Bottled SAFT: a Web App Providing SAFT-\(\gamma\) Mie Force Field Parameters for Thousands of Molecular Fluids

Åsmund Ervik,† Andres Mejía,‡ and Erich A. Müller∗,¶

†Department of Energy and Process Engineering, Norwegian University of Science and Technology (NTNU), Trondheim, N-7491, Norway
‡Departamento de Ingeniería Química, Universidad de Concepción, Concepción, 4089100, Chile
¶Department of Chemical Engineering, Imperial College London, London, SW20 8EB, UK

E-mail: e.muller@imperial.ac.uk

Abstract

Coarse-grained molecular simulation has become a popular tool for modelling simple and complex fluids alike. The defining aspects of a coarse grained model are the force field parameters, which must be determined for each particular fluid. Since the number of molecular fluids of interest in nature and in engineering processes is immense, constructing force field parameter tables by individually fitting to experimental data is a futile task. A step towards solving this challenge was taken recently by Mejia et al., who proposed a correlation that provides SAFT-\(\gamma\) Mie force field parameters for a fluid provided one knows the critical temperature, the acentric factor and a liquid density, all relatively accessible properties. Building on this, we have applied the correlation to more than 6000 fluids, and constructed a web application, called “Bottled SAFT” which makes this data set easily searchable by CAS number, name or chemical formula.
Alternatively, the application allows the user to calculate parameters for components not present in the database. Once the intermolecular potential has been found through Bottled SAFT, code snippets are provided for simulating the desired substance using the “raaSAFT” framework, which leverages established molecular dynamics codes to run the simulations. The code underlying the web application is written in Python using the Flask microframework; this allows us to provide a modern high-performance web app while also making use of the scientific libraries available in Python. Bottled SAFT aims at taking the complexity out of obtaining force field parameters for a wide range of molecular fluids, and facilitates setting up and running coarse-grained molecular simulations. The web application is freely available at [http://www.bottledsaft.org](http://www.bottledsaft.org).

The underlying source code is available on Bitbucket under a permissive license.

**Introduction**

Coarse-grained molecular simulation has become a popular simulation tool\(^1\). By coarse-grained we mean that the particles in a simulation do not represent individual atoms, but perhaps two to four atoms bunched together into a “bead”. In comparison with atomistic simulation this trades resolution for speed; by resolving less details of the system we can simulate larger and more complicated systems at the same cost. But lower resolution is not the only difference: consider as an example 10 different chemical elements used in atomistically-detailed models for some class of molecules. Then the number of coarse-grained bead types that can be constructed from two, three or four of these elements is 375, and this number grows combinatorically with the number of elements. While of course many of these combinations are chemically irrelevant, it nevertheless illustrates the greater challenge faced by coarse-grained models when it comes to constructing the required force field parameters for a wide range of molecules. Furthermore, it has been demonstrated\(^2\) that a simple integration of atomistic information inevitably leads to an information loss, and the potentials obtained through this “bottom-up” approach lack robustness and transferability.
A step towards solving this problem is offered by top-down coarse-graining methods, of which the SAFT-\(\gamma\) Mie approach is an example; see Ref. [2] for a thorough review. An overview of the approach is given in the supporting information. In principle, this approach obtains the force field parameters by fitting an equation of state to experimental thermophysical data, and invoking the unique one-to-one link between the equation of state and the underlying molecular force field. Recognising that the procedure could be simplified further, Mejía et al. [3] developed the “M&M” correlation which takes as input just the critical properties, the acentric factor of the compound, and the liquid density at a reduced temperature \(T_r = 0.7\); which if unavailable can be estimated from a Rackett-type equation [4]. Thus the force field parameters for a given liquid can be computed directly without any iteration. Furthermore, large tables with the required critical properties are found in the literature. For the current application, we have used as a basis the book by Yaws [5] which gives data for more than 7800 compounds. Applying the correlation to this data set, and filtering out models which are outside the correlation range of application, SAFT models for more than 6000 compounds are available. A description of how the data set was filtered is found in the supporting information.

In the web application Bottled SAFT, each model is identified by the name, chemical formula and the CAS number (if it has one). The application lets the user search by either of these, including searching for parts of a name. Once the user has arrived at the desired model, Bottled SAFT constructs an input file for raaSAFT [6], our open-source Python framework for easily setting up molecular dynamics simulations with the SAFT-\(\gamma\) Mie potential. raaSAFT employs widely-used molecular dynamics codes to run the actual simulations; the default is HOOMD-blue [7], but GROMACS [8] can also be used. In addition to searching the database, users may also run the M&M correlation with custom input, e.g. if they have thermophysical properties for a compound of interest which is not in the database. Finally, we provide as a proof-of-concept a page with a molecule drawer integrated with the Python backend. The name Bottled SAFT stems from the fact that “saft” in German and Scandinavian languages...
means “fruit juice”, while the acronym SAFT stands for “Statistical Associating Fluid Theory”.

**Implementation**

Bottled SAFT is implemented in the Python programming language. The recent development of web microframeworks in Python (e.g. Flask) has enabled scientists to write web applications in a language with many scientific libraries. For building Bottled SAFT, we have used Flask and its extensions together with the libraries Pandas and RDkit. By using the Bootstrap CSS extension of Flask, we achieve a modern look which works equally well on personal computers and mobile devices. We also integrate into the application the ChemDoodle Web Components JavaScript library, and demonstrate how to get a molecule drawn by the user into an RDkit representation in the Python code for further processing.

The complete code underlying the web application is open source, available from Bitbucket, and can be run locally on a user’s computer for testing and development. The address of the web application is http://www.bottledsaft.org. The full code running this web application can be obtained from the repository at [https://bitbucket.org/asmunder/bottledsaft](https://bitbucket.org/asmunder/bottledsaft). Executing the saftflask.py file will run the web application on the local computer, assuming the dependencies are installed. For more details of the technology stack, see the supporting information.

**Using Bottled SAFT**

When the user accesses http://www.bottledsaft.org they come to the front page. At the top of this page, three cards display short information and links to the three different components: search the database, run the correlation with custom input, or draw a molecule with the sketcher. Below this, a brief review of the SAFT-γ Mie approach is given, with references to the literature.

When the user clicks "Search the bottle", the /search page is rendered from the code
shown in the supplementary information. This page is shown on the left of figure 1.

Figure 1: Left: the /search page. Right: the top table displays the properties of the compound. The bottom cards display the resulting force field parameters, with an image showing the number of beads superimposed on the molecule. At the bottom is a button for getting the script implementing it in raaSAFT. More cards for higher bead numbers are below this (not shown). (For higher resolution and more details please visit the actual webpage.)

When searching by formula or name, multiple results may be returned. The next page then displays a drop-down list where the user can select the desired compound. Once a unique compound has been selected, the user is given the option of refining the model, by entering a more accurate value for the liquid density at a reduced temperature $T_r = 0.7$. As mentioned, this value is otherwise estimated from a Rackett-type equation. Once the result has been refined, or the user has decided to use the value from the Rackett equation, the result is presented. This is shown on the right of figure 1. The figure of the molecule is linked in from the NIH Chemical Identifier Resolver.
Once one or more models are obtained, they can be used in molecular dynamics simulations. As a rule of thumb, the number of beads should be chosen to represent 3 or 4 of the atoms heavier than hydrogen. When several models are used together in a simulation, i.e. in a mixture or a heterogeneous molecule, the cross-interaction between beads of different type can be obtained from the Lafitte combination rules. In these there is one free parameter $k_{ij}$, which as a reasonable starting point may be set to zero. When possible, the value of $k_{ij}$ should be tuned by comparing the simulation results to experimental values for a macroscopic property, such as solubilities, interfacial tension, vapour-liquid compositions, etc. It is important to stress that the parameters give Mie potentials representing different compounds, and there is a risk when trying to “mix and match” these with other coarse-grained models, e.g. MARTINI, Elba, Sirah, Palace or other force fields for lipids, proteins or DNA.

Results and Discussion

To demonstrate how Bottled SAFT enables the study of a vast number of complicated fluid systems, we consider three cases: a) thermophysical properties of the refrigerant R134a, b) boiling point predictions for the entire database compared to literature values, and c) a ternary mixture of sulfolane, octane and benzene. Models for several of the most common components in the database, such as the shorter alkanes, aromatics, gases like $N_2, CO_2$ etc. have previously been studied in Ref. 10.

Case a: Thermophysical properties of R134a

R134a (1,1,1,2-Tetrafluoroethane, $CH_2FCF_3$) is an ozone-layer friendly refrigerant which is widely used in domestic and automobile applications. Interestingly, simultaneous fitting of the densities and vapor pressures of this fluid can not be obtained by the use of a simple Lennard-Jones potential. Using the parameters in Bottled SAFT, we have computed the coexistence densities, vapor pressure, interfacial tension and enthalphy of vaporization as a
function of temperature from 220 K to 340 K. The details of these simulations are given in the supporting information. We have also used the link between the force field parameters and the SAFT Equation of State (EoS) to compute these properties as function of temperature directly from the EoS. The results are shown in figure 2, together with experimental data from the NIST webbook[11], where excellent agreement is found.

Case b: Boiling points of the entire database

One of the strengths of the SAFT-\(\gamma\) Mie approach is the one-to-one link between the molecular force field and the SAFT EoS, and that computations using one or the other are in very close agreement, as shown in the previous figures and in the original SAFT-\(\gamma\) Mie publications[2,9,12,13]. While running molecular simulations for the entire database is a herculean undertaking, obtaining results from the corresponding EoSs is a simpler task. To test the accuracy of the entire database of parameters, the corresponding EoSs were used to predict the boiling points of all components. These results are compared to the boiling points provided by Yaws[5]. It is important to note here that these latter values were not used when fitting the EoS/force field parameters. The result is shown in the parity plot in figure 3. As this plot shows, the agreement is overall very good, with absolute deviations below 2% for 99.9% of the compounds in the database. This indicates the consistent performance of the force field parameters in Bottled SAFT.

Case c: Liquid-liquid equilibria of sulfolane + octane + benzene

This system has important applications in the chemical processing industry, where sulfolane is used for liquid-liquid extraction of aromatics during petroleum refinement. Sulfolane is also used to purify natural gas in the sulfinol process. The sulfolane + octane + benzene system has a liquid-liquid phase split where an octane-rich phase and a sulfolane-rich phase are formed, with the benzene soluble in both phases.

To be able to predict the two liquid phase compositions for the ternary system and
Figure 2: Results for coexistence densities, vapor pressure, interfacial tension and enthalpy of vaporization for R134a. Filled circles: simulation results. Lines: EoS results. Open circles: values from NIST[11]. Details of the simulations and EoS calculations are given in the supporting information.

subsequently compare the results with experiments, we consider the system at a temperature of 403.15 K, atmospheric pressure, and with an overall composition of 40 % sulfolane, 40 % octane and 20 % benzene (all in mole percent). Models for this system were obtained from
Figure 3: Parity plot of the predicted values of boiling points using the parameters in Bottled SAFT versus the values tabulated by Yaws. The inset shows density contours for the absolute deviations (in %).

Bottled SAFT. [For details of the models and the simulation, see the supporting information.] The result is shown in figure 4 as a ternary phase diagram, where excellent agreement with experimental results is observed.

Conclusions

Coarse-grained molecular simulation has shown promise as a powerful tool to study complex fluid mixtures. Its widespread adoption requires the availability of accurate and reliable force field parameters for a wide range of substances. With the SAFT-γ Mie force field, a top-down approach to coarse-graining, we can compute force field parameters directly from
Figure 4: A ternary phase diagram, showing the experimental results by Lee and Kim[14] (black points) connected with tie-lines (dark gray), and with a thin gray curve as a guide to the eye showing the two-phase region. The blue open point in the middle is the initial composition in the simulation, and the two solid blue points are the compositions of the two liquid phases which separate. Note that the blue tie-line obscures the gray tie-line for these points.
the critical properties of the compound in question, by use of the M&M correlation. For
the web application demonstrated here, Bottled SAFT, we have applied the correlation
to a large database of critical properties. This makes Bottled SAFT a frontend to a
database with parameters for over 6000 compounds that can be simulated, with the additional
possibility of tailor-made calculations for substances not in the database. To facilitate running
molecular dynamics simulations with the resulting parameters, the user can download a script
implementing the result in our raaSAFT framework. The end result is an unprecedented
ease-of-use up to the point of running an initial simulation.

Acknowledgement

The authors acknowledge the contributions of Muhammed Jansi and Karson Wong, who
compiled the initial table of parameters. We also thank Prof. Bernhard Müller and Dr. Svend
Tollak Munkejord for their comments on the manuscript.

ÅE acknowledges the financial support from the project Fundamental understanding of
electrocoalescence in heavy crude oils coordinated by SINTEF Energy Research and funded
by the Petromaks programme of the Research Council of Norway (206976), Petrobras, Statoil
and Wärtsilä Oil & Gas Systems. AM acknowledges the financial support from FONDECYT
(Chile) under the Project 1150656. EAM acknowledges financial support for the Molecular
Systems Engineering group from the Engineering and Physical Sciences Research Council

Supporting Information Available

Supporting Information includes a brief outline of the SAFT-γ Mie approach, how results
from the correlation were tabulated and filtered, details of the Bottled SAFT implementation,
performance testing, and details concerning the molecular simulations. This material is
available free of charge via the Internet at http://pubs.acs.org/.
References


(6) Ervik, Å.; Serratos, M. G. J.; Müller, E. A. raaSAFT: a Framework Enabling Coarse-Grained Molecular Dynamics simulations Based on the SAFT-$\gamma$ Mie Force Field. Submitted to Comp. Phys. Comm. 2016,


Graphical TOC Entry