Automatic Parameterization of Force Fields for Liquids by Simplex Optimization

Roland Faller, Heiko Schmitz, Oliver Biermann, Florian Müller-Plathe
Max-Planck Institut für Polymerforschung, Ackermannweg 10, D-55128 Mainz
March 31, 2022

Abstract We demonstrate an automatic method of force field development for molecular simulations. Parameter tuning is taken as an optimization problem in many dimensions. The parameters are automatically adapted to reproduce known experimental data such as the density and the heat of vaporization. Our method is more systematic than guessing parameters and, at the same time saves human labour in parameterization. It was successfully applied to several molecular liquids: As a test, force fields for 2-methylpentane, tetrahydrofurane, cyclohexene and cyclohexane were developed.

Keywords: force fields, molecular dynamics, parameter optimization, molecular liquids, simulation techniques

1 Introduction

In atomistic molecular dynamics simulations, one of the central problems is the choice of the proper parameters for modeling the desired system. There is a variety of approaches to this problem. Ab initio quantum chemistry would be an ideal tool for this purpose if it were able to handle interactions of big molecules in reasonable time. The standard solution, however, is quite pragmatic. One either chooses a force field that reproduces certain experimental data or one takes standard values for the different atoms. Hence, force field design is either a cumbersome trial-and-error procedure or relies heavily on the transferability of parameters.

There are attempts to make the computer do this job, e.g. force field development by weak coupling [1, 2]. However, that procedure relies on the requirements that one force field parameter dominates the behavior of one property and that their relationship is monotonic. As, in more complex force fields, one property may be influenced significantly by several parameters, a more general multidimensional optimization algorithm is needed. In our approach, we consider the experimentally measured properties as multi-dimensional functions of
the parameters. Then we use the well-known simplex algorithm \cite{3} to find the optimum parameter set.

2 Algorithm and Implementation

2.1 Simplex algorithm

The simplex method is a well-known algorithm for minimization in many dimensions \cite{3}. It is not constrained by conditions like monotonicity, convexity or differentiability of the function being optimized. It minimizes any single-valued function of an arbitrary number of variables. Additionally, it is very robust in finding a local optimum. Its main drawback is the large number of necessary function evaluations, i.e. in our case MD simulation runs, which are quite time consuming. In the following we briefly summarize the simplex algorithm used in this work.

A simplex (a “d-dimensional distorted tetrahedron”) is a set of \(d+1\) points in the \(d\)-dimensional parameter space. It is transformed geometrically depending upon the “quality” of the function values. There are three geometric transformations in the algorithm.

1. In a reflection, the point \(x_i\) with the highest function value is reflected through the hyper-plane defined by the other points (see figure 1a).

\[
x'_i = \frac{2}{d} \sum_{j=1}^{d+1} x_j - \left(\frac{2}{d} + 1\right) x_i
\]  

2. An expansion by the factor \(\lambda\) is a linear transformation of one point along the normal of the hyper-plane defined by the others (fig. 1b).

\[
x'_i = \frac{1 - \lambda}{d} \sum_{j=1}^{d+1} x_j - \left(\frac{1 - \lambda}{d} + 1\right) x_i
\]  

Thus, a reflection is just the special case \(\lambda = -1\).

3. A (d-dimensional) contraction is a linear transformation of all but one point \(x_j\) towards the lowest point (fig. 1c). Contractions by a factor of 2 are applied.

\[
x'_i = \frac{1}{2}(x_i + x_j), \forall i \neq j
\]  

The algorithm runs iteratively. Each iteration starts with a reflection of the highest point. Depending on the function value at the new point, an expansion or a contraction is performed. If the new point is better than the best point an additional expansion with the factor \(\lambda = 2\) (i.e. the distance to the hyper-plane of the others is doubled) is applied to explore further into this “promising”
direction. If the new point point is very far away from the minimum (i.e. worse
than the second worst point up to now) an expansion with $\lambda = 0.5$ is applied.
If this resulting point is still very bad (in the above sense) a contraction around
the best point is performed. Then the next iteration, starting again with a
reflection, follows.

## 2.2 The target function evaluation

As the algorithm only knows about scalar functions in $\mathcal{R}^d$, we have to con-
struct a single-valued function $f_{\text{target}}(p_1, \ldots, p_d)$ of our force field parameters
$p_1, \ldots, p_d$. The function to be minimized should indicate the deviation of physi-
cal properties of the simulated model system from the real system as observed in
experiments. Typically, one chooses a set of physical properties $\{P_i\}$, which are
well characterized experimentally and converge rapidly in simulations. A natural
choice for $f_{\text{target}}$ is the square root of the weighted sum of relative squared
deviations

$$f_{\text{target}}(\{p_n\}) = \left( \sum_i w_i \left( 1 - \frac{P_i(\{p_n\})}{P_{i,\text{target}}} \right)^2 \right)^{1/2}, \quad (4)$$

where $P_{i,\text{target}}$ is the experimental value of property $P_i$. The square root is
chosen because it comes steeper to the minimum. The weights $w_i$ account for the
fact that some property may be easier to reproduce than others. Thus, the
algorithm can be forced to focus stronger on the difficult properties. Typically,
the density $\rho$ is easier reproduced than the enthalpy of vaporization $\Delta H_{\text{vap}}$,
which are the two properties we optimize our force fields against. They converge
rapidly and experimental data is readily available for many fluids (see e.g. [4, 5]).

If the number of parameters to be optimized is about 2 to 4 the flexibil-
ty to fit the data is normally sufficient and the computational time is still man-
ageable. If there are more target properties it may be necessary to increase the
dimensionality of the optimization space at the cost of more computer time.

In the beginning, a simplex of parameter sets has to be constructed by the
user. These data may be guessed from parameters for similar compounds or
from standard force fields [6, 7, 8]. Furthermore, a starting configuration of the
system is needed which should be close to the supposed real state. That means
that geometry and density should be almost correct. The starting configura-
tion is relaxed some picoseconds with a guessed force field in order to obtain
a proper liquid structure. The target function for the initial parameter sets is
first evaluated before the simplex algorithm starts.

## 2.3 Parameters to optimize

Since the dimensionality of parameter space is limited, we have to decide which
parameters of the force field we want to optimize. This number is mainly limited
by the available computing resources.
Typically, a Lennard-Jones potential is used to model the non-bonded interactions.

\[ V_{LJ} = 4\epsilon \left( \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^6 \right). \]  

(5)

The density \( \rho \) depends quite strongly on the Lennard-Jones radius \( \sigma \) whereas the enthalpy of vaporization \( \Delta H_{\text{vap}} \) depends stronger on \( \epsilon \). It is recommended to optimize non-bonded interaction parameters or charges and not the molecular geometry, because of simulational stability. The fact, that the geometry is mostly quite well known, supports this choice. There are several experimental methods to determine geometries, e.g. x-ray or neutron diffraction in the crystal or electron or microwave diffraction in the gas phase. Ab initio quantum chemistry, too, gives molecular structures with useful accuracy. These geometries can, in most cases, be used for the liquid phase as well. Hence, we did not try the algorithm on geometry optimization although this may be possible in principle. Our simulations focused on the liquid phase, whose macroscopic properties depend only weakly on internal force field parameters. Therefore, the force field parameters for angles and dihedral angles may be adopted from similar force fields.

2.4 Equilibration

A MD run can produce reliable results only if the system has been equilibrated. Therefore, we need a scheme to test for equilibration which has to fulfill several requirements: It has to reject reliably non-equilibrated configurations because otherwise all following results are meaningless. It has to work fully automatic inside the overall algorithm, and it has to equilibrate as fast as possible in order not to waste resources.

If the force field parameters (i.e. the Hamiltonian) of a simulation change between iterations, like in our case, a configuration equilibrated with respect to the old parameters is no longer equilibrated with respect to the new ones. Hence, after each change of parameters, i.e. in each step of the simplex algorithm we have to re-equilibrate with respect to the actual parameters. In order to do this, we take as the starting configuration the final configuration from a simulation with a parameter set, which is close to the new one. As “distance” in parameter space we define the sum of squared deviations

\[ |\{p^{(new)}\} - \{p^{(old)}\}|^2 := \sum_{i=1}^{n} (p_i^{(new)} - p_i^{(old)})^2. \]  

(6)

If, for some reason, the equilibration did not converge for that set or some other problem occurred a standard configuration is used.

Using the configuration selected in this way we start a number of successive equilibration runs (typical length 50ps with 1fs timestep). These runs are analyzed for equilibration until they are either accepted or a maximum number of
runs (in our case 10) is exceeded. In the latter case, the parameters are consid-
ered not useful and the target function $f_{\text{target}}$ is set to an arbitrary high value
in order to indicate the failure.

How does the automatic determination of equilibration work? To our knowl-
edge, there is no strict criterion for equilibration. The standard procedure is
to inspect visually the time development of a typical quantity (like the density
for low molecular weight liquids). Then one decides if it “settled” to stochastic
oscillations around a converged mean value. In our case, we use the following
test: The time series of the density is cut into 3 to 5 intervals, for each of which
the mean and the standard error are calculated. If all these averages agree
within their errors the configuration is considered equilibrated. In comparison
with the “human eye”-method, this method proved to be rather strict. How-
ever, this is necessary because we cannot accept non-equilibrated configurations
which would mislead the simplex algorithm. The equilibration scheme worked
well and led on average to an equilibrated configuration in about 3 to 4 runs.
Naturally, the number of runs decreases during the optimization because the
changes in the parameters get less drastic. We also checked a second equilibra-
tion test where the last third of the simulation was fitted by linear regression.
If the slope is zero within its error the configuration is assumed equilibrared.
The outcomes of the two tests differed only slightly.

Only very few parameter sets (less than 10%) had to be discarded due to
non-equilibration. Even fewer led to instabilities in the simulation.

2.5 Convergence criterion

The simplex algorithm finishes if the target function falls below a given thresh-
old $l$ which is usually set to about 1% (i.e. $f_{\text{target}} < l \approx 0.01$). If this is achieved
the parameters are deemed to be satisfactory. It does not make sense to repro-
duce experimental data more closely because the typical simulation error limits
the reliability anyway. In addition, the target values themselves carry some
uncertainty.

If the desired accuracy $l$ is not achieved and the simplex ends up in a local
minimum the algorithm is aborted. Therefore, the highest and lowest value of
the target function in the actual simplex are compared. Hence, if

$$\max(f_{\text{target}}) - \min(f_{\text{target}}) < \delta f \approx 0.001$$

(7)
is achieved further optimization makes no sense. In this case, either the num-
ber of parameters is too small to reproduce the desired number of properties
(overdetermination) or the appropriate parameter values are far off the initial
guess. We note that other convergence and abortion criteria are possible, for
example based on the size of the simplex. However, ours have proven to work
well in practice.
2.6 Implementation

The parts of the algorithm were implemented in different programming languages. The backbone is a tcsh script which calls all auxiliary programs and controls the overall flow of the procedure. It uses standard UNIX utilities like awk. The routine for producing a new topology from a set of parameters is a PERL script whereas the programs for calculating the distance in parameter space and the determination of equilibration are implemented in C++. Several programs from the YASP simulation package [9] are used: the MD program itself as well as the utilities for calculating enthalpy of vaporization and density. Any program or utility may be easily exchanged without affecting the overall structure, e. g. for using another MD program or a different equilibration scheme.

The structure of the procedure for obtaining a function value from a given set of parameters is shown in the flow diagram in figure 3.

3 Examples

The optimization procedure was tested with different model systems in order to explore its ability to produce force fields.

The all-atom nonbonded force field consists of a Lennard-Jones 12-6 potential and an electrostatic potential using reaction field and a finite cutoff (of 0.9nm)

\[
V^{(\text{nonb})} = 4\varepsilon_{ij} \left( \left( \frac{\sigma_{ij}}{r} \right)^{12} - \left( \frac{\sigma_{ij}}{r} \right)^{6} \right) + \frac{q_i q_j}{4\pi\varepsilon_0 \varepsilon} \left( \frac{1}{r} + \frac{\varepsilon_{RF} - 1}{2\varepsilon_{RF} + 1} \frac{r^2}{r_{\text{cutoff}}} \right). \tag{8}
\]

This potential is applied to atoms belonging to different molecules, internal non-bonded interactions are excluded in our test cases. The Lennard-Jones parameters between unlike atoms are derived by the Lorenz-Berthelot mixing rules [10]

\[
\varepsilon_{ij} = (\varepsilon_{ii}\varepsilon_{jj})^{\frac{1}{2}}, \quad \sigma_{ij} = \frac{1}{2}(\sigma_{ii} + \sigma_{jj}). \tag{9}
\]

A bond angle potential

\[
V^{(\text{angle})} = \frac{k^{(\text{angle})}}{2}(\Theta - \Theta_0)^2, \quad \Theta : \text{bond angle} \tag{10}
\]

and, for some molecules, torsional potentials with threefold symmetry

\[
V^{(\text{tors})} = \frac{k^{(\text{tors})}}{2}(1 - \cos(3\tau)), \quad \tau : \text{dihedral angle} \tag{11}
\]

or a harmonic dihedral potential

\[
V^{(\text{hd})} = \frac{k^{(\text{hd})}}{2}(\tau - \tau_0)^2 \tag{12}
\]
are applied in order to keep the correct molecular shape.

The bond lengths were constrained using the SHAKE algorithm [11, 12]. Our systems are subject to cubic periodic boundary conditions. The simulations were run at ambient conditions (T=298 K, p=1013 hPa). The neighbor-list [10] is calculated up to 1.0nm every 10 to 15 time-steps. We use the Berendsen algorithm for constant pressure and temperature [13]. The coupling times were 0.2 ps and 2 ps, respectively. The simulation runs lasted 50 ps at a timestep of 1 fs for each equilibration run and 100 ps at a timestep of 2 fs for the evaluation runs. The errors of the properties were obtained by a binning analysis [10].

3.1 Methylpentane

As a first test, a system of 125 uncharged 2-methylpentane molecules was optimized with respect to density $\rho$ and enthalpy of vaporization $\Delta H_{vap}$. All Lennard-Jones parameters are subject to optimization. However, all like atoms (C and H) are constrained to have the same LJ parameters. The internal part of the force field is taken from the AMBER force field [6]. This comprises the rule that the Lennard Jones $\epsilon$ (eq. 5) is scaled by a factor of 0.5 for 1-4 interactions.

We used the following parameter file to start the algorithm. The number “4” in the second line indicates the dimensionality of the parameter space. The following five lines are the guesses of the parameters, the initial simplex. The last column shows the results after evaluation of the target function. The Lennard-Jones energies $\epsilon$ and $\Delta H_{vap}$ are measured in kJ/mol the radii $\sigma$ in nm, the density $\rho$ in kg/m$^3$.

```
## $\epsilon_C$ $\sigma_C$ $\epsilon_H$ $\sigma_H$ $f_{\text{target}}$ $\Delta H_{\text{vap}}$ $\rho$
4
0.291643 0.339215 0.154545 0.258859 0.030287 29.42 636.1
0.290554 0.340351 0.151371 0.260571 0.052871 28.88 626.5
0.290167 0.340656 0.151116 0.260825 0.057492 28.81 623.8
0.290545 0.341183 0.149968 0.260825 0.043120 29.18 629.6
0.290421 0.341161 0.150191 0.260763 0.051831 28.94 626.2
```

These parameters produce properties which are already quite close to the target values. The simplex algorithm now does the fine tuning. First, the simplex is reflected away from parameter set 3. The new set is

```
## $\epsilon_C$ $\sigma_C$ $\epsilon_H$ $\sigma_H$ $f_{\text{target}}$ $\Delta H_{\text{vap}}$ $\rho$
0.291414 0.340299 0.151922 0.259653 0.045721 29.04 629.6
```

After 11 optimization steps, which took about 2 weeks altogether on a DEC 433MHz processor, the optimization finally finished with the following values:

```
## $\epsilon_C$ $\sigma_C$ $\epsilon_H$ $\sigma_H$ $f_{\text{target}}$ $\Delta H_{\text{vap}}$ $\rho$
0.294477 0.336339 0.162144 0.254668 0.008629 30.15 652.7
```

Figure 3 illustrates the progress of the optimization by means of a previous run. The circles in fig. 3a) show the results of function evaluations and the solid line
shows the current best values of $f_{\text{target}}$. In the beginning, the function values scatter quite strongly. In the run of the algorithm, this starts to decrease. Figure 3b) shows how density and enthalpy of vaporization reach their target values. The only maintenance which had to be done was restarting the algorithm after a shutdown of the computer system. The whole algorithm proved to be stable and worked fully automatically. Only once the equilibration failed due to exceeding the limit of 10 runs. It is shown by the spike in figure 3 (which goes up to 100000). The final force field is shown in table I. These values reproduced the experimental data in a satisfactory way (table II).

3.2 Tetrahydrofurane

As a different test system we used tetrahydrofurane (THF). Here, we especially focused on the optimization of partial charges. The hydrogens did not carry any partial charges but oxygen and carbon did. The charges of the carbons 2 and 5 and the carbons 3 and 4 are the same for symmetry reasons. With the constraint of electroneutrality, there were two charge parameters to be optimized. We chose $q_O$ and $q_{C2/C5}$, then we have $q_{C3/C4} = \frac{1}{2}(-q_O - 2q_{C2/C5})$. Additionally, the oxygen parameters $\epsilon_O$ and $\sigma_O$ were included in the optimization. The first guess for the partial charges was taken from a quantum chemical Hartree-Fock calculation with a 6-311G** basis set using Gaussian 94 (Mulliken charges with hydrogens summed into heavy atoms) [14]. This yielded also the bond angle values. The bond lengths are taken from electron diffraction [5]. The simulated system contained 216 molecules. The electrostatic interactions were simulated with a reaction field correction ($\epsilon_{RF} = 7.5$) using the same cutoff $r_c = 0.9$ nm as for the Lennard-Jones potential. Here the following starting simplex was taken:

\[
\begin{array}{cccccccc}
\# & q_O & q_{C2/C5} & \epsilon_O & \sigma_O & f_{\text{target}} & \Delta H_{\text{vap}} & \rho \\
4 & 0.581241 & 0.225443 & 0.516818 & 0.208594 & 0.084176 & 32.72 & 816.96 \\
 & 0.659870 & 0.251733 & 0.325788 & 0.300391 & 0.137257 & 35.39 & 811.81 \\
 & 0.480765 & 0.251793 & 0.635725 & 0.316797 & 0.186735 & 27.68 & 774.02 \\
 & 0.684265 & 0.276431 & 0.729962 & 0.264345 & 0.232852 & 39.29 & 847.92 \\
 & 0.582970 & 0.220928 & 0.535152 & 0.192715 & 0.088840 & 33.57 & 823.39
\end{array}
\]

The first optimization attempt, which tried to optimize procedure the above parameters, ended up in a local minimum with $f_{\text{target}} \approx 0.07$ after 53 evaluations because the experimental liquid density could not be reproduced satisfactorily. It was systematically too low. Therefore, the best parameters so far were frozen and a new optimization was started where only the Lennard-Jones radii of all species were optimized. Finally, convergence ($f_{\text{target}} \leq 0.01$) was achieved. The resulting THF force field is described in table III. These parameters lead to the physical properties shown in table IV. Our force field has about the same accuracy as an earlier Monte Carlo simulation of a united atom OPLS model for THF [17].
3.3 Cyclic Hydrocarbons

Finally, the method was applied in order to obtain force fields for cyclohexene and cyclohexane with 125 molecules in the periodic box. The geometries were taken from electron diffraction data [5]. The geometric data are shown in table V. In the cyclohexene force field, harmonic dihedral angles are used in order to keep the atoms around the double bond in plane, since the sp\(^2\) hybridisation prevents the double bond from rotating. Additionally, standard torsional potentials with three-fold symmetry are used. Cyclohexane was simulated without any dihedral angle potentials. For the angular force constants we used a standard value, since they are believed to be of minor importance for the desired properties. Additionally, they may be compensated by the nonbonded parameters.

The optimized Lennard-Jones 12-6 parameters are shown in table VI. The parameters included in the optimization procedure are denoted with opt in the table. No charges were used. All the parameters which are not optimized as well as the initial simplices have been taken from similar force fields.

Except for the Lennard-Jones \(\epsilon\) of the hydrogens, the resulting final parameters are very similar for the two molecules. This shows that force field parameters are not a unique description of a certain atom type but rather they are only a part of the overall molecular description. Mostly, however, the same atoms in similar environments may be described by similar parameters.

We compare our thermodynamic data with experiment in table VII. A more detailed analysis of transport properties of these cyclic hydrocarbons will be published elsewhere [16]. The cyclohexane force field yields a slightly better comparison to experiment than in a recent study using a commercial force field [8] whereas the study of cyclohexene is the first to our knowledge.

4 Conclusions

We applied the simplex algorithm to the problem of force field optimization for MD simulations. Given a good initial guess for the force field parameters and the experimental data for some properties, our method tunes the parameters to optimum values. Once the routine has been set up, very little human interference is required for maintenance. The algorithm proved to be robust and found local minima if set up properly. The resulting force fields are able to reproduce experimental data of low molecular weight liquids in a satisfactorily.

In the examples of this contribution, we typically optimized 4 force field parameters against 2 observables. Hence, the solutions are most likely not unique. This, however, is a feature of the problem of finding a force field given a small number of observables, not of the algorithmic solution presented here. Density and enthalpy of vaporization are the two properties most commonly used to derive force fields, as they are experimentally available for many fluids and quickly converging in a simulation. At present, our method has to be used with a judicious choice of starting values for the parameters to prevent it
from optimizing towards an unphysical, non-transferable set of parameters. It
shares this restriction with all other methods of finding force fields, including
“optimization by hand”. On the other hand, it is mostly not difficult to come
up with a reasonable first guess for the parameters. What is time consuming is
the fine tuning and it is at this point where our method offers help.

A possible way out of the dilemma is to increase the base of experimental
observables used in the target function. In a few selected cases we have used
other liquid properties than \( \rho \) and \( \Delta H \text{vap} \) together with the other refinement
scheme \([1, 2]\). However, one has to note that there are not too many suitable
fluid properties. Some properties are of similar character to what we already
have. For example, the excess chemical potential \( \mu_{ex} \) probes almost the same
regions of the force field as \( \Delta H \text{vap} \) and, thus, does not add much independent
information. Dynamic properties often converge too slowly in simulations to
be useful (shear viscosity, dielectric constant) or the experimental data are not
of sufficient quality (tracer diffusion coefficient, molecular reorientation times).
We, therefore, follow the strategy of optimizing towards \( \rho \) and \( \Delta H \text{vap} \) and sub-
sequently checking the final force field against other liquid properties. For our
models of cyclic hydrocarbons we have, for instance, calculated tracer and bi-
nary diffusion coefficients as well as molecular reorientation times for both the
pure liquids and binary mixtures, and the results agree well with experimental
data where available \([16]\).

The automatic parameterization scheme presented has the small disad-
vantag of probably requiring moderately more computer time than an optimiza-
tion by hand. This is more than offset by the invaluable advantage of freeing
researchers from the labour of parameter optimization. In a reasonable use of
computing time (a few weeks workstation time) one is able to cope with dimen-
sionalities of parameter space of about 4. This depends, however, strongly on
the actual simulations to be performed. On the other hand, the full potential of
speeding up our algorithm has not yet been realized. We foresee possibilities of
substantial improvement by using a less rigorous and maybe adaptive equilibra-
tion scheme and by substituting the simplex algorithm by a faster converging
optimizer (e.g. Fletcher) in the final stages of minimization. This remains an
interesting starting point for future research.

References

[1] Njo, S. L.; van Gunsteren, W. F.; Müller-Plathe, F. J Chem Phys 1995,
102, 6199–6207.

[2] Berweger, C. D.; van Gunsteren, W. F.; Müller-Plathe, F. Chem Phys
Letters 1995, 232, 429–436.

[3] Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. Numerical
Recipes in C: The Art of Scientific Computing; Cambridge University Press:
New York, second edition 1992.
[4] Landolt, H.; Börnstein Numerical data and functional relationships in science and technology; Springer: Berlin, 1993.

[5] Lide, D. R., ed. CRC handbook of chemistry and physics; CRC Press: Boca Raton, 76th edition 1995.

[6] Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz Jr., K. M.; Ferguson, D. M.; Spellmeyer, C. D.; Fox, T.; Caldwell, J. W.; Kollman, P. A. J Am Chem Soc 1995, 117, 5179–5197.

[7] van Gunsteren, W. F.; Billeter, S. R.; Eising, A. A.; Hünenberger, P. H.; Krüger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G. Biomolecular simulation: The GROMOS manual and user guide; Vdf: Zürich, 1996.

[8] Sun, H. J Phys Chem B 1998, 102, 7338–7364.

[9] Müller-Plathe, F. Comput Phys Commun 1993, 78, 77–94.

[10] Allen, M. P.; Tildesley, D. J. Computer Simulation of Liquids; Clarendon Press: Oxford, 1987.

[11] Ryckaert, J.-P.; Cicotti, G.; Berendsen, H. J. C. J Comput Phys 1977, 23, 327–341.

[12] Müller-Plathe, F.; Brown, D. Comput Phys Commun 1991, 64, 7–14.

[13] Berendsen, H. J. C.; Postma, J.; van Gunsteren, W.; DiNola, A.; Haak, J. J Chem Phys 1984, 81, 3684–3690.

[14] Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheeseman, J. R.; Keith, T. A.; Petersson, G. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; Defrees, D. J.; Baker, J.; Stewart, J. P.; Head-Gordon, M.; Gonzalez, C.; Pople, J. A. Gaussian 94; Gaussian, Inc.: Pittsburg PA, 1995.

[15] Briggs, J. M.; Matsui, T.; Jorgensen, W. L. J Comput Chem 1990, 11, 958–971.

[16] Schmitz, H.; Faller, R.; Müller-Plathe, F. Molecular mobility in cyclic hydrocarbons - a simulation study J. Phys. Chem. B, submitted.

[17] Harris, K. R.; Dunlop, P. J. Ber Bunsenges Phys Chem 1994, 98, 560–562.
### Table I: Details of the Methylpentane force field

| bonded parameters | non-bonded parameters |
|-------------------|------------------------|
| parameter         | value                  | parameter         | value                  |
| \(|C-C|\)           | 0.1526 nm              | \(m_C\)          | 12.01 amu              |
| \(|C-H|\)           | 0.109 nm               | \(m_H\)          | 1.00782 amu            |
| \(k_{C-O-C}^{(angle)}\) | 167.47 kJ/mol radi\(^2\) | \(\epsilon_C\)   | 0.294 kJ/mol           |
| \(k_{C-O-C}^{(angle)}\) | 209.34 kJ/mol radi\(^2\) | \(\epsilon_H\)   | 0.162 kJ/mol           |
| \(k_{H-O-H}^{(angle)}\) | 146.54 kJ/mol radi\(^2\) | \(\sigma_C\)     | 0.336 nm               |
| \(\sigma_{H}\)     |                        | \(\sigma_{H}\)   | 0.254 nm               |

| \(H-C-H\)         | 109.5°                 | \(C-C-C\)        | 109.5°                 |
| \(C-C-H\)         | 109.5°                 | \(C-C-C\)        | 109.5°                 |
| \(k_{C-O-C}^{(tors)}\) | 11.5kJ/mol             | \(k_{C-O-C}^{(tors)}\) | 11.5kJ/mol             |
| \(k_{H-O-H}^{(tors)}\) | 11.5kJ/mol             | \(k_{H-O-H}^{(tors)}\) | 11.5kJ/mol             |

**Table II: Experimental and simulated properties of 2-methylpentane**

|                     | exp.       | sim.        |
|---------------------|------------|-------------|
| \(\Delta H_{vap}[kJ/mol]\) | 29.89[5]  | 29.92±0.03  |
| \(\rho[kg/m^3]\)    | 653.0[5]   | 653.4±0.5   |
| \(D[cm^2/s]\)       | (2.5±0.2)×10\(^{-5}\) |             |

**Figure 1:** Transformations of the simplex used during the algorithm: a) reflection, b) expansion, c) contraction.
| parameter | value       | parameter | value       |
|-----------|-------------|-----------|-------------|
| $\varepsilon_O$ | 0.509 kJ/mol | $|C-O|$ | 0.1428 nm |
| $\varepsilon_H$ | 0.200 kJ/mol | $|C-H|$ | 0.1115 nm |
| $\varepsilon_C$ | 0.290 kJ/mol | $|C-C|$ | 0.1536 nm |
| $\sigma_O$ | 0.243 nm | $k_{\text{angle}}$ | 450.0 kJ/mol rad $^2$ |
| $\sigma_H$ | 0.193 nm | C-O-C | 111.2° |
| $\sigma_C$ | 0.306 nm | O-C-C | 106.1° |
| $q_O$ | $-0.577 e$ | C-C-C | 101.4° |
| $q_C2$ | 0.228 e | O-C-H | 109.0°, 109.3° |
| $q_C3$ | 0.061 e | C$_3$-C$_2$-H | 111.0°, 113.2° |
| $m_O$ | 15.9949 u | H-C$_2$-H | 108.2° |
| $m_C$ | 12.0 u | H-C$_3$-H | 108.1° |
| $m_H$ | 1.00787 u | C$_2$-C$_3$-H | 110.4° (2×) |
|          |             |           | 112.8° (2×) |
|          |             |           |               |
|          |             |           | C$_3$-C$_4$-H | 113.7° (2×) |
|          |             |           |               |
|          |             |           | 110.4° (2×) |

Table III: Optimized force field for tetrahydrofurane. In the case of two angles in one line one is applied to the first hydrogen, the other to the second hydrogen, otherwise the angles would not be consistent with each other.

| Parameter | Experiment | Simulation (this work) | Simulation [13] |
|-----------|------------|------------------------|-----------------|
| $\rho$   | 889.0 kg/m$^3$ | (886.0 ± 1.3) kg/m$^3$ | (882±1) kg/m$^3$ |
| $\Delta H_{vap}$ | 31.99 kJ/mol | (32.0±0.1) kJ/mol | (31.57±0.08) kJ/mol |

Table IV: Properties of tetrahydrofurane
Table V: Geometry of the cyclic hydrocarbons and their intramolecular potentials

| property                        | \( \text{C}_6\text{H}_{10} \) | \( \text{C}_6\text{H}_{12} \) |
|---------------------------------|-------------------------------|-------------------------------|
| \(|C_{sp2}=C_{sp2}|\)           | 0.1334nm                      |                               |
| \(|C_{sp2} \cdot C_{sp3}|\)     | 0.150nm                       |                               |
| \(|C_3 \cdot C_4|, |C_5 \cdot C_6|\) | 0.152nm                       |                               |
| \(|C_4 \cdot C_5|\)           | 0.154nm                       |                               |
| \(|C-C|\)                       |                               | 0.1526nm                      |
| \(|C_{sp2} \cdot H|\)         | 0.108nm                       |                               |
| \(|C_{sp3} \cdot H|\)         |                               | 0.109nm                       |
| \( k_{(angle)} \text{C-C-C} \) | 450 kJ/mol rad\(^2\)         | 335 kJ/mol rad\(^2\)         |
| \( k_{(angle)} \text{C=C-C} \) | 500 kJ/mol rad\(^2\)         |                               |
| \( k_{(angle)} \text{C-C-H} \) | 500 kJ/mol rad\(^2\)         | 420 kJ/mol rad\(^2\)         |
| \( k_{(angle)} \text{H-C-H} \) | 500 kJ/mol rad\(^2\)         | 290 kJ/mol rad\(^2\)         |
| C=C-C                           | 112.0°                        |                               |
| C-C-C                           | 110.9°                        | 109.5°                        |
| C_{sp2}-C-C                    | 123.45°                       |                               |
| C-C-H                           |                               | 109.5°                        |
| C_{sp2}-C-H                    | 119.75°                       |                               |
| H-C-H                           |                               | 109.5°                        |
| \( k_{(hd)} \text{C=C=C-C} \)  | 250 kJ/mol rad\(^2\)         |                               |
| \( k_{(hd)} \text{H-C=C-C} \)  | 200 kJ/mol rad\(^2\)         |                               |
| \( k_{(tor)} \text{C=C=C-C} \) | 10kJ/mol                      |                               |

Table VI: Cyclohexene and cyclohexane non-bonded parameters

| parameter         | opt/fix | \( \text{C}_6\text{H}_{10} \) | \( \text{C}_6\text{H}_{12} \) |
|-------------------|---------|-------------------------------|-------------------------------|
| \( \epsilon_C \) | opt     | 0.296kJ/mol                   | 0.299kJ/mol                   |
| \( \epsilon_H \) | opt     | 0.265kJ/mol                   | 0.189kJ/mol                   |
| \( \sigma_H \)   | opt     | 0.252nm                       | 0.258nm                       |
| \( \sigma_C \)   | opt     |                               | 0.328nm                       |
| \( \sigma_{C_{sp2}} \) | fix    | 0.321nm                       |                               |
| \( \sigma_{C_{sp3}} \) | fix    | 0.311nm                       |                               |
| \( m_C \)        | fix     | 12.01amu                      |                               |
| \( m_H \)        | fix     | 1.00787amu                    |                               |
Table VII: Properties of cyclohexene and cyclohexane

| Property         | cyclohexene       | cyclohexane       |
|------------------|-------------------|-------------------|
| $\rho$ [kg/m$^3$] | $805.8 \pm 1.5$   | $777.6 \pm 1.5$   |
| $\Delta H_{\text{vap}}$ [kJ/mol] | $33.47 \pm 0.1$ | $33.33 \pm 0.05$ | $33.41$ |

Figure 2: Flow diagram of the algorithm, one iteration.
Figure 3: Convergence of a previous methylpentane optimization run: a) Target function: Solid line: best value of $f_{\text{target}}$; Circles/dotted line: actual value of $f_{\text{target}}$. b) Properties: density and enthalpy of vaporization.