Fluid Phase Equilibria 373 (2014) 100–108

Contents lists available at ScienceDirect

Fluid Phase Equilibria

journal homepage: www.elsevier.com/locate/fluid

Multicriteria optimization of molecular force fields by Pareto approach

K. Stöbener[a], P. Klein[a], S. Reiser[b], M. Horsch[b,*], K.-H. Küfer[a], H. Hasse[b]

[a] Fraunhofer Institute for Industrial Mathematics ITWM, Kaiserslautern, Germany
[b] Laboratory of Engineering Thermodynamics, University of Kaiserslautern, Germany

ARTICLE INFO

Article history:
Received 25 November 2013
Received in revised form 7 April 2014
Accepted 11 April 2014
Available online 21 April 2014

Keywords:
Pareto optimization
Force field models
Lennard-Jones fluid

ABSTRACT

Molecular force fields are widely used for simulating thermodynamic properties of fluids. In developing such force fields, usually some of their parameters are adjusted to experimental data sets, which are often of different type. The adjustment is commonly carried out by minimizing a single objective function which represents the deviations between the model and the data. In the present work, a different approach is explored. Individual objective functions are defined for each data set and a multicriteria optimization task is solved. It is explicitly acknowledged that the different objectives are usually conflicting. The multicriteria optimization problem is solved by determining the Pareto set. By definition this set includes all solutions for which no further improvement in one objective can be achieved without having to accept a decline in at least one other objective and, hence, contains best compromises. The user can then choose out of these solutions one which is particularly suited for his application. The procedure is illustrated using the parameterization of the Lennard-Jones model for argon and methane as examples. Six different objective functions are included in the optimization. They represent the deviations between the model and the following properties at boiling conditions over a wide temperature range: (a) liquid density, (b) vapor pressure, (c) enthalpy of vaporization, (d) liquid shear viscosity, (e) liquid thermal conductivity, and (f) surface tension. First single objective fits are carried out for all properties. Then Pareto sets are determined for two triples of objectives namely, (a, b, c) on one side and (d, e, f) on the other side. An unexpected topology of the Pareto set is observed and explained. Then the full Pareto set for all six properties is determined and all results are compared. They show that good results can be achieved with the simple Lennard-Jones model for the two studied fluids, even when the goal is to simultaneously describe many different thermodynamic properties. The work also illustrates the benefits of using Pareto optimization for developing force fields, and, more generally thermodynamic models.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Molecular simulations with atomistic force fields are widely used for solving problems in physics, biology, chemistry, and engineering. They rely on the availability of suitable force fields. Usually, the functional form of the force field is known based on physical grounds, but the model parameters still need to be determined. This is often done using a combination of setting parameters based on quantum chemical data and adjusting the remaining parameters to experimental data [1]. The resulting optimization problem is usually solved by minimizing a single objective function, which contains the information on the deviations between the simulation results and the experimental data. Different solver strategies can be employed to find the minimum. Faller et al. [2] and Reith et al. [3] each presented an automatic scheme based on simplex algorithms. Wang and Kollman [4] introduced an automatic engine based on systematic search as well as a genetic algorithm. Bourasseau et al. [5] used a Gauss-Legendre least-squares estimator to find the minimum of the objective function. Hülsmann et al. [6,7] compared the performance of several algorithms. Deublein et al. [8] also present an automated method for the development of force fields based on a gradient method.

By finding the minimum of the objective function one specific parameter set for the molecular model is identified, which is optimal for the chosen objective function.

The multicriteria optimization approach used in this work considers several objective functions. These can e.g. stem from using experimental data sets of different type (e.g. densities or vapor pressures) or at different conditions (e.g. liquid or gas phase). In general, the different objective functions are conflicting, i.e. they

* Corresponding author. Tel.: +49 631 205 3227; fax: +49 631 205 3835.
E-mail address: martin.horsch@mv.uni-kl.de (M. Horsch).

http://dx.doi.org/10.1016/j.fluid.2014.04.009
0378-3812/© 2014 Elsevier B.V. All rights reserved.

N	number of temperatures
O	thermodynamic property
k	Boltzmann's constant
f	objective function
p	dimension objective space
p_C	critical pressure
p^S	saturated vapor pressure
q	dimension design space
u	Lennard-Jones potential
x	decision vector
T	temperature
T_C	critical temperature
T_{TP}	triple point temperature
Δh_V	enthalpy of vaporization
δ	mean relative deviation
ϵ	Lennard-Jones energy parameter
γ	surface tension
λ'	liquid thermal conductivity
η'	liquid shear viscosity
ρ'	saturated liquid density
ρ_C	critical liquid denisty
σ	Lennard-Jones size parameter

cannot be minimized simultaneously. In the present work the Pareto approach is used for solving this multicriteria optimization problem. It relies on identifying the set of Pareto optimal solutions (Pareto set). The Pareto set represents those solutions, for which one objective function can only be improved by having to accept a decline in at least one other objective function. Hence, Pareto sets represent best compromises. Once the Pareto set is determined the user gets an overview of what can be achieved with a certain model. Based on that knowledge he can then choose from the Pareto set the model he considers to be most attractive for his application.

Pareto optimization was previously employed in the context of molecular models. Mostaghim et al. [9] describe an optimization of bond terms for primary alcohols in which three objective functions based on the reproduction of ab initio information were designed and five different force field parameters considered. To solve the problem multi-objective evolutionary algorithms and particularly multi-objective Particle Swarm Optimization were applied to determine a Pareto set.

This work is, to the best of our knowledge, the first to use Pareto optimization for determining parameters of intermolecular interaction potentials. By a brute force enumeration, we systematically study the application of Pareto-optimization for developing atomistic force fields using the one center Lennard-Jones 6–12 potential as an example. The performance of that simple model for representing thermo-physical properties of two fluids, argon and methane, is investigated. For each of these fluids six objective functions are defined which represent the deviation of the model and the experimental data sets for the following properties: liquid density, vapor pressure, enthalpy of vaporization, liquid shear viscosity, liquid thermal conductivity, and surface tension.

In a first step two three dimensional objective spaces, each considering three properties are explored. An interesting and unexpected topology is observed and explained. To complete the analysis, the full six dimensional problem is addressed. The results illustrate the benefits from using Pareto optimization for parameterizing atomistic force fields. They also show that the Lennard-Jones model performs astonishingly well for describing fluid properties of simple substances.

2. Multicriteria optimization

A multicriteria optimization problem is characterized by several objective functions $f_i(x)$ which have to be minimized simultaneously:

$$\min f(x) = (f_1(x), \ldots, f_p(x)) \in \mathbb{R}^p \tag{1}$$

They span the objective space $\mathbb{R}^p$ and depend on the decision vector $x \in \mathbb{R}^q$ where $\mathbb{R}^q$ is the design space. The solution to such a problem is a set of best compromises: For any improvement in a single objective $f_i(x)$, a decline in at least one other objective $f_j(x)$, $i \neq j$ has to be accepted. The set of all best compromises in the objective space is called the Pareto frontier. The corresponding solutions in the design space are called the Pareto set. Once they are identified, a trade-off discussion of the individual objectives is possible. Based on this knowledge a decision vector which is particularly suitable for the studied application can be chosen.

The Pareto frontier is a subset of all feasible points in the objective space and needs to be approximated by a suitable numerical strategy. The most obvious approach for the construction of the Pareto frontier starts from a construction of the set of all feasible points in the objective space and a subsequent identification of the Pareto frontier by brute force comparison of the different objectives. This strategy is employed in the present work. However for high dimensions of the design space and for computationally intensive calculations of the objective functions, a full sampling of all feasible solutions is not possible. Then multicriteria optimization algorithms, available in the literature, have to be applied to identify the Pareto set (see e.g. [10,11]). For more information on multicriteria optimization (see e.g. [12–14]).

In the context of force field parameterization the design space is spanned by the parameters describing the model. Thus we refer to it as parameter space in the subsequent text. A point in the parameter space, hence, corresponds to a certain force field parameter set and would commonly be called a model of the substance. Mapped to each point in the parameter space is one point in the objective space. The objective functions contain the information on the quality of the force field model, e.g. as represented by the mean relative deviations between the simulation results and the chosen experimental data sets. Pareto optimization crucially depends on a suitable choice of the objective function. Different approaches are possible and can individually be adapted to the thermodynamic properties, whose quality needs to be assessed.

Once the Pareto frontier is identified, full insight over the achievable model performance is gained. It is then possible to navigate on the Pareto frontier, i.e. to carry out the trade-off discussion of the individual objectives, and to finally choose the most attractive of all Pareto optimal solutions.

3. Studied systems

In this work the parameterization of the Lennard-Jones (LJ) potential

$$u(r) = -4\epsilon \left[\left(\frac{\sigma}{r}\right)^6 - \left(\frac{\sigma}{r}\right)^{12} \right] \tag{2}$$

for argon and methane is studied as an example. The studied properties are all related to vapor–liquid equilibria: saturated liquid density, saturated vapor pressure, enthalpy of vaporization, saturated liquid shear viscosity, saturated liquid thermal conductivity, and surface tension. The two fluids are studied for technical reasons: A large variety of data and many literature models are available to compare the findings of this work to.

Download English Version:

https://daneshyari.com/en/article/202433

Download Persian Version:

https://daneshyari.com/article/202433

Daneshyari.com