Multicriteria optimization of molecular force field models

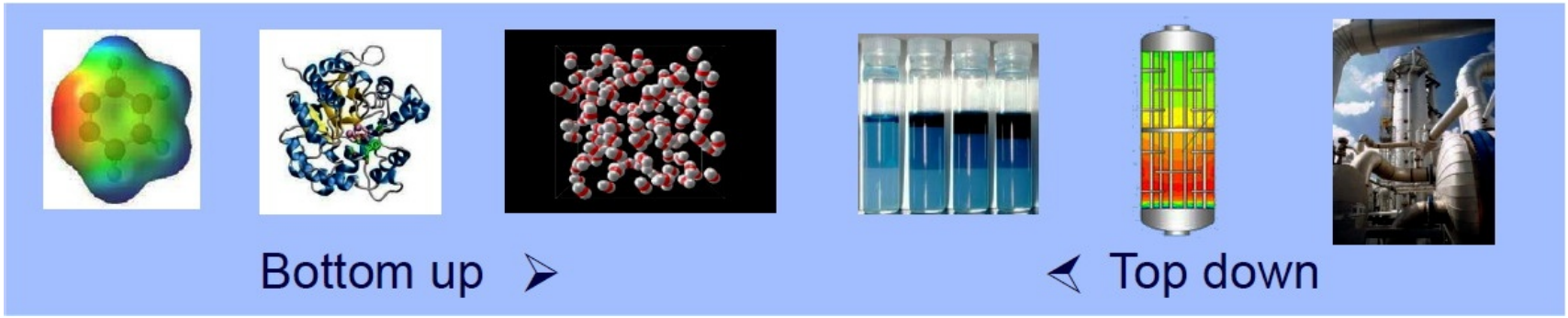
Martin Horsch,[1] Katrin Stöbener,[1, 2] Stephan Werth,[1] and Hans Hasse[1]

[1] Laboratory of Engineering Thermodynamics, University of Kaiserslautern
[2] Fraunhofer Institute for Industrial Mathematics, Kaiserslautern

Leipzig, 27th November 2015
NTZ CompPhys15 Workshop

Computational Molecular Engineering

Computational molecular engineering



From Physics (qualitative accuracy)

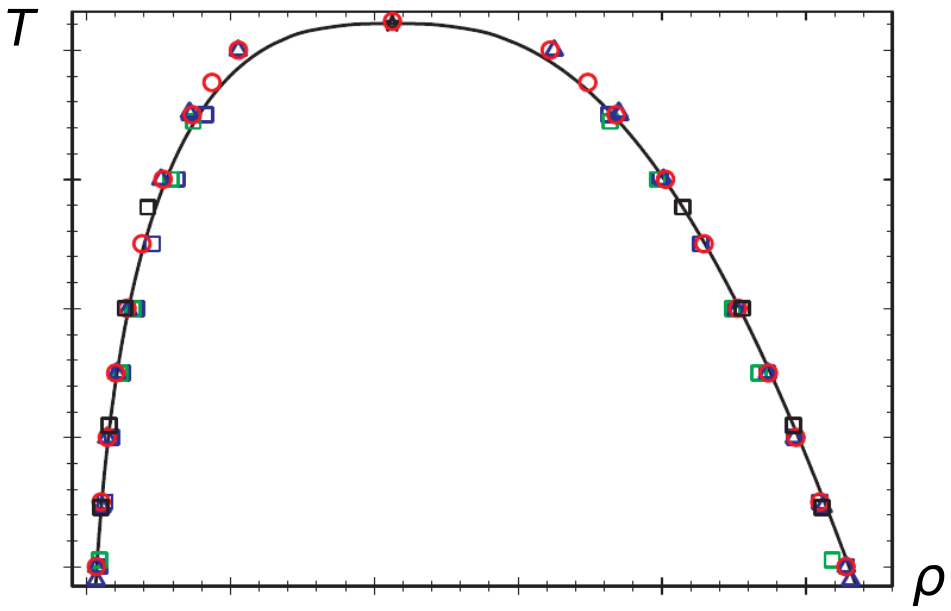
- Physically realistic modelling of intermolecular interactions
- Separate contributions due to repulsive and dispersive as well as electrostatic interactions

To Engineering (quantitative reliability)

- No blind fitting, but parameters of *effective pair potentials* are adjusted to experimental data
- Physical realism facilitates reliable interpolation and extrapolation

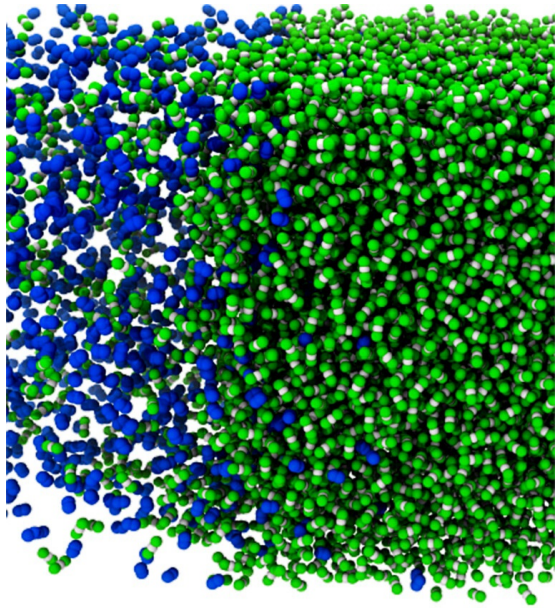
Vapour-liquid equilibria

Bulk properties



vapour pressure, saturated densities, composition, enthalpy of vaporization, etc., by Grand Equilibrium simulation

Interfacial properties

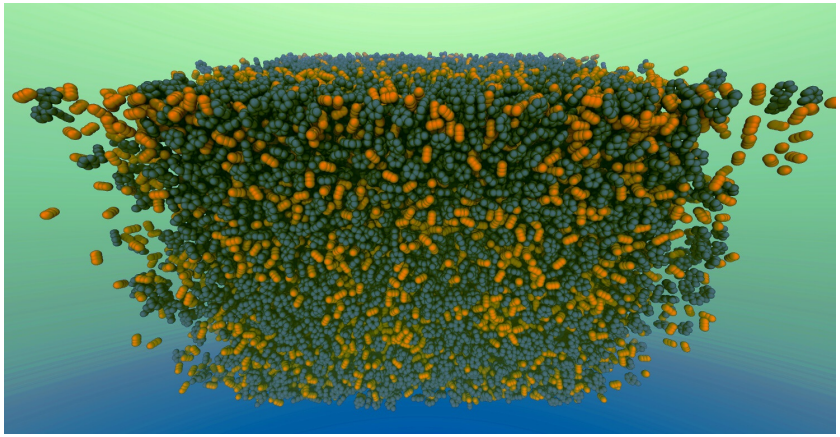


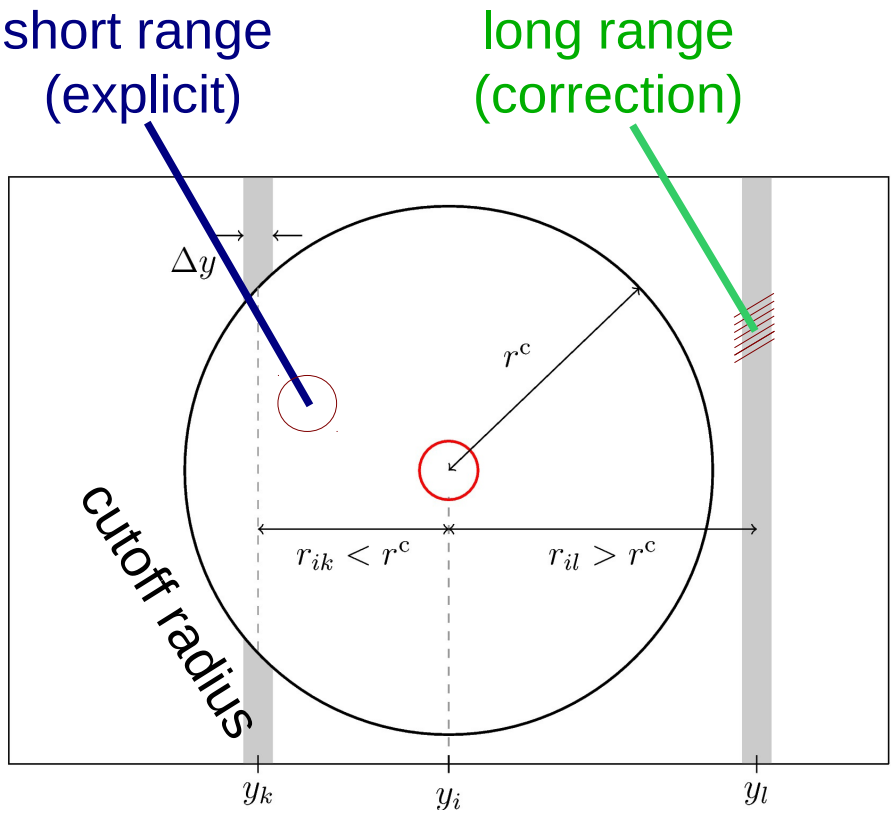
heterogeneous systems with finite-size effects and long-range interactions

Long-range correction at planar interfaces

For planar interfaces:

Long-range correction from the density profile, following **Janeček**.



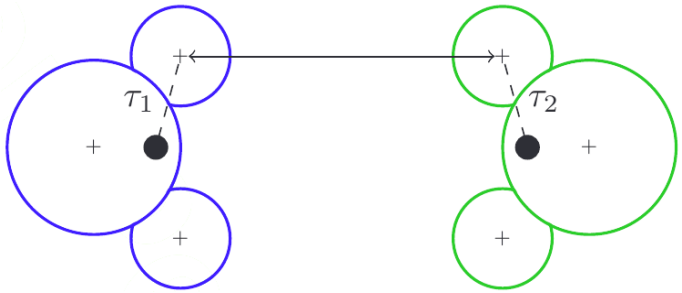


Full evaluation of all pairwise interactions is too expensive ...
... **short-range interactions** are evaluated only for **neighbours**.

Long-range correction at planar interfaces

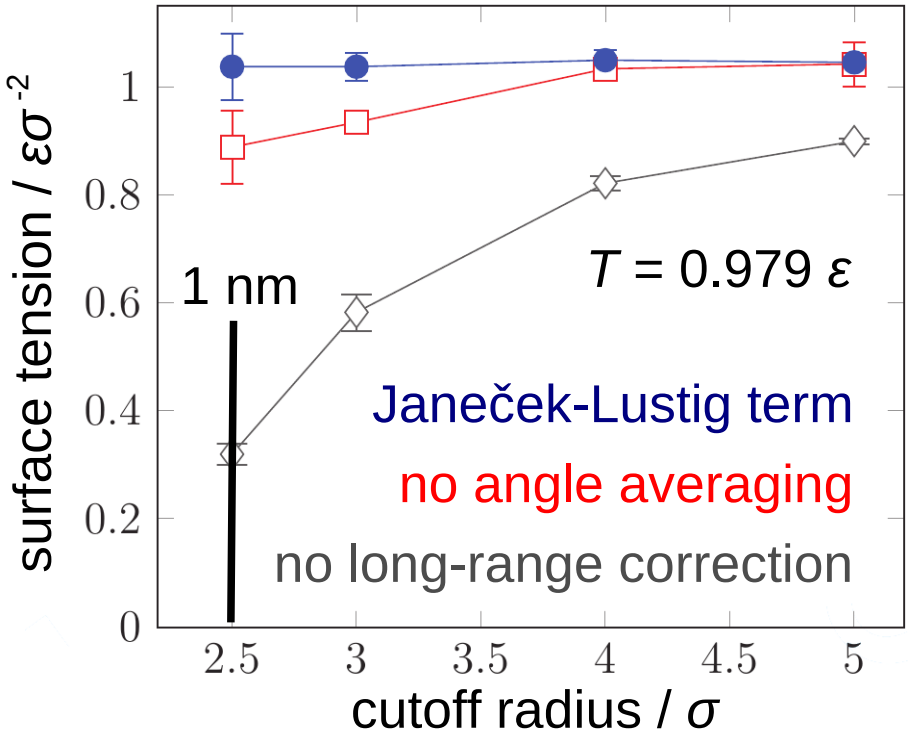
For planar interfaces:

Long-range correction from the density profile, following **Janeček**.



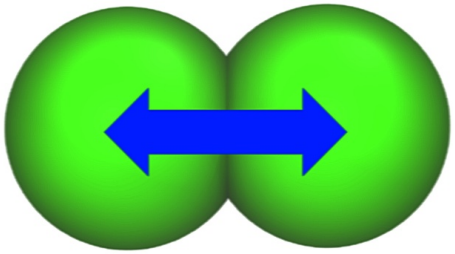
Angle-averaging expression for multi-site models, following **Lustig**.

Two-centre LJ fluid (2CLJ)



Dipole and dispersion lead to analogous long-range correction expressions.
The long-range contribution of the quadrupole can be neglected.

Validation of molecular force field models



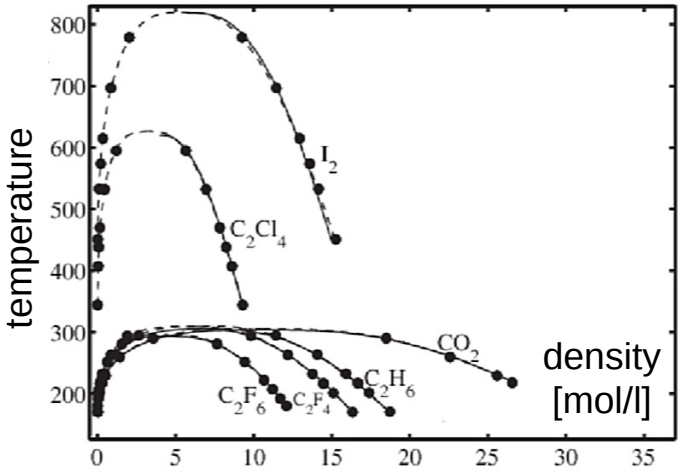
2CLJQ models:

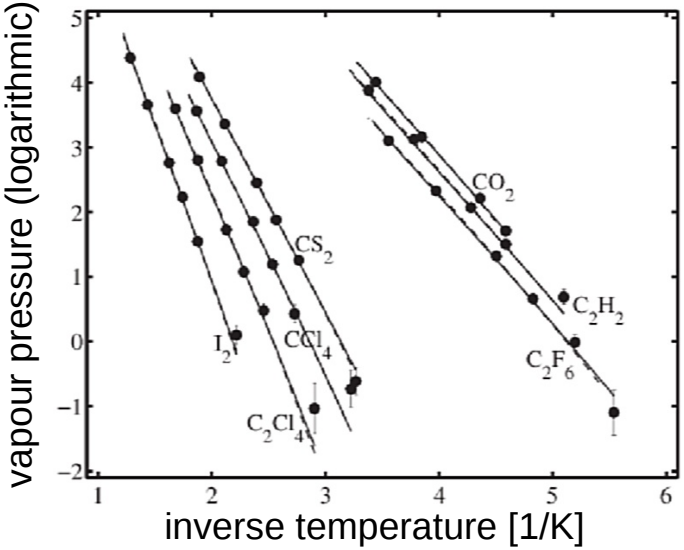
- 2 LJ centres
- Quadrupole

Fit of parameters σ, ε, L, Q to VLE data of 29 fluids by Stoll *et al.*

Deviation:

- $\delta\rho' \approx 1$ %
- $\delta P^{sat} \approx 5$ %



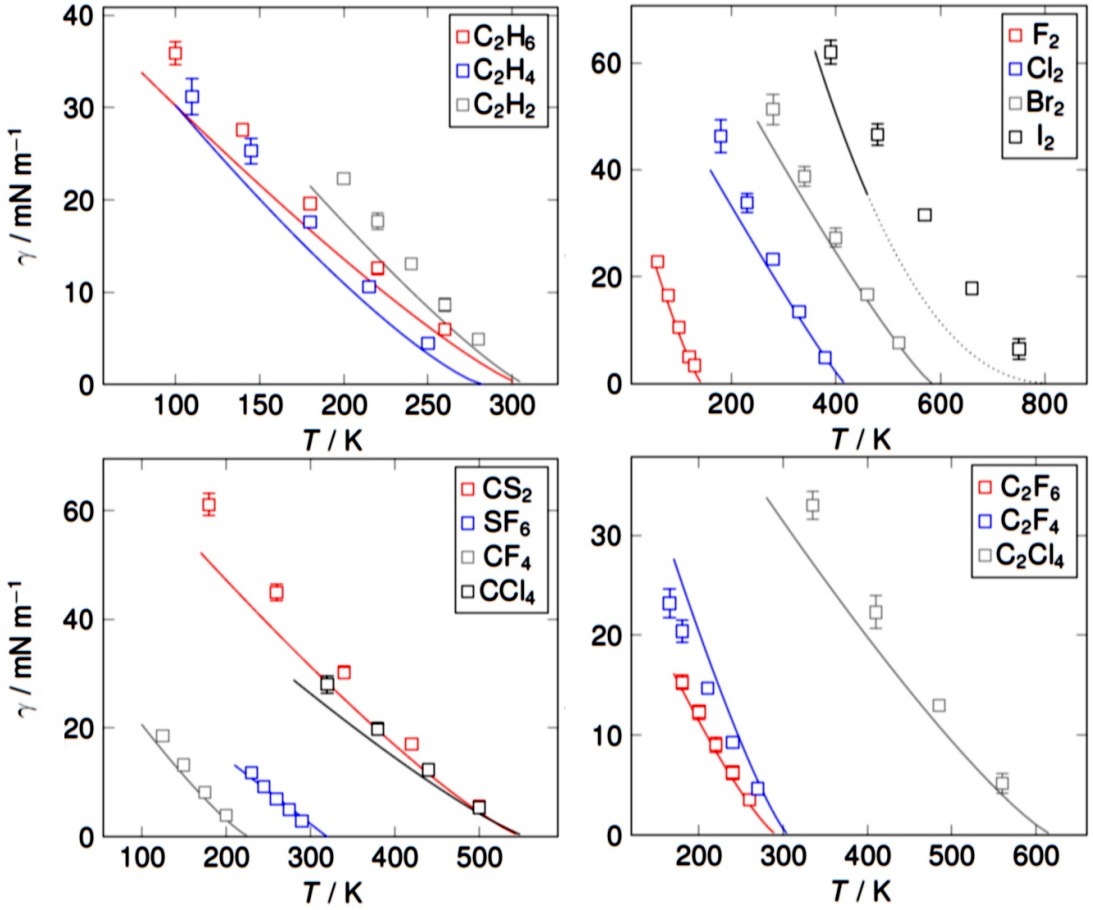


- simulation
— DIPPR correlation

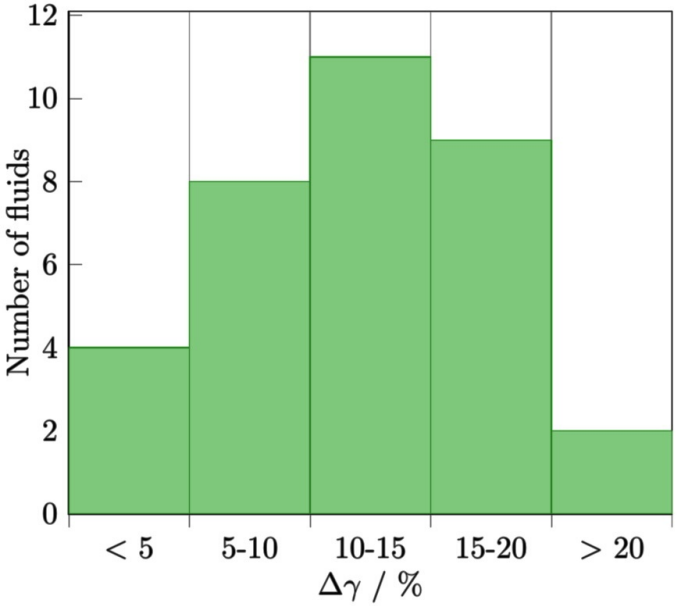
No interfacial properties were considered for the parameterization.

Validation of molecular force field models

Two LJ + quadrupole (2CLJQ)



Two LJ + dipole (2CLJD)

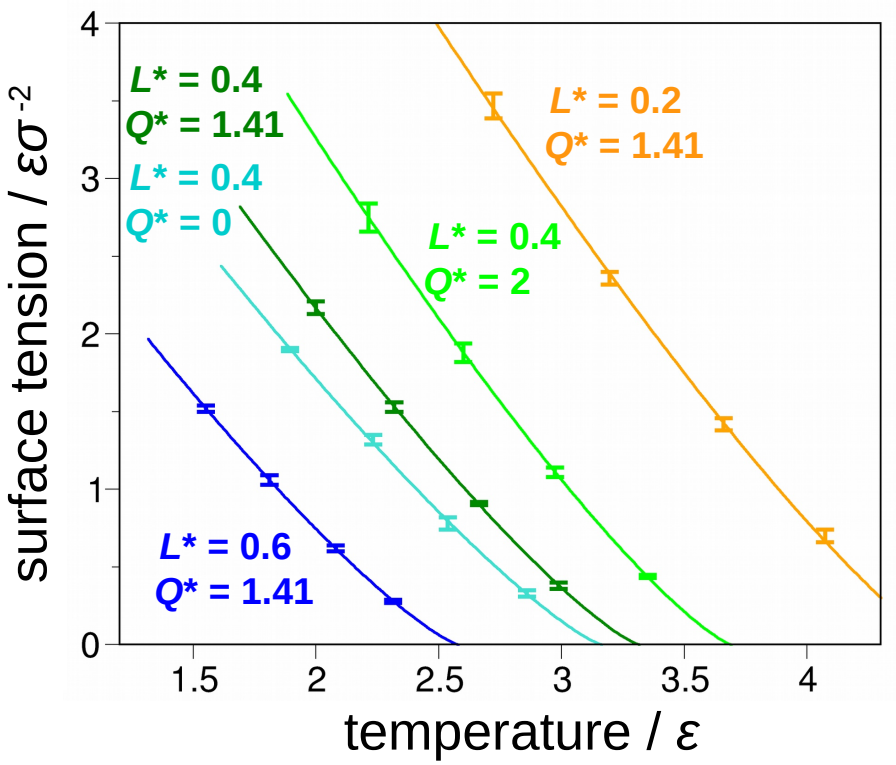


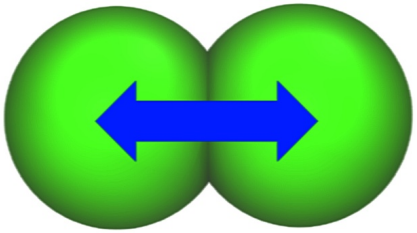
Fit to bulk properties

10 to 20 % overestimation of vapour-liquid surface tension

Massively parallel molecular modelling

Two LJ + quadrupole (2CLJQ)





2CLJQ

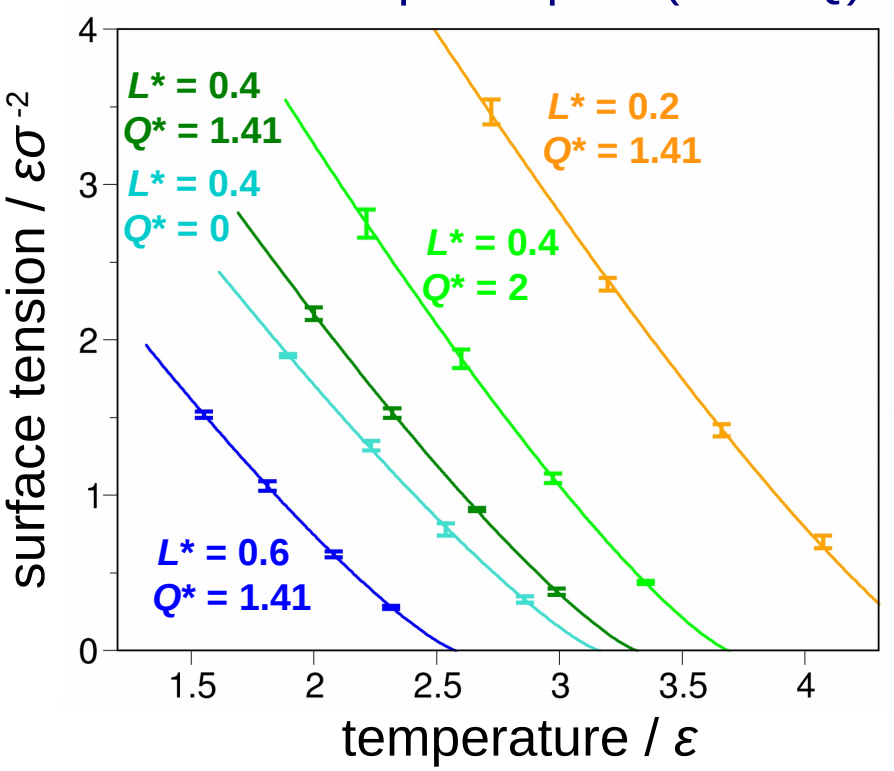
Model parameters:

- LJ size parameter σ
- LJ energy parameter ε
- Elongation L
- Quadrupole moment Q

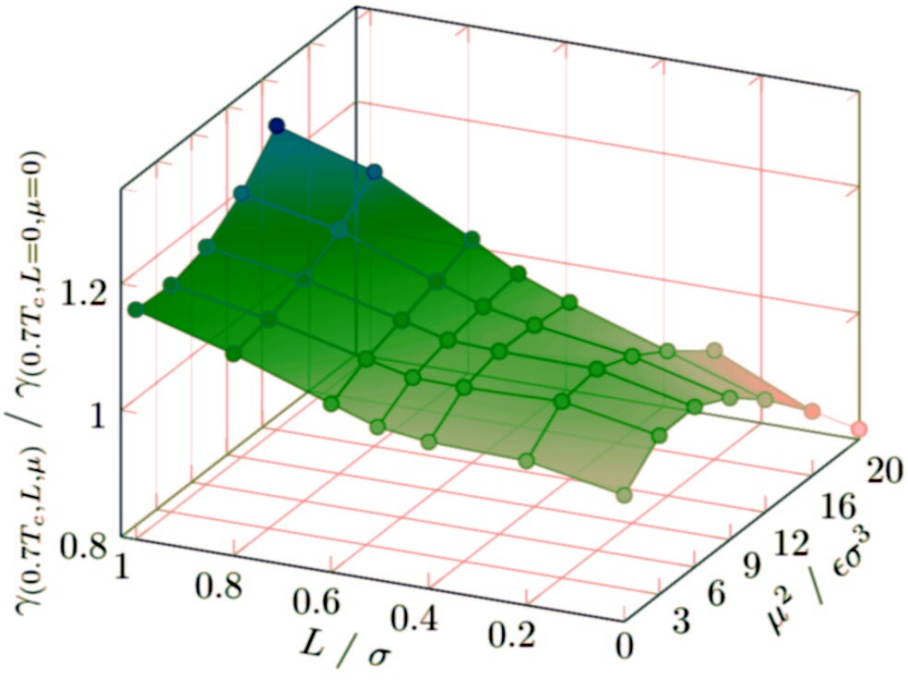
- Systematic exploration of the four-dimensional model parameter space

Massively parallel molecular modelling

Two LJ + quadrupole (2CLJQ)



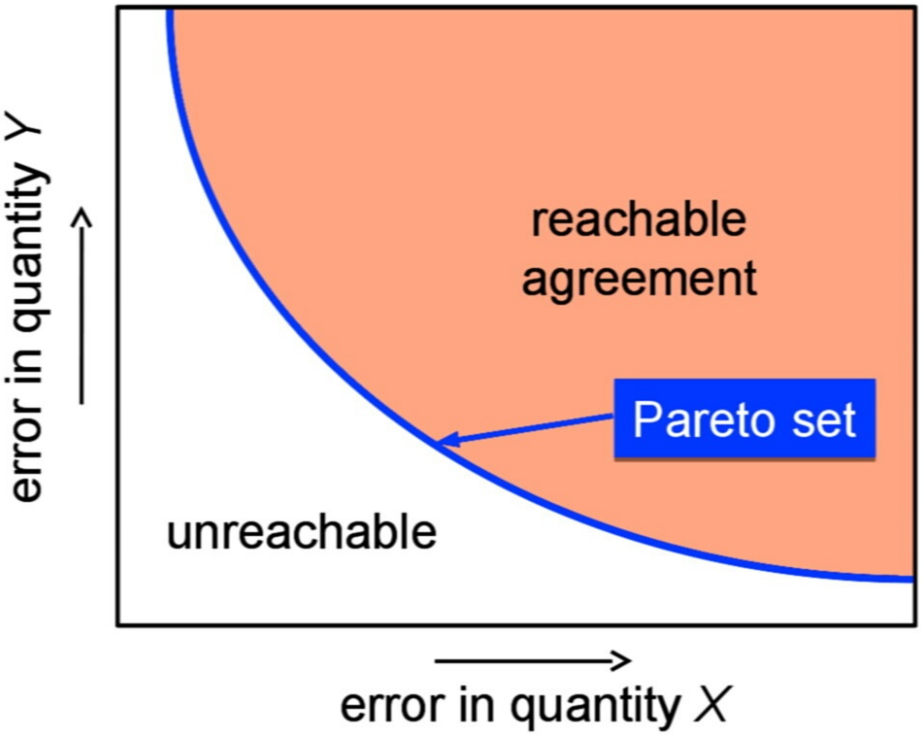
Two LJ + dipole (2CLJD)



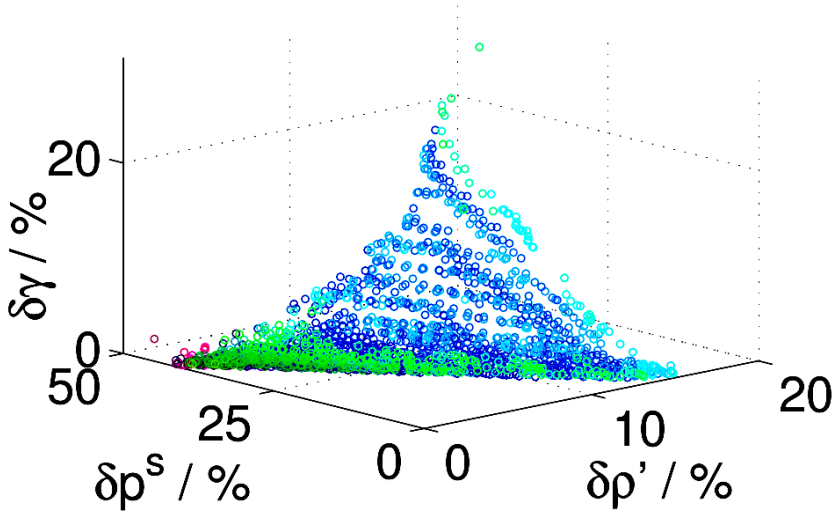
- Systematic exploration of the four-dimensional model parameter space
- Correlation of the surface tension by a critical scaling expression

Multicriteria model optimization

Pareto optimality criterion



Multiple objectives



(2CLJQ for carbon dioxide)

Multicriteria optimization requires massively parallel molecular modelling.

Computation of the Pareto set

Multicriteria optimization problem

Simultaneously minimized objective functions f_ξ with $\xi \in \{\rho', p^s, \gamma\}$ given by

$$f_\xi = \left\langle \delta \xi^2 \right\rangle_{0.55 T_c^{exp} < T < 0.95 T_c^{exp}} = \lim_{N \to \infty} \frac{1}{N+1} \sum_{i=0}^{N} \left(1 - \frac{\xi^{sim}(T)}{\xi^{exp}(T)} \right)^2_{T/T_c = 0.55 + 0.4 i/N} \quad \text{(here: } N = 9\text{)}.$$

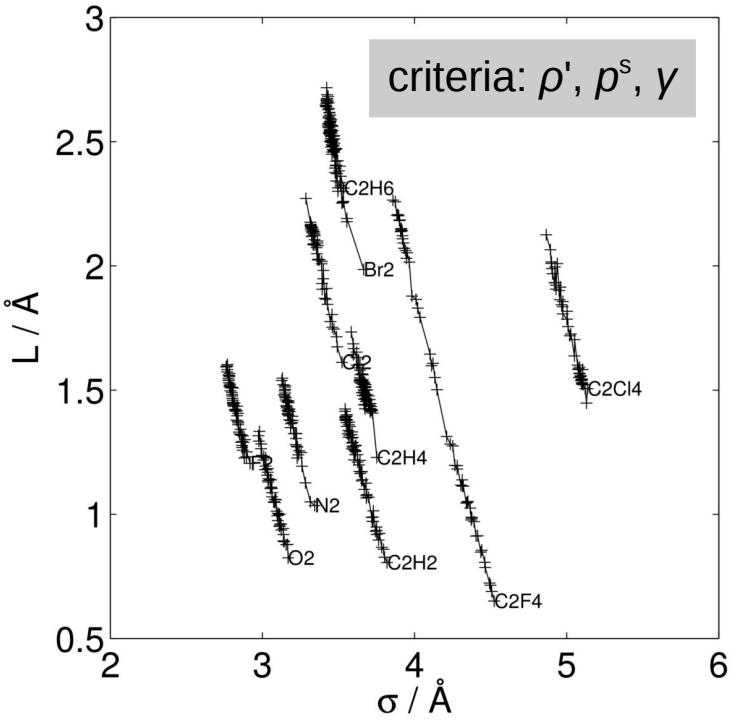
Sandwiching

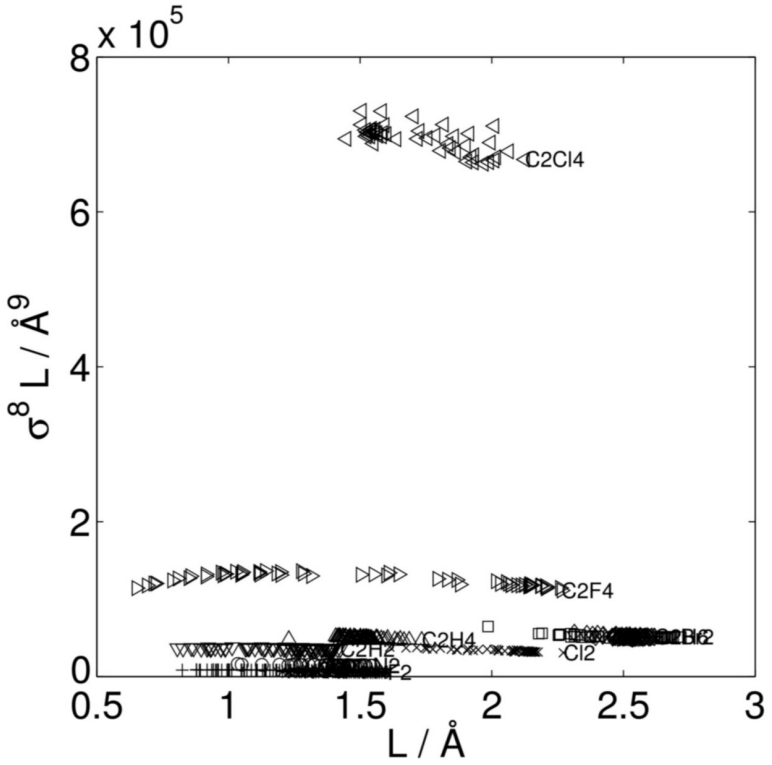
Alternating construction of inner (reachable) and outer (unreachable) approximations, assuming *local convexity* of the Pareto set.

Hyperboxing

In non-convex regions (“hyperboxes”), Pascoletti-Serafini scalarization is employed to obtain a suitable local single-criterion optimization problem.

Invariants of Pareto-optimal models

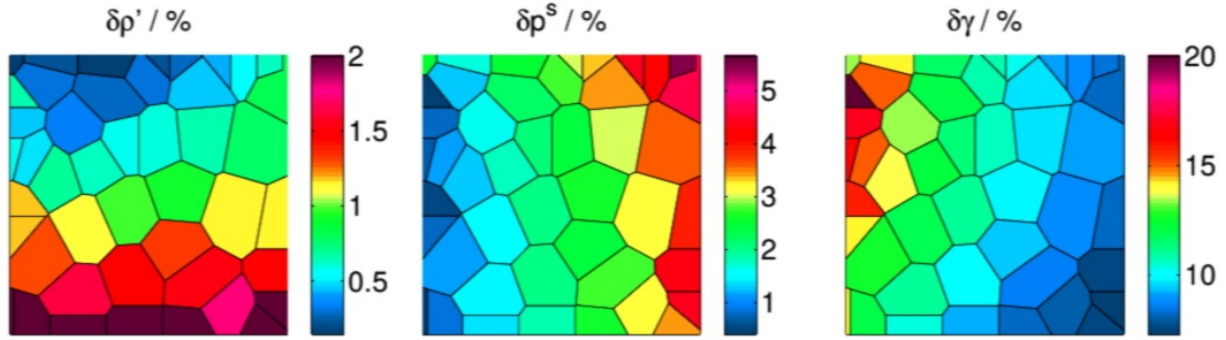




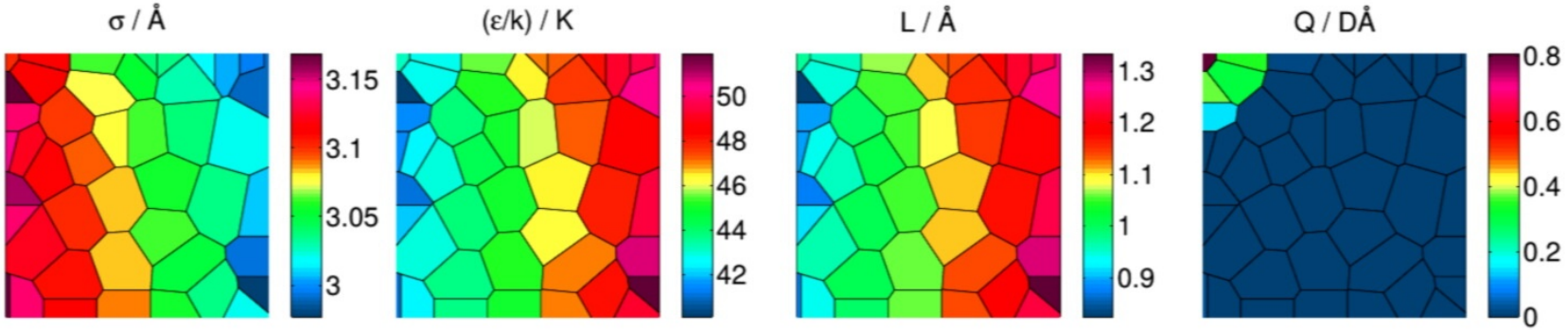
For obtaining a rough approximation of the Pareto set, the dimension of the parameter space can be reduced from four to three (or even two).

Pareto sets for 2CLJQ models of real fluids

Representation of objective and parameter spaces by **patch plots**:

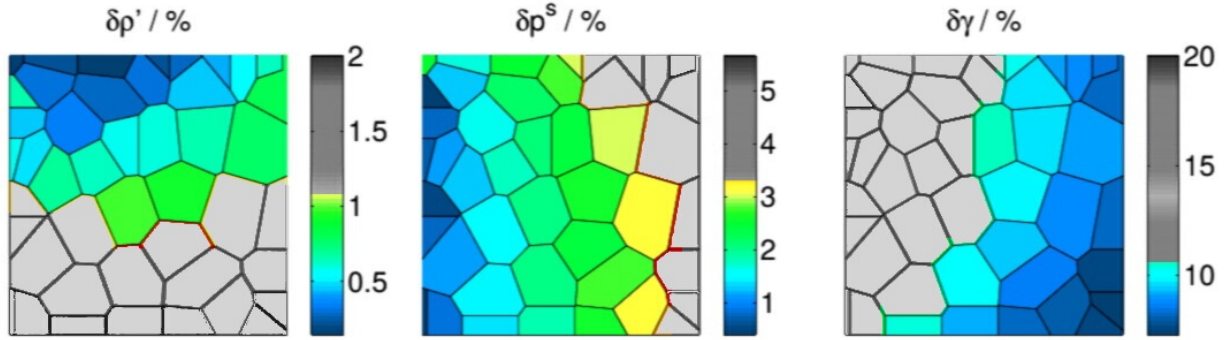


Pareto-optimal 2CLJQ models of molecular oxygen

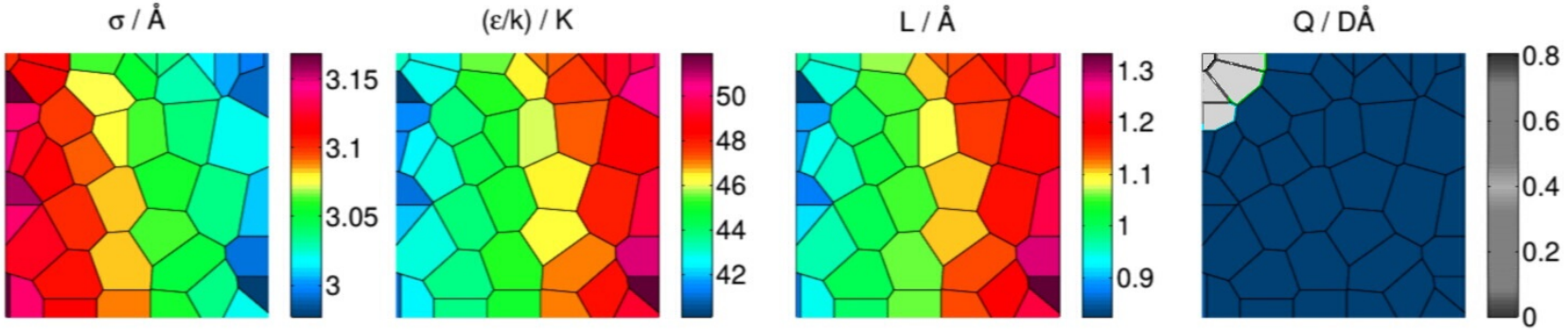


Model tailoring by the end user

For each specific application, accuracy requirements can be specified:

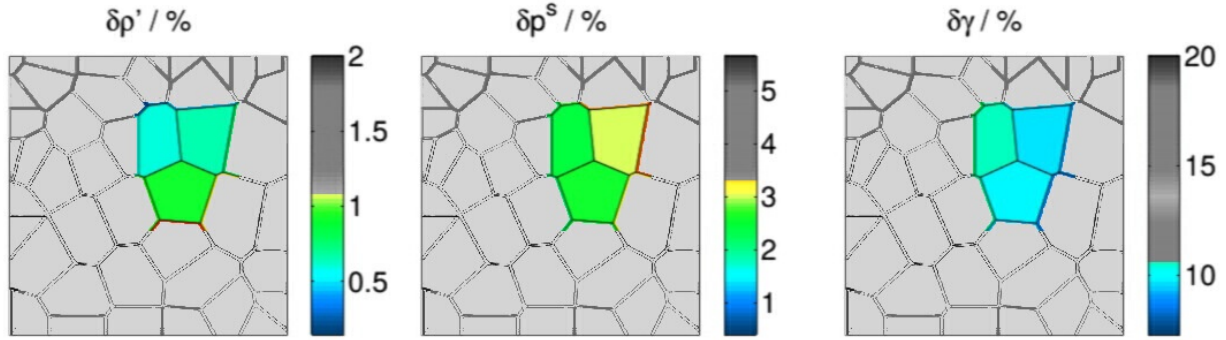


Restrictions imposed on 2CLJ models of molecular oxygen

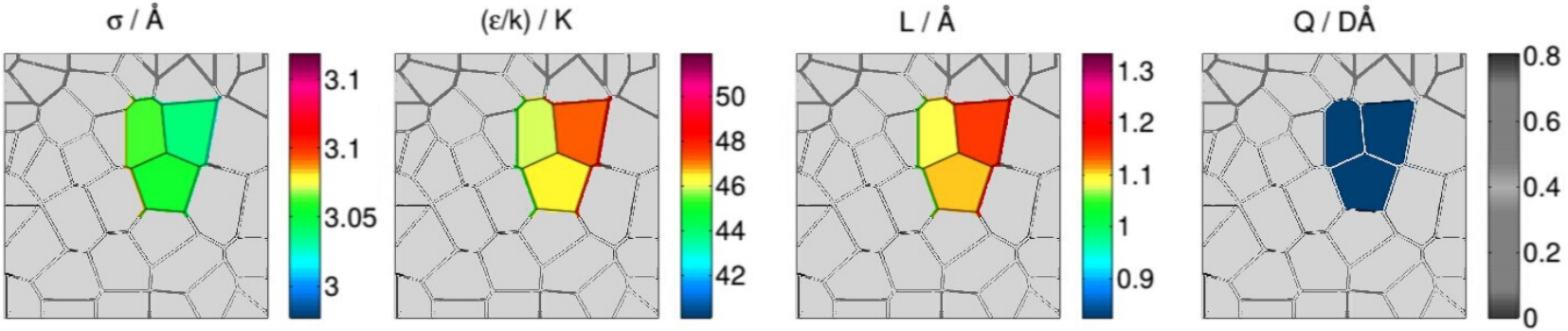


Model tailoring by the end user

Intersection of the highlighted areas within all replicas of the patch plot:



2CLJ models of molecular oxygen fulfilling all requirements



Summary

The traditional art of molecular modelling

An **expert modelling artist** designs and publishes

- a single optimized model for a particular fluid,
- according to his choice of criteria (often unknown to the public),
- users are passive, they have to live with the artists' decision.

Summary

The traditional art of molecular modelling

An **expert modelling artist** designs and publishes

- a single optimized model for a particular fluid,
- according to his choice of criteria (often unknown to the public),
- users are passive, they have to live with the artists' decision.

Scientific modelling by multicriteria optimization

For established model classes and multiple thermodynamic criteria,

- the dependence of thermodynamic properties on the model parameters is determined and correlated,
- the deviation between model properties and real fluid behaviour is characterized, and the Pareto set is published,
- users can design their own tailored model **with minimal effort**.