ensemble}}$$

- Not iterative
- Splines used to represent F_I^{CG} $\rightarrow$ linear least squares fit
- Tries to match multibody PMF (motivated from liquid state theory)



[4] S. Izvekov and G. Voth, J. Chem. Phys. **123**, 134105 (2005).

Votca

Design Concepts

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

Parts of VOTCA[5]- www.votca.org

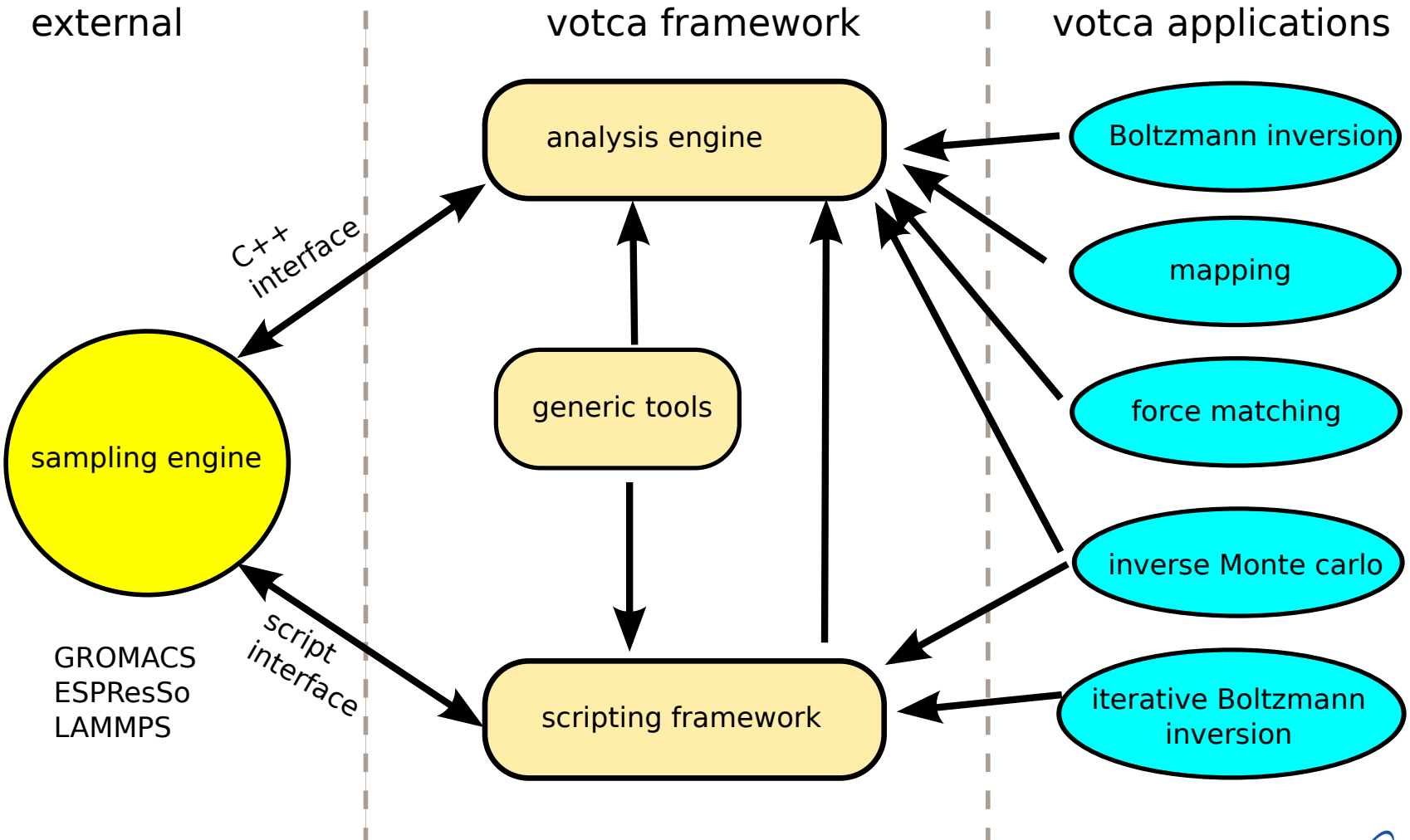
- Mapping engine
- Parallel analysis framework
- Iterative coarse-graining
- Charge transport modules

- Ohloh: 10 Person Years / 39.8k Lines / $ 528.4K
- 15 Developers
- Packages in Fedora, OpenSuse, Gentoo

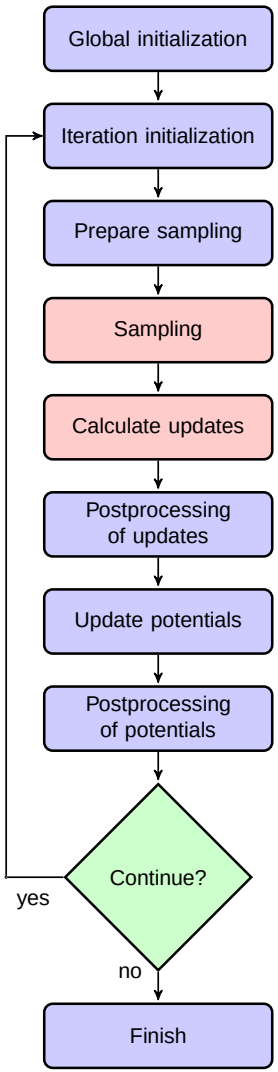
[5] V. Rühle, C. Junghans et al, J. Chem. Theo. Comp. 5 (12), 3211 (2009) & Macromol. Theo. Simul. 20, 472 (2011)

Votca

Votca Components



Votca
Iterative Coarse-Graining



Basic Idea

Iterative process until a desired accuracy:

$$U_{n+1}^{\mathrm{CG}} = U_n^{\mathrm{CG}} + \Delta U(n\text{th CG Simulation})$$

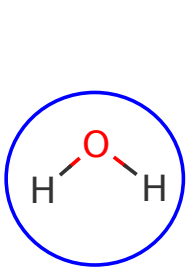
- Automated
- Steering options in XML file
- Customizable (e.g. replace sampling engine)
- Partially parallelized
- Post-processing options:
  - Smoothing
  - Spline fitting
  - Inter/Extrapolation
  - Additional corrections

Los Alamos NATIONAL LABORATORY EST.1943

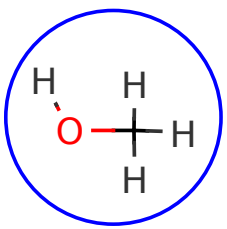
Examples
Overview

Examples from the Paper[6]

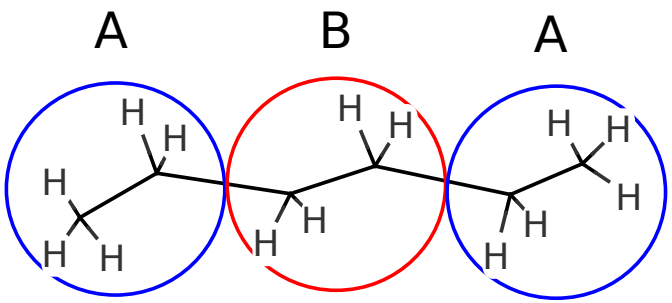
- SPC/E water - good reference system?
- Methanol - finite size effects in IMC
- Hexane - limited basis set
- Propane - multicomponent example



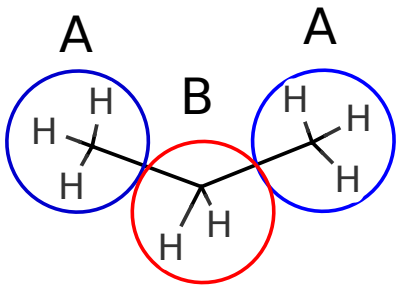
water



methanol



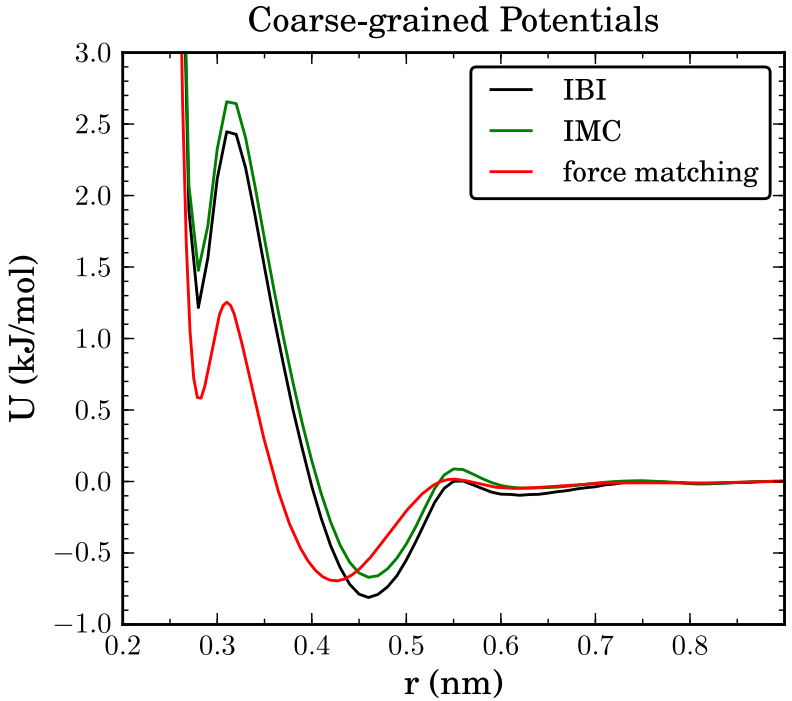
hexane

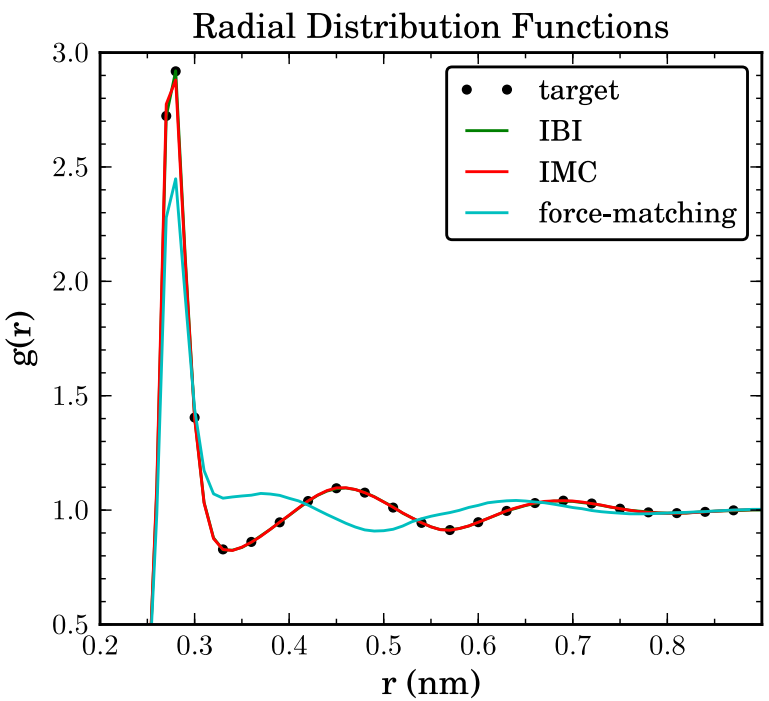


propane

[6] V. Rühle, C. Junghans et al, J. Chem. Theo. Comp. 5 (12), 3211 (2009)

Examples
SPC/E water

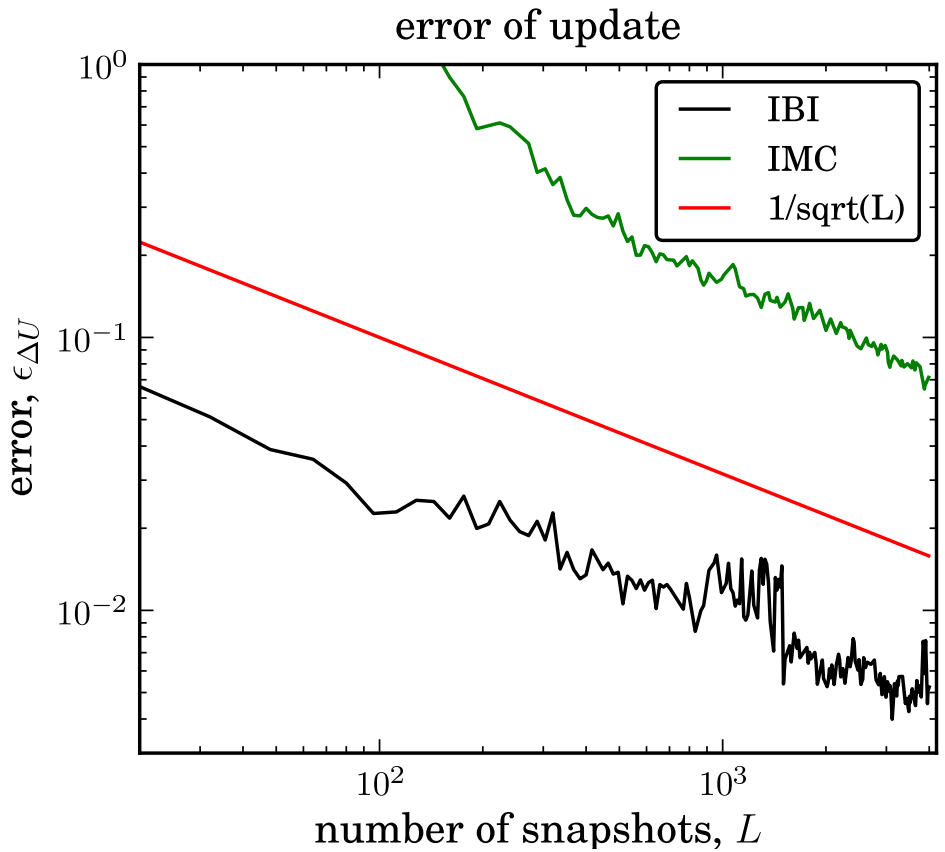




- IBI and IMC have similar potentials and both reproduce the radial distribution function
- Force-matching potential is similar, but RDF is wrong
  $\rightarrow$ pair potential not a good approximation

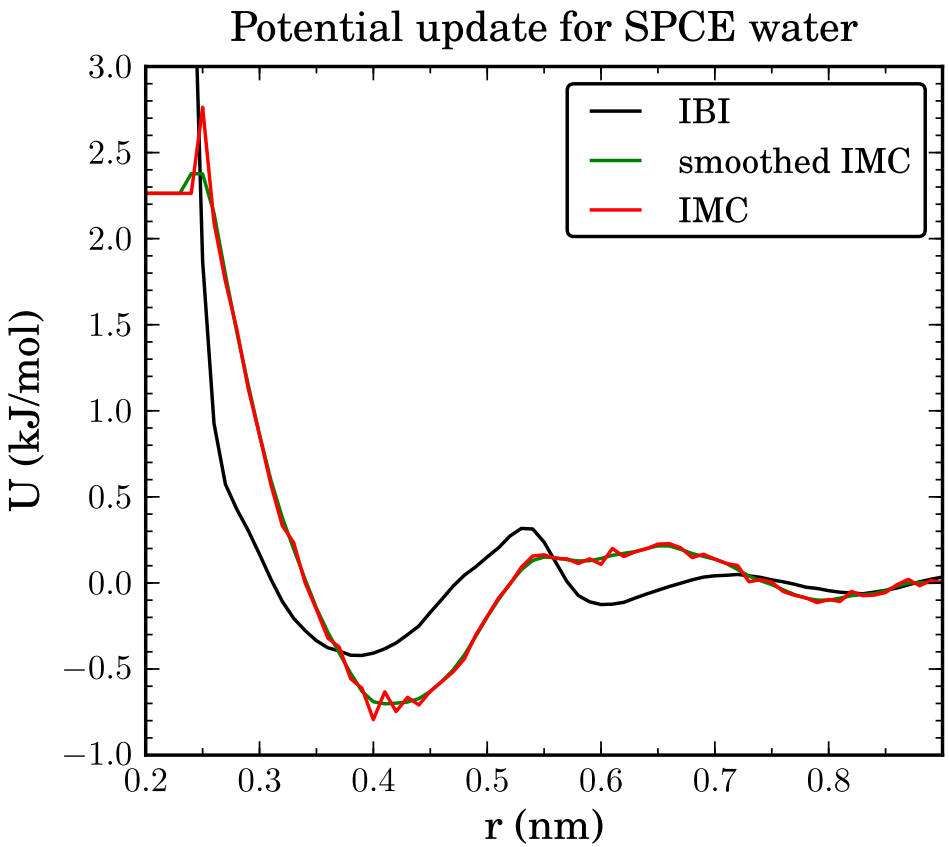
Los Alamos NATIONAL LABORATORY EST.1943

Examples
SPC/E water



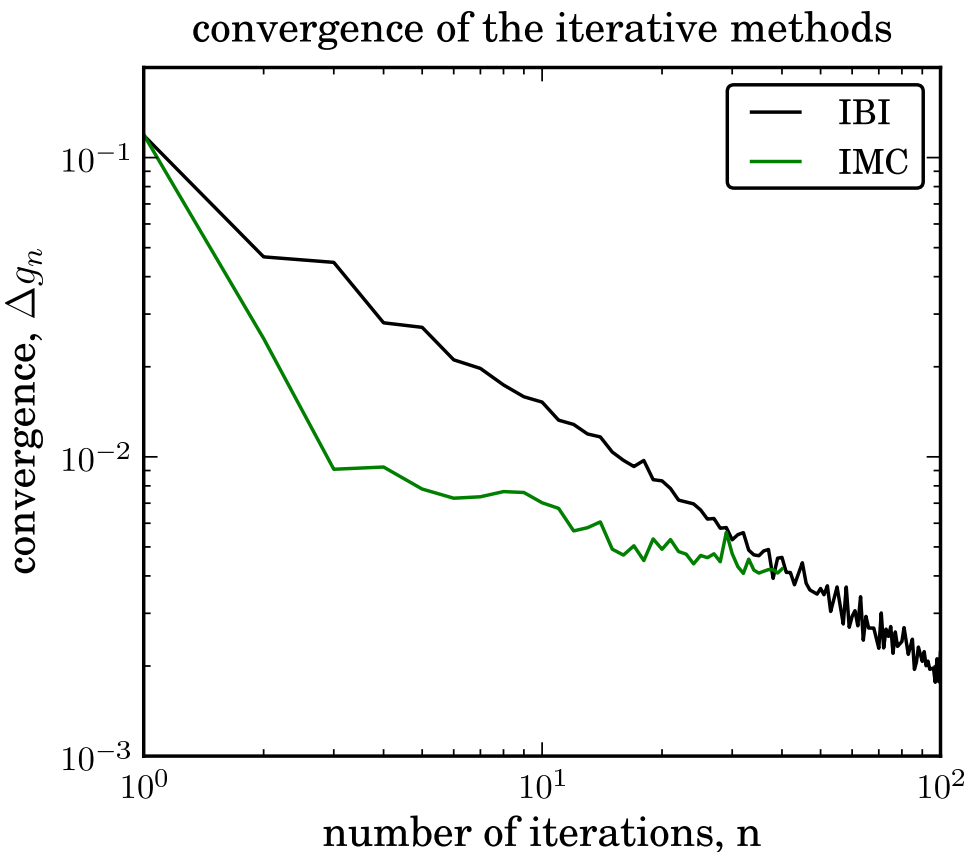
- Errors calculated using jackknife analysis
- IMC needs more data to calculate update
- Smoothing applied to IMC update

Examples
SPC/E water



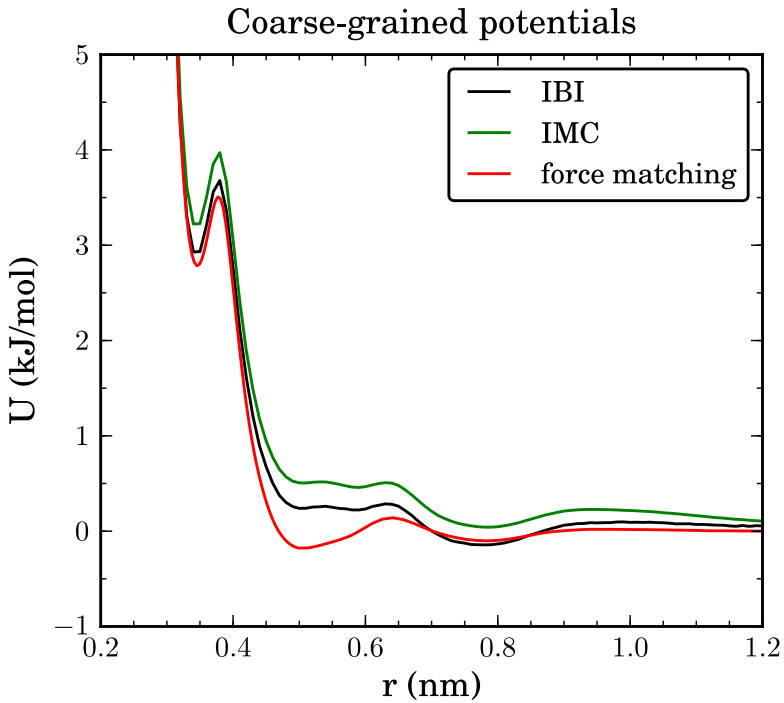
- Potential update has less sharp features than potential or RDF
  $\rightarrow$ smoothing of ΔU instead of U
- Only a few snapshots required for IBI if smoothing is applied

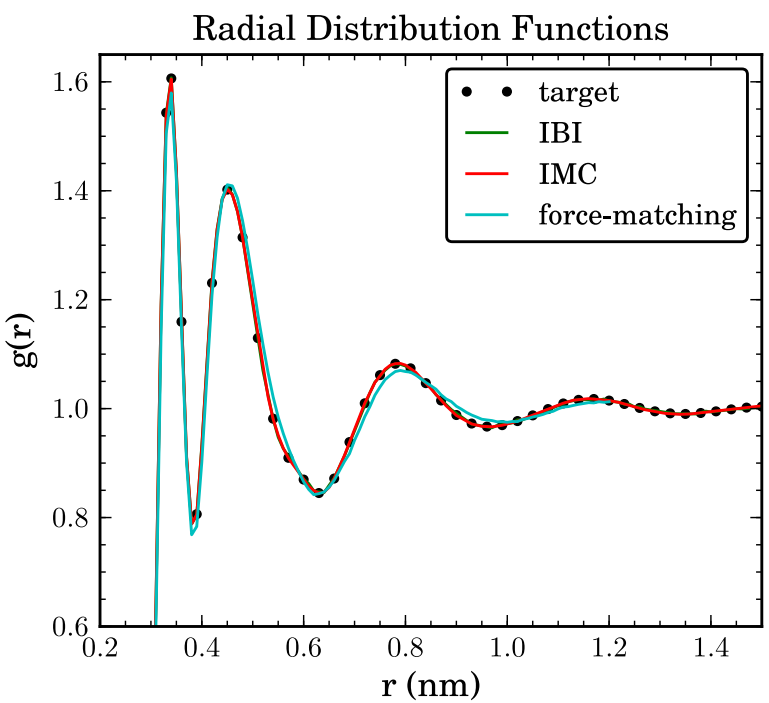
Examples
SPC/E water



- IMC converges faster than IBI
- Convergence of IMC reaches a plateau
- Computational costs for both runs are comparable

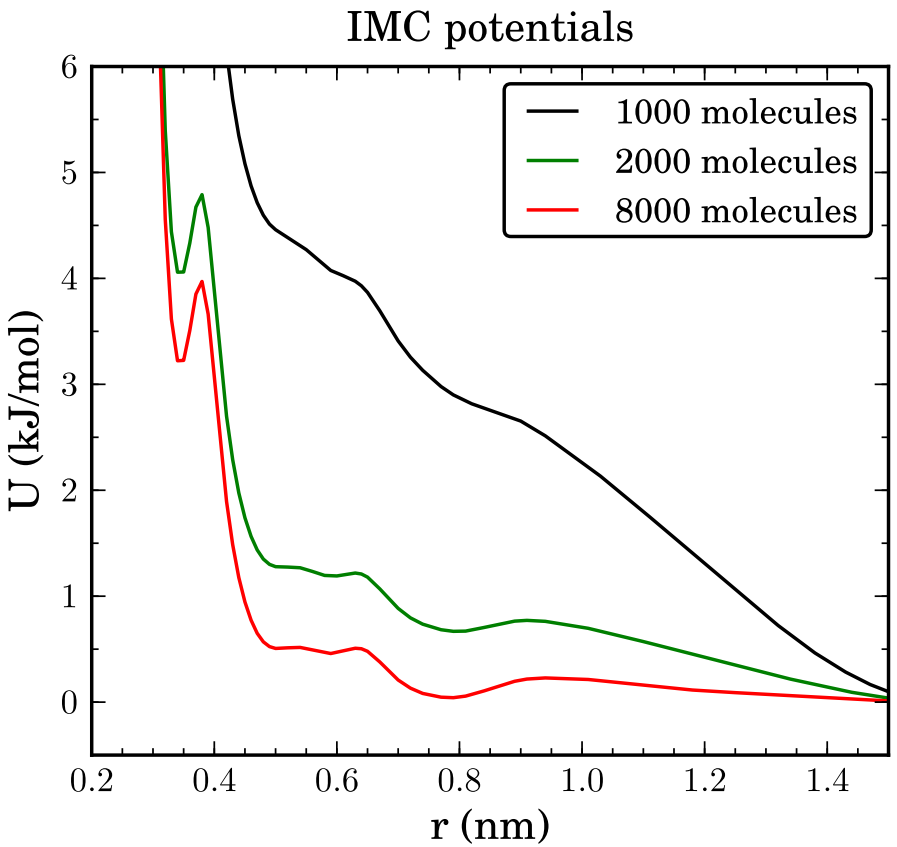
Examples
Liquid Methanol





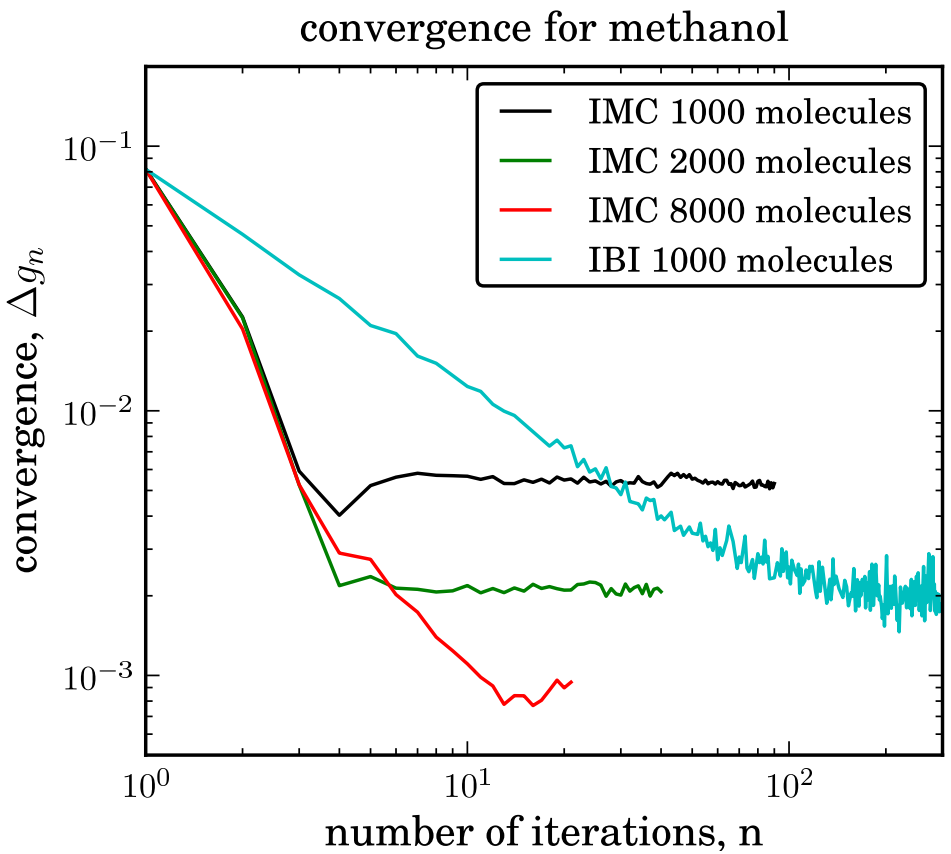
- All methods reproduce the radial distribution function
- Force- and structure-based methods give similar potentials

Examples
Liquid Methanol



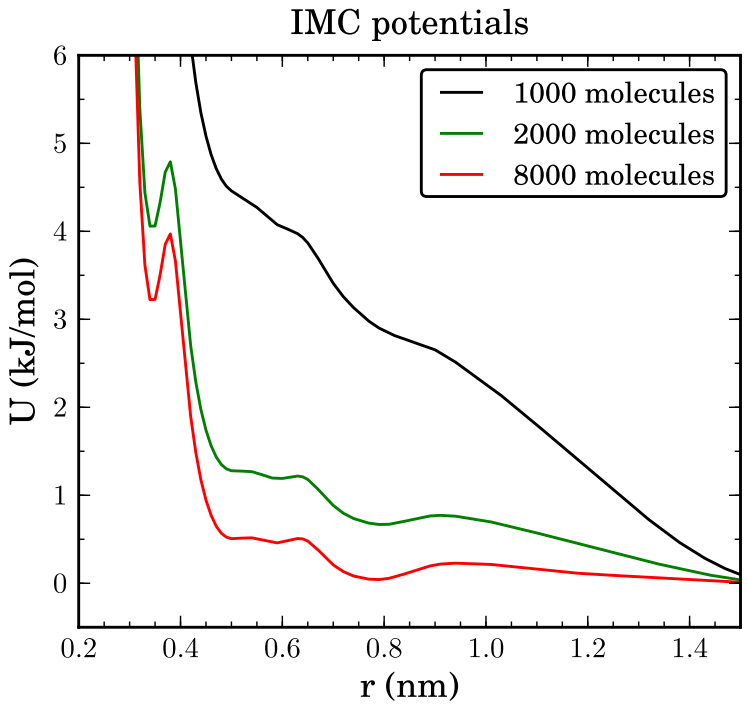
- IMC update sensitive to system size
- Need for bigger systems
- Problem arises due to cross-correlations

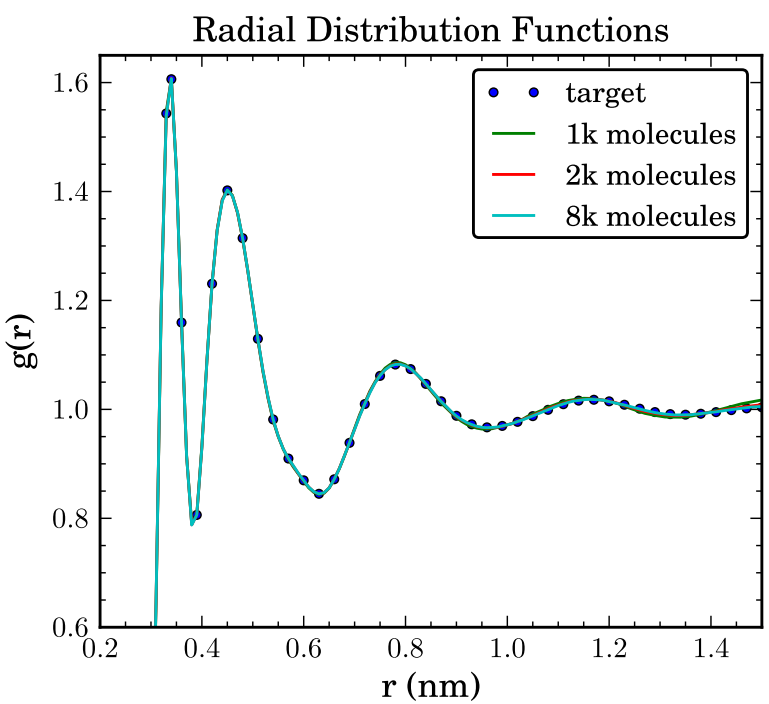
Examples
Liquid Methanol



- Small systems + IMC results in deviations of RDF at large distances
- Bigger system have increased computational costs

Examples
Liquid Methanol

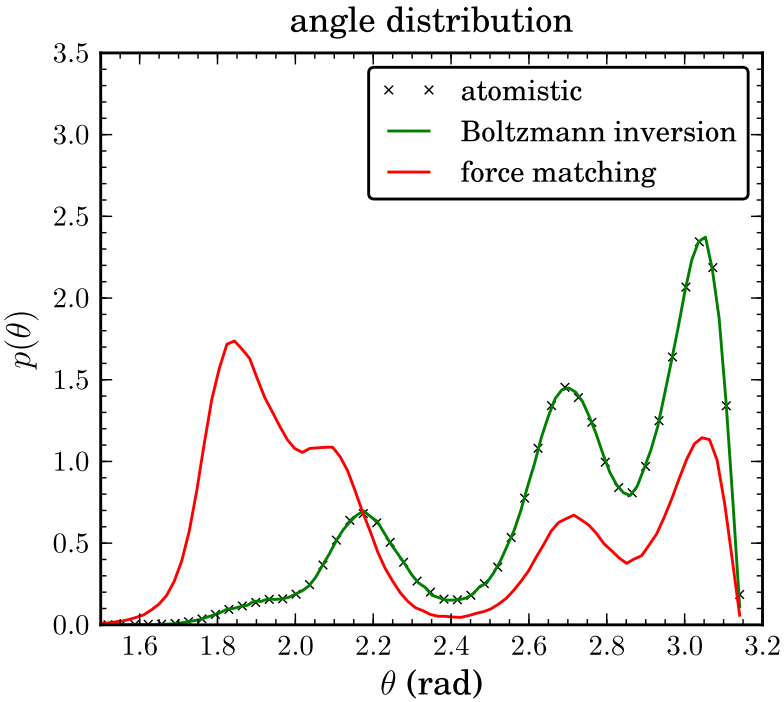


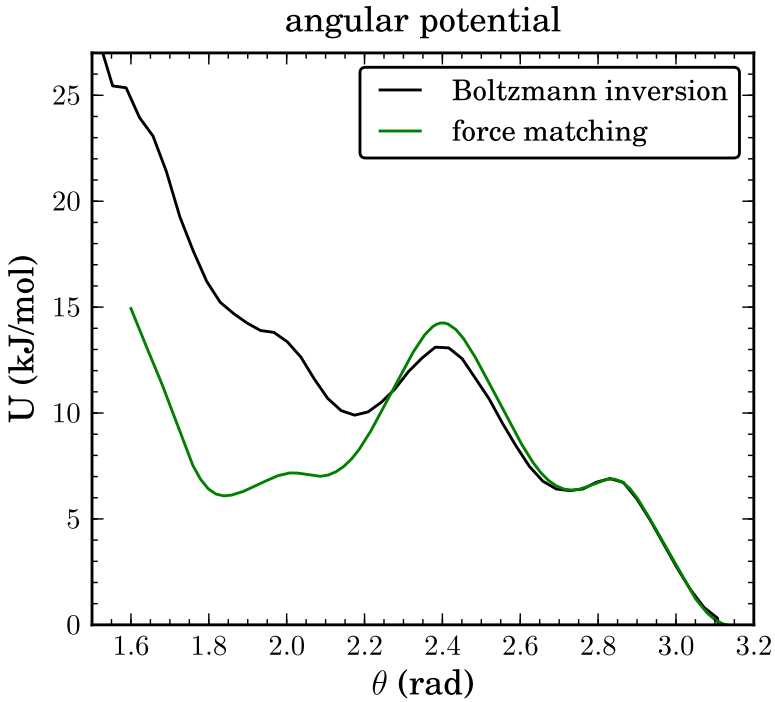


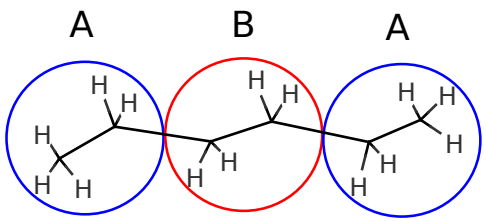
- Completely different potentials give similar RDFs
- Is the RDF a good measure to parameterize for?

Examples

Hexane - Single molecule in vacuum





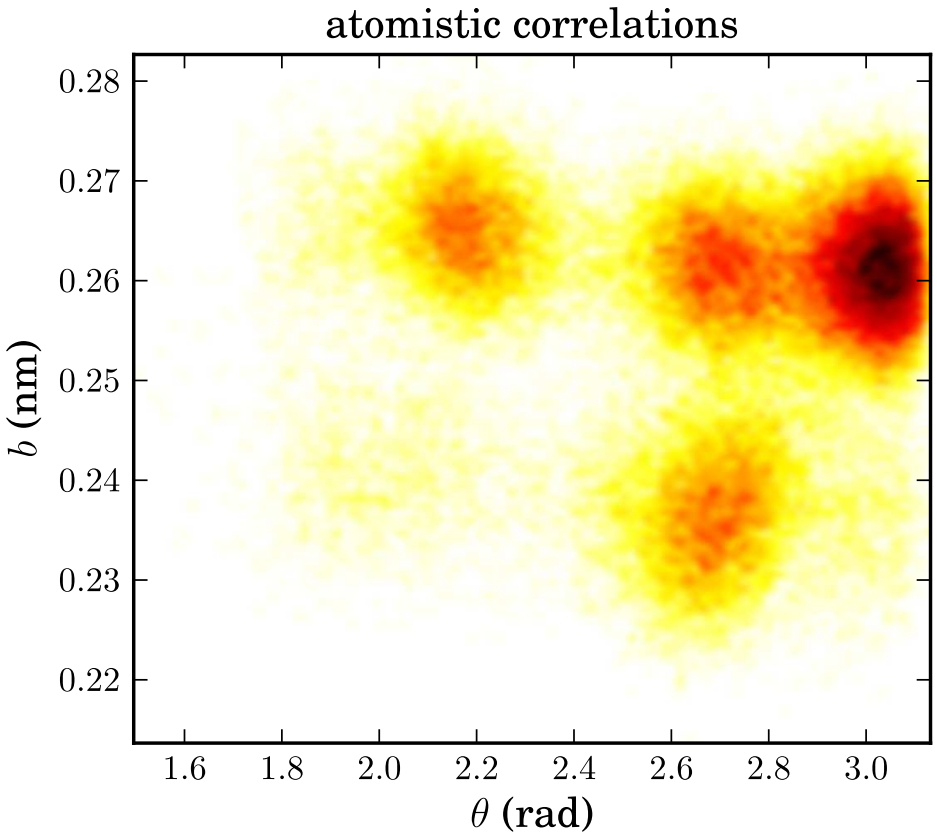


- Single molecule of hexane
- Force-matching does not reproduce angle distribution $\rightarrow$ hybrid approaches[7]

[7] Macromol. Theo. Simul. 20, 472 (2011)

Examples

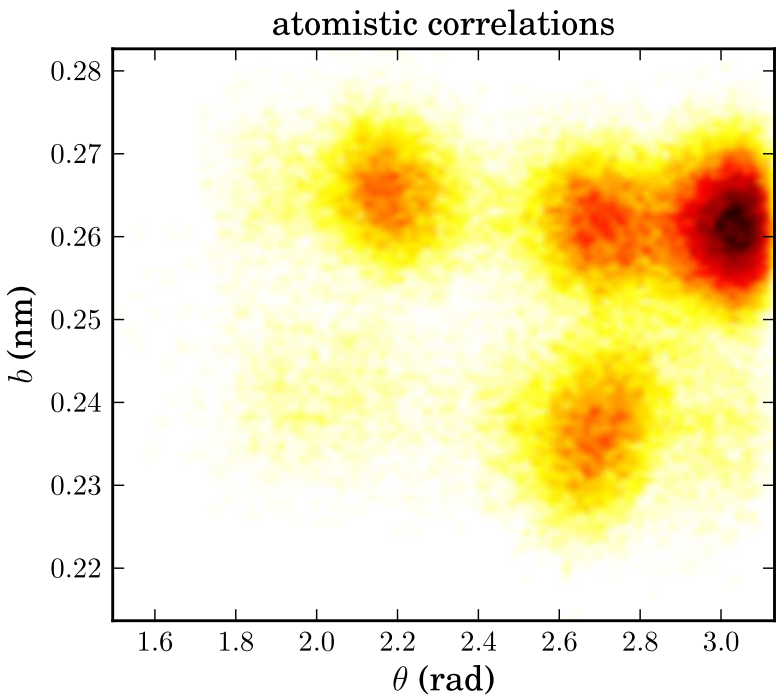
Hexane - Single molecule in vacuum

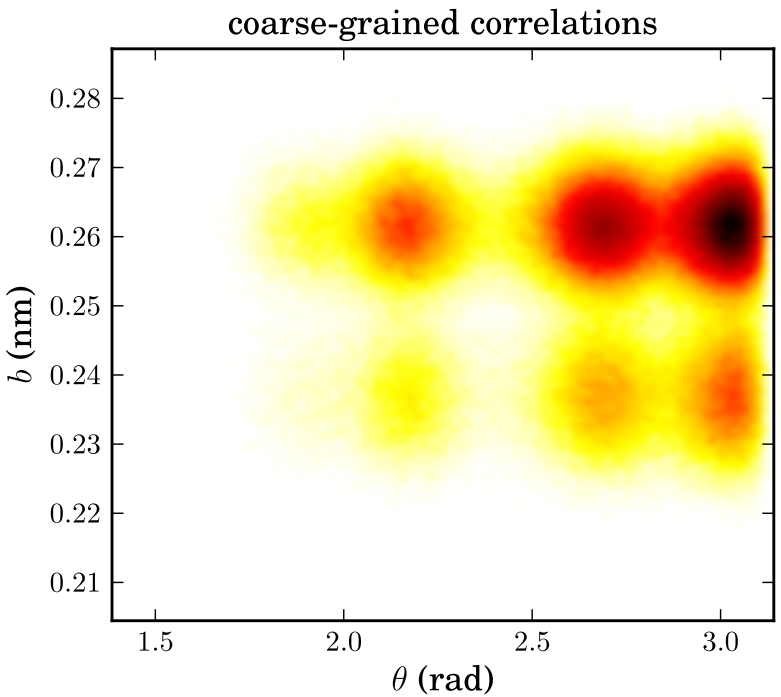


- Bonds + angles are correlated
- Correlations wrong with the given IBI potentials
- Additional cross-coupling term needed

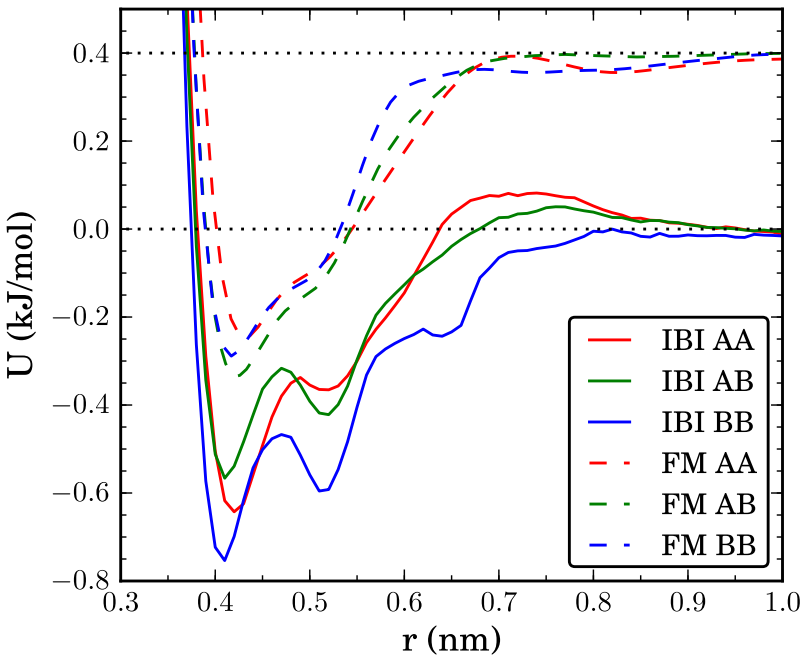
Examples

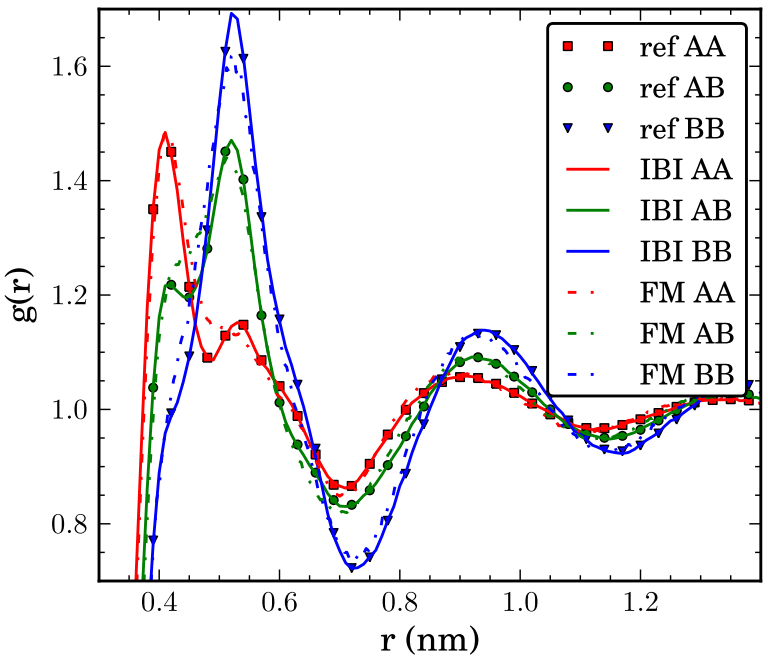
Hexane - Single molecule in vacuum





Examples
Liquid Propane





- Multi component example (inner and outer beads)
- Bonded interaction by Boltzmann inversion
- Different update schemes possible

Conclusion

Be careful[8]:

- CG models never reproduce all properties
- Correlations in interactions
- System size
- Completeness of the CG force-field
- CG force-fields parameterized at a specific state point
- Dynamics

[8] V. Rühle, C. Junghans et al, J. Chem. Theo. Comp. 5 (12), 3211 (2009)

Conclusion

Conclusion

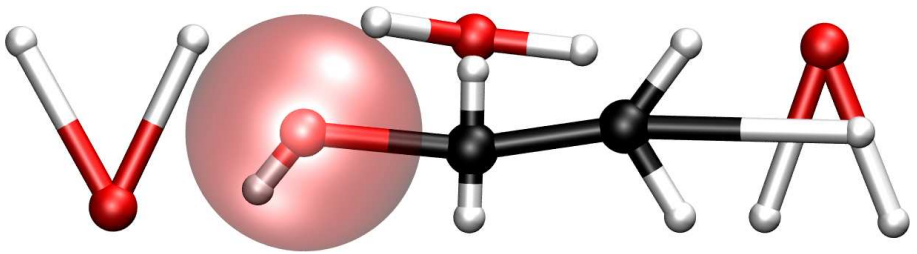
Features to come in Votca 1.3:

- IBI for bonded interactions ✓
- Relative Entropy method (✓)
- Optimization of potential with functional form ✓
- More optimizers and targets properties (✓)
- Kirkwood-Buff correction ✓

Conclusion
Team

Project supervisor
Denis Andrienko
Kurt Kremer

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

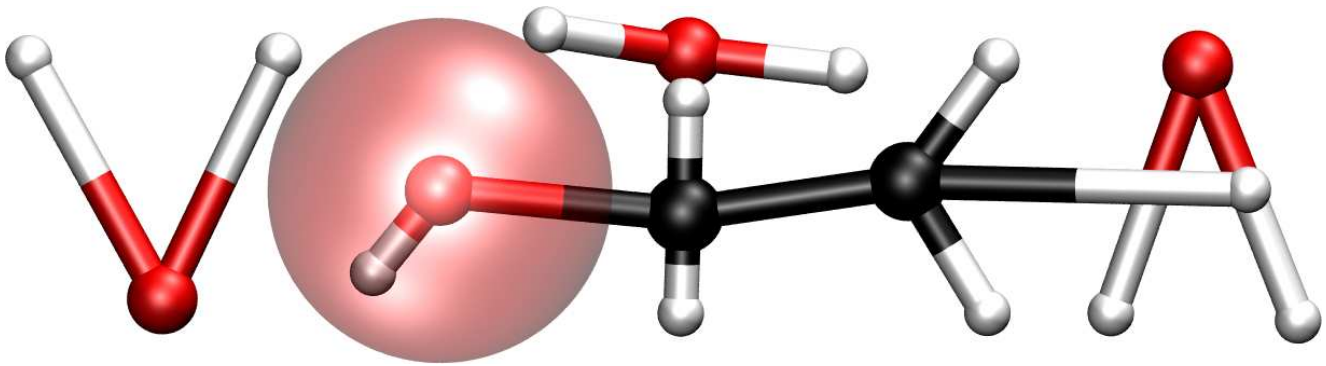
Tristan Bereau	interface to ESPResSo
Sebastian Fritsch	interface to AdResS
Mara Jochum	SIMPLEX algorithm
Konstantin Koschke	parallel analysis engine
Alexander Lukyanov	force-matching
Sikandar Mashayak	relative entropy method

Los Alamos NATIONAL LABORATORY EST.1943

Conclusion
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

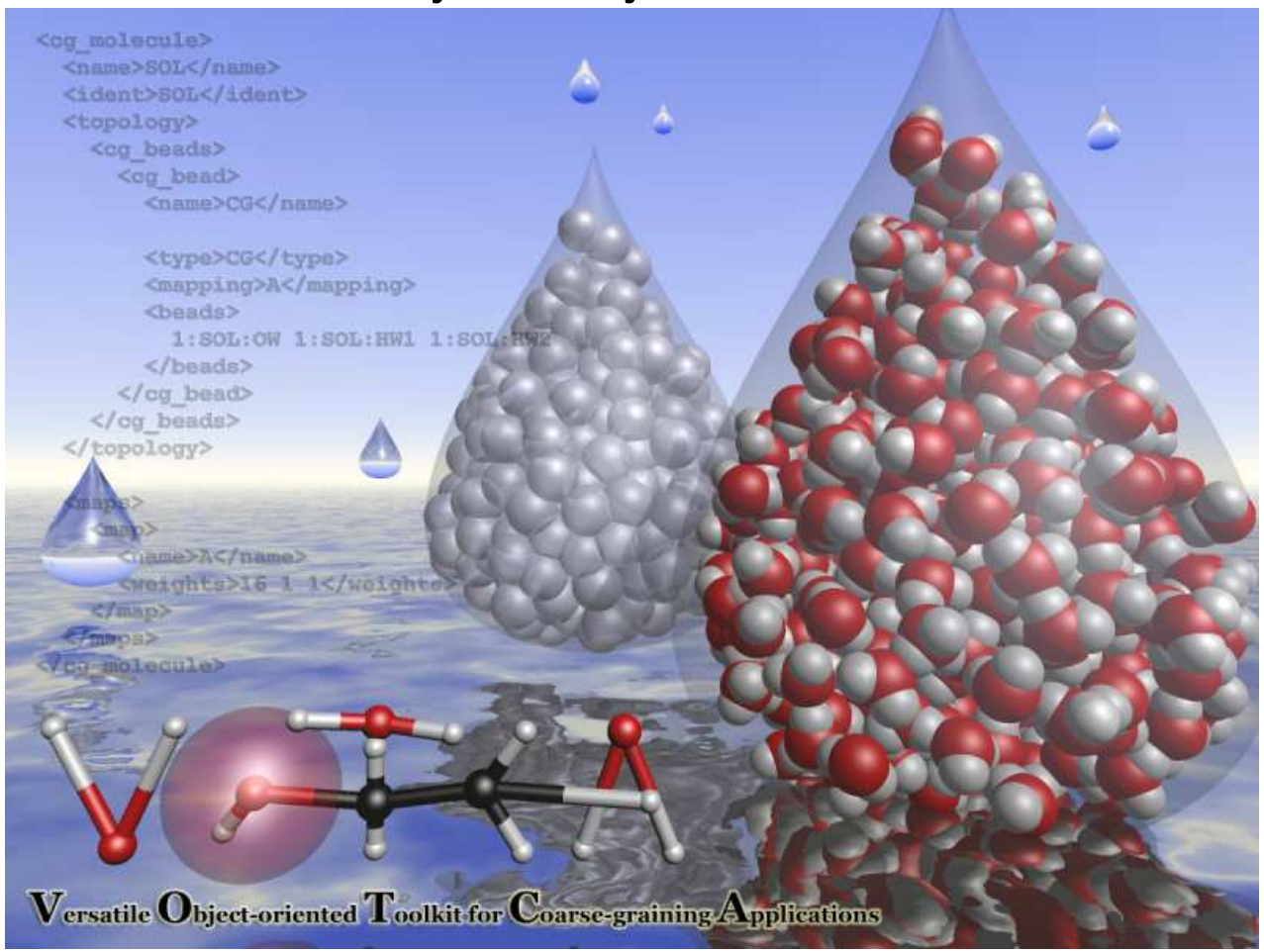
- It's free
- All examples are in the tutorial
- It's flexible and expandable

Visit us at www.votca.org

Los Alamos NATIONAL LABORATORY EST.1943

The End

Thank you for your attention !



Visit us at www.votca.org