Journal of Molecular Graphics and Modelling 139 (2025) 109035

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

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

Fine-tuning property domain weighting factors and the objective function in force-field parameter optimization

Robin Strickstrock [a], Alexander Hagg [a], Marco Hülsmann [b], Karl N. Kirschner [b], Dirk Reith [a,*]

[a] Department of Engineering and Communication (DEC), University of Applied Sciences Bonn-Rhein-Sieg, Grantham-Allee 20, 53757 Sankt Augustin, Germany
[b] Department of Computer Science (CS), University of Applied Sciences Bonn-Rhein-Sieg, Grantham-Allee 20, 53757 Sankt Augustin, Germany

ARTICLE INFO

Keywords:
Force field models
Lennard-Jones parameters
Multiscale parameterization
Local optimization
Gradient-based optimization
Weighting factors
Objective function
Non-linear projection

ABSTRACT

Force field (FF) based molecular modeling is an often used method to investigate and study structural and dynamic properties of (bio-)chemical substances and systems. When such a system is modeled or refined, the force-field parameters need to be adjusted. This force-field parameter optimization can be a tedious task and is always a trade-off in terms of errors regarding the targeted properties. To better control the balance of various properties' errors, in this study we introduce weighting factors for the optimization objectives. Different weighting strategies are compared to fine-tune the balance between bulk-phase density and relative conformational energies (RCE), using *n*-octane as a representative system. Additionally, a non-linear projection of the individual property-specific parts of the optimized loss function is deployed to further improve the balance between them. The results show that the combined error for the reproduction of the properties targeted in this optimization is reduced. Furthermore, the transferability of the force field parameters (FFParams) to chemically similar systems is increased. One interesting outcome is a large variety in the resulting optimized FFParams and corresponding errors, suggesting that the optimization landscape is multi-modal and very dependent on the weighting factor setup. We conclude that adjusting the weighting factors can be a very important feature to lower the overall error in the FF optimization procedure, giving researchers the possibility to fine-tune their FFs.

1. Introduction

Over the past decades, computer-based molecular simulations have gained an increasingly important role in many scientific fields, for example, in material, (bio-)chemical, and pharmaceutical sciences. They are used to model structural properties or dynamic processes of diverse molecules and provide insight into the modeled molecular system at an atomistic level of detail and a fine temporal resolution. Commonly used methods are, for instance, molecular mechanic (MM) and molecular dynamic (MD) simulations. They are based on parameterized potential functions, called FF, that model the interactions between the atoms, molecules, or particles. The quality of the molecular simulations strongly depends on the used FF and the deployed FFParams. Numerous good FF, usually including dedicated FFParams for different applications, are available and have been developed over the past decades. Examples include General AMBER Force Field (GAFF) [1] for small organic molecules compatible with Amber's [2] biological force fields, the very recent Amber-consistent small molecule FF developed by Xue et al. [3], CHARMM general force field (CGenFF) [4] for bio- and drug-like molecules (an extension of the CHARMM [5,6] force field and simulation tool), and optimized potentials for liquid simulations force field (OPLS) [7,8] for liquids, organic molecules and peptides. Another well-established example is the Open Force Field Initiative (OpenFF) [9–11] using the SMIRKS-native Open Force Field (SMIRNOFF) [12] in combination with their versatile simulation toolkit called OpenMM [13]. A more detailed overview and discussion of recent FF development is provided by He et al. [14].

When optimizing FFParams, usually the bonded interactions are fitted to quantum mechanic (QM) calculations of energies, frequencies and structures, while the van der Waals (vdW) parameters are optimized towards experimental quantities like densities or heat of vaporization [7]. The electrostatic terms are typically parameterized to reproduce QM electrostatic potentials [15,16] or energies of water-molecule interactions [17]. Many of the available FFParams optimization methods are either not automated or treat the optimization objectives sequentially, instead of simultaneously. To address the latter, we introduced a multiscale optimization workflow that simultaneously optimizes Lennard-Jones parameters (LJParams) (i.e. σ_C, σ_H, ϵ_C and ϵ_H) towards different property domains (i.e. single-molecule or atomistic

* Corresponding author.
E-mail address: dirk.reith@h-brs.de (D. Reith).

https://doi.org/10.1016/j.jmgm.2025.109035
Received 26 September 2024; Received in revised form 10 December 2024; Accepted 23 March 2025
Available online 24 April 2025
1093-3263/© 2025 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

properties and multi-molecular properties) [18]. It was shown that the RCE reproduction of *n*-octane could be significantly improved with a minimal reduction in the accuracy of its bulk-phase density. Herein, a modified approach is presented that provides more control over the balancing of the optimized properties, improving the accuracy of the density calculations, while maintaining the RCE reproduction quality. This is achieved by a NLP of the individual property-specific parts of objective function and adapting their WFs. Four different approaches are investigated to determine the intra-domain WFs (i.e. those used within the same property domain). Additionally, inter-domain property balancing is introduced. Using this technique, the balance between the various properties subject to optimization is finely tuned. Both intra- and inter-domain balancing are combined with the NLP approach of the loss function's components. Note that the terminology "reweighting" in FFParams optimizations also refers to ensemble reweighting that allows to recalculate properties with altered FFParams [19–24]. This method can be very efficient, especially for large systems, where additional simulations can be very expensive and time consuming. The reweighting or rebalancing herein, however, refers to the weighting factors of the loss function and is a different approach. We present the best set of optimized FFParams, improving the previously achieved results [18] and show that different weighting factors result in different parameter sets, corresponding to different minima in the objective function landscape, indicating that FFParams optimization not only is a multiobjective but also a multi-modal problem and should be treated as such. Finally, the parameter set's transferability is tested by reproducing the surface tension, viscosity and density of *n*-hexane, *n*-heptane, *n*-octane and *n*-nonane at 293.15 K, 315.15 K and 338.15 K.

2. Methodology

To quantify the error of the parameter optimization, the following loss function $F(x)$ is defined [18]:

$$F(x) = \sum_{i=1}^{n} \omega_i^{\mathrm{MM}} \left(\frac{f_i^{\mathrm{QM}} - f_i^{\mathrm{MM}}(x)}{f_i^{\mathrm{QM}}} \right)^2 + \sum_{j=1}^{m} \omega_j^{\mathrm{MD}} \left(\frac{f_j^{\mathrm{exp}} - f_j^{\mathrm{MD}}(x)}{f_j^{\mathrm{exp}}} \right)^2 \quad (1)$$

The loss function is dependent on the force-field parameters x and includes properties from different domains (i.e. the bulk-phase density and the RCE). The simulation results $f_i^{\mathrm{MM}}(x)$ and $f_j^{\mathrm{MD}}(x)$ originate from MM minimization and MD calculation, respectively, and are dependent on the force-field parameters x. They are subtracted from the particular property target values f_i^{QM} and f_j^{exp}, which are, for example, obtained by experiments. They are divided by the target values to get the relative deviations, squared to eliminate the sign, and scaled by a weighting factor ω. Due to this, the same absolute error values can have different contributions to the overall loss, depending on the magnitudes of corresponding target values. The target values consist of f_i^{QM}, $i = 1, \ldots, n$, containing the ith RCE obtained from QM calculations (i.e. a single-molecular target property). For this work there are 96 unique RCE of *n*-octane ($n = 96$) that are obtained at an MP2.5/AVTZ theory level [25]. The f_j^{exp}, $j = 1, \ldots, m$, part contains the jth experimentally obtained bulk-phase density (i.e. a multi-molecular target property). In this work, *n*-octane's bulk-phase density at 293.15 K and 1 bar pressure is used [26]. The associated WFs are ω_i^{MM} and ω_j^{MD}, which allows control of the relative importance of the optimization objectives. The number of individual properties in each domain is denoted with n and m. Note that any number of additional properties can be added to the optimization by adding a suitable term to the loss function.

2.1. Determining the weighting factor

For fine-tuning the WFs two aspects are considered. First, the **intra-domain** distribution of the WFs and second, their **inter-domain** balancing. The former concentrates on the distribution of the WFs values for one property independently from the other properties — for example assigning values for every ω_i^{MM} without concerning the values for ω_j^{MD} (see Eq. (1)). The latter takes the balance between the overall weights of the target properties from different domains (i.e. $\sum_i \omega_i^{\mathrm{MM}}$ vs. $\sum_j \omega_j^{\mathrm{MD}}$) into account.

Herein four approaches for the intra-domain weight balancing are presented: "uniformly distributed WFs", "linearly decreasing WFs", "exponentially decreasing WFs" and "stepwise decreasing WFs" (see Fig. 1), which are discussed in the following sections alongside the inter-domain balancing approach.

Because the RCE are the energies relative to the structure with the lowest energy and, by common convention, the lowest energy conformation is assigned the relative energy value of 0. Consequently, the difference between the simulation result and the target value is always 0 for this conformation and the corresponding weight ω_1^{MM} is also set to 0 in every approach. By setting $\omega_1^{\mathrm{MM}} = 0$, it is ensured that the total amount of weight assigned for this property (i.e. $\sum_i \omega_i^{\mathrm{MM}}$) is distributed among the non-zero RCE. Note that intra-domain balancing is not applied to the density WFs, because for this property domain herein only one value is used and thus a balancing is unnecessary.

2.1.1. Uniform distribution

The uniform distribution scheme assigns the same value to each RCE weighting factor. While simplistic in its implementation, this approach does not follow the chemical reasoning that in a real bulk-phase system of *n*-octane – at a given temperature and pressure — the possible conformations will not be equally populated, and are thus not equally important for the simulation modeling. This approach was used in previous work [18] and serves as a baseline to compare the other approaches to.

The allocation of the uniform distribution WFs to *n*-octane's RCE is visualized in Fig. 1, where every factor ω_i^{MM} for $i = 2, \ldots, n$ is assigned the same value. The WFs corresponding to each of the RCE are calculated using the following equation:

$$\begin{aligned} \omega_{i,\mathrm{uni}}^{\mathrm{MM}} &= \frac{1}{(n-1)} \cdot b, \quad i = 2, \ldots, n \\ \omega_{1,\mathrm{uni}}^{\mathrm{MM}} &= 0 \end{aligned} \quad (2)$$

Thereby, n is the total number of WFs (here: $n = 96$) and b is the overall amount of weight that is assigned to the property domain (i.e. $\sum_i^n \omega_{i,\mathrm{uni}}^{\mathrm{MM}} = b$; here: $b = 0.96$). Note that b can, in general, be freely chosen and is only dependent on the inter-domain balancing, (see Section 2.1.5, which also explains why $b = 0.96$ and not just $b = 1$).

2.1.2. Linearly decreasing distribution

An alternative simplistic approach is to assign a linear relationship between the WFs and the RCE that better highlights the importance of the lower energy conformations. At room temperature and atmospheric pressure (293.15 K and 1 bar), higher-energy conformations are less likely to occur than lower-energy conformations. Thus, the WFs have been adapted to linearly decrease (i.e. $\omega_{i+1,\mathrm{lin}} = \omega_{i,\mathrm{lin}} - d$), while still preserving the overall value for all RCE WFs (i.e. $\sum_i \omega_{i,\mathrm{lin}}^{\mathrm{MM}} = 0.96$). Fig. 1 shows the WFs calculated using Eq. (3). *Note*, that the WFs are linearly decreasing w.r.t. the energy state IDs (i.e. $i = 1, \ldots, n$) and not a linear function dependent on the RCE values.

$$\begin{aligned} \omega_{n,\mathrm{lin}} &= k \\ \omega_{n-r,\mathrm{lin}} &= \omega_{n-r+1,\mathrm{lin}} + d, \quad r = 1, \ldots, n-2 \\ d &= 2 \cdot \frac{b - ((n-1) \cdot k)}{(n-2) \cdot (n-1)} \\ \omega_{1,\mathrm{lin}} &= 0 \end{aligned} \quad (3)$$

Thereby, n is the total number of WFs (here: $n = 96$) and b is the overall amount of weight that is assigned to the property domain (i.e. $\sum_i^n \omega_{i,\mathrm{lin}}^{\mathrm{MM}} = b$; here: $b = 0.96$). k is a user-chosen value for the lowest weighting factor $\omega_{n,\mathrm{lin}}$. It determines the slope of linearly decreasing WFs and how much weight is assigned to the other WFs, because the sum of all WFs cannot exceed b. Furthermore, $k > 0$ and is chosen such, that the WFs clearly differ from the uniformly distributed WFs (here: $k = 0.001$).

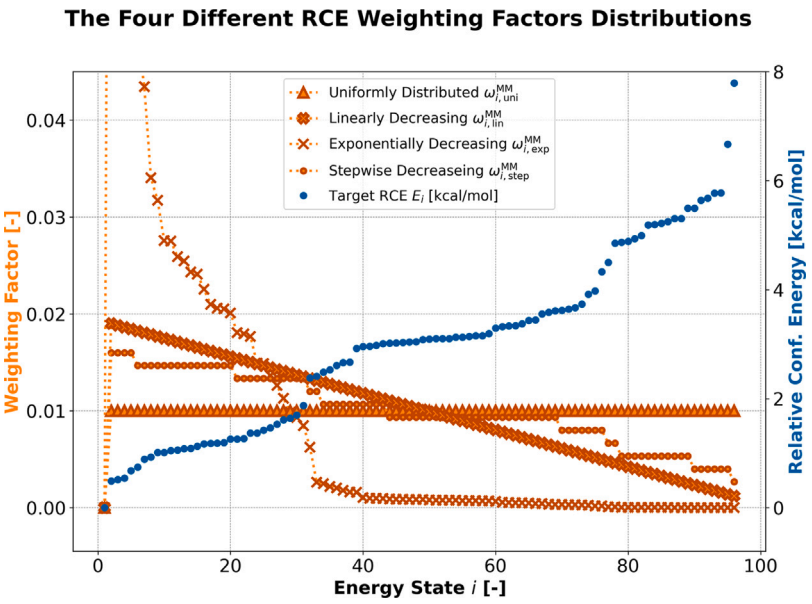

Fig. 1. Examples for the four different intra-domain balancing approaches for the RCE WFs (orange graphs). Additionally, for every ω_i^{MM} the corresponding RCE is plotted (blue). Note that the highest values for $\omega_{i,exp}^{MM}$ are cut off because this figure aims to illustrate the trend of the WFs. For that purpose the exact values of every ω_i^{MM} have a low priority and, additionally, can vary based on the optimization problem and inter-domain balancing setup. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2.1.3. Exponentially decreasing distribution

The exponentially decreasing WFs for the RCE follows an idea similar to the linearly decreasing WFs. In contrast, however, this approach is based on a Boltzmann distribution (see Eq. (4)) and is visualized in Fig. 1. This approach attempts to mimic the probability that a conformation occurs based on its internal energy in combination with the system's temperature. Thus, its importance for the parameter optimization is based on those properties.

$$\begin{aligned}\tilde{\omega}_{i,\mathrm{exp}}^{\mathrm{MM}} &= \mathrm{e}^{-E_i(k_\mathrm{B}\cdot T)^{-1}}\\ \omega_{i,\mathrm{exp}}^{\mathrm{MM}} &= \tilde{\omega}_{i,\mathrm{exp}}^{\mathrm{MM}}\cdot\frac{b}{\sum_i^n \tilde{\omega}_{i,\mathrm{exp}}^{\mathrm{MM}}},\quad i=2,\ldots,n\\ \omega_{1,\mathrm{exp}}^{\mathrm{MM}} &= 0\end{aligned}\tag{4}$$

Thereby, E_i is the ith RCE (Note, that for E_i the units are converted from kcal/mol to joule), $k_\mathrm{B} = 1.38064852\cdot 10^{-23}\,\mathrm{m^2\,kg\,s^{-2}\,K^{-1}}$ (Boltzmann's constant) and T the temperature (here: $T = 293.15\,\mathrm{K}$). The number of WFs is n (here: $n = 96$) and b is the overall amount of weight that is assigned to the property domain (i.e. $\sum_i^n \omega_{i,\mathrm{exp}}^{\mathrm{MM}} = b$; here: $b = 0.96$). After calculating the unscaled WFs $\tilde{\omega}_{i,\mathrm{exp}}^{\mathrm{MM}}$ they are scaled such that $\sum_i^n \omega_{i,\mathrm{exp}}^{\mathrm{MM}} = b$, resulting in the actual WFs $\omega_{i,\mathrm{exp}}^{\mathrm{MM}}$.

2.1.4. Stepwise decreasing distribution

The underlying idea for a stepwise decreasing distribution of the WFs for the RCE is that room temperature introduces an energy of approximately $0.6\,\mathrm{kcal\,mol^{-1}}$ to a system. Thus, the RCE are grouped into multiples of $0.6\,\mathrm{kcal\,mol^{-1}}$ and every subsequent group, containing energetically higher RCE than the previous group, are weighted less because they are less likely to occur. They are calculated following Eq. (5) and are visualized in Fig. 1.

To calculate the stepwise decreasing WFs, first, the set $\mathcal{E}$ containing all RCE $e_1,\ldots,e_n$ is divided into subsets $\mathcal{A}_1,\ldots,\mathcal{A}_p$ containing the energies that are less or equal than multiples of the stepsize k (here: $k = 0.6\,\mathrm{kcal\,mol^{-1}}$):

$$\begin{aligned}\mathcal{E} &= \{e_2,\ldots,e_n\} = \cup_{q=1}^{p}\mathcal{A}_q\\ \mathcal{A}_q &= \{e_i\in\mathcal{E}\,|\,(q-1)\cdot k < e_i \le q\cdot k\},\quad i=2,\ldots,n\end{aligned}$$

Where p is the number of different subsets $\mathcal{A}$. For every RCE e_i, the corresponding unscaled weighting factor $\tilde{\omega}_{i,\mathrm{step}}^{\mathrm{MM}}$ is assigned based on the subset $\mathcal{A}_q$ the RCE e_i is an element of. Thereby, the WFs corresponding to the RCE with the lowest energy are assigned the highest value:

$$\tilde{\omega}_{i,\mathrm{step}}^{\mathrm{MM}} = p-(q-1),\quad \forall i : e_i\in\mathcal{A}_q$$

Finally, the WFs $\tilde{\omega}_{i,\mathrm{step}}^{\mathrm{MM}}$ are scaled such that their sum equals the overall amount of weight that is assigned to the property domain b (here: $b = 0.96$), resulting in the scaled WFs:

$$\begin{aligned}\omega_{i,\mathrm{step}}^{\mathrm{MM}} &= \frac{\tilde{\omega}_{i,\mathrm{step}}^{\mathrm{MM}}}{\sum_i^n \tilde{\omega}_{i,\mathrm{step}}^{\mathrm{MM}}}\cdot b\\ \omega_{1,\mathrm{step}}^{\mathrm{MM}} &= 0\end{aligned}\tag{5}$$

2.1.5. Inter-domain balancing

In the previous work [18] the summed values of the WFs of the two domains (i.e. $\sum_{i=1}^n \omega_i^{\mathrm{MM}}$ vs. $\sum_{j=1}^m \omega_j^{\mathrm{MD}}$) were equally weighted, putting no emphasis on either domain. In order to improve the force field for both domains, in this work different inter-domain weights are evaluated to re-balance the loss function. For example, by setting $\sum_{j=1}^m \omega_j^{\mathrm{MD}} = 1.44$ (that herein simplifies to $\omega^{\mathrm{MD}} = 1.44$) and $\sum_{i=1}^n \omega_i^{\mathrm{MM}} = 0.48$, a re-balancing is reached that assigns 75% of the overall weight to the density, and thus is called "75-25". The different intra-domain balancing setups used herein are listed in Section 2.4.

2.2. Non-linear projection of the loss function

A difficulty in this simultaneous multiscale optimization approach is that the loss function values for the different target domains might be poorly comparable. For example, the error for the RCE ($\mathrm{err_{RCE}} \approx 10\%\ldots100\%$) is approximately one to two orders of magnitude higher than the error for the density ($\mathrm{err_{dens}} \approx 0.5\%\ldots15\%$) and while an error of approximately 10% for the RCE is acceptable, it would be too large for the density, rendering the parameter sets unusable for reproducing experimental results.

The loss function therefore needs to have different sensitivities in the two domains. To address this problem, the loss function value for every single property is projected to the closed interval [0, 1], herein called non-linear projection (NLP), before being further processed in

the loss function. The function z^* is used to convert the original loss function value z:

$$z^*(c, z) = 1 - e^{(-c \cdot z)} \tag{6}$$

The hyperparameter $c > 0$ controls the sensitivity of how z^* responds to changes in the loss function values z. For z the original loss function values of the different properties from Eq. (1) are used:

$$z_i^{\mathrm{MM}}(x) = \left(\frac{f_i^{\mathrm{QM}} - f_i^{\mathrm{MM}}(x)}{f_i^{\mathrm{QM}}} \right)^2 , \; z_j^{\mathrm{MD}}(x) = \left(\frac{f_j^{\mathrm{exp}} - f_j^{\mathrm{MD}}(x)}{f_j^{\mathrm{exp}}} \right)^2$$

Combining the NLP (see Eq. (6)) with the loss function $F(x)$ (see Eq. (1)) gives the resulting NLP loss function $F^*(x)$:

$$F^*(x) = \sum_{i=1}^{n} \omega_i^{\mathrm{MM}} \left(z_i^{*,\mathrm{MM}}(x) \right) + \sum_{j=1}^{m} \omega_j^{\mathrm{MD}} \left(z_j^{*,\mathrm{MD}}(x) \right) \tag{7}$$

2.3. Determining the sensitivity hyperparameter c

The hyperparameter c is used to control the sensitivity of the NLP loss function (see Section 2.2). Since we already established that different domains need different loss sensitivity, how actual values for c can be obtained is shown in the following examples for the density and RCE.

In general, c needs to be adjusted such that an improvement in the simulation results yields a lower loss function value z^*, while a worsening yields a higher value (i.e. $\frac{\mathrm{d}}{\mathrm{d}z} z^*(c, z) \neq 0$). This is required because otherwise the gradient-based optimization algorithm cannot calculate a descent direction to update the parameter set and suggest a better solution. Additionally, c must be chosen such that the loss function value for unwanted results does not fall below the threshold that causes the optimization to stop. Note, that we cannot give a general recommendation for a suitable value for c because the desired loss-function behavior varies for every property depending on the expected and accepted error ranges and how sensitive the property is to changes in the FFParams. Thus, the user has to assess individual property-specific thresholds, which depend both on the acceptable error and the shape of the loss landscape around the optimal point. However, following the hereinafter described procedures to determine $c_{density}$ and c_{energy} can help to assess c for other properties.

Analyzing various values for c with respect to the density, showed that $c_{\mathrm{density}} = 1000$ induce the aforementioned desired behavior. In Fig. 2 the loss function values, using the NLP loss function, for a hypothetical and unsuitable $c_{\mathrm{density}} = 10$ and a suitable $c_{\mathrm{density}} = 1000$ are shown (solid and dashed orange lines). The gray line shows the corresponding original loss function values (see Eq. (1)); Note, that those values can exceed 1, although this cannot be seen in the figure. Higher c_{density} values result in a narrower area where $z^* < 1$, and a steeper slope results in the transition areas. When c is too large, high-loss function values can overlap with the range of acceptable errors (blue area), prohibiting the optimization algorithm from terminating. However, when c is too low, unwanted optimization results are assigned a low loss function value, erroneously causing the optimization to finish. For the optimization performed herein, the density sensitivity hyperparameter is set to $c_{\mathrm{density}} = 1000$. For that value, a well-balanced configuration between the aforementioned aspects is achieved.

The same scheme and rules need to be applied when determining c_{energy}. In Fig. 3 (subfigure (b)), it is shown that using $c_{\mathrm{energy}} = c_{\mathrm{density}} = 1000$ would result in the same loss function values for the RCE before and for an intermediate iteration of the optimization (black and gray lines, respectively). This is a problem because when the loss function value of a given parameter set $F(x)$ (e.g. the loss function values for the initial RCE; black graph) is the same as or very similar to another value $F(x + h)$ (e.g. the loss function values for the RCE of an intermediate iteration of the optimization, when the force-field parameters x are changed by a small amount h; gray graph), the resulting gradient $\nabla F(x) = 0$ and no descent direction can be recognized. Analyzing various values for c with respect to the RCE showed that by choosing $c_{energy} = 1$, the aforementioned problem does not occur and yields the desired behavior regarding the reproduced RCE and loss function values (see Fig. 3, subfigures (a) and (c), respectively).

Table 1
Simulation configurations and evaluated observables for the transferability test. The properties are surface tension (γ), viscosity (η) and density (ρ).

System	T [K]	t_{sim}	Ensemble	Property
$C_6H_{14} + H_2O$, $C_7H_{16} + H_2O$, $C_8H_{18} + H_2O$, $C_9H_{20} + H_2O$	293.15 315.15 338.15	1 ns	NVT	γ
C_6H_{14}, C_7H_{16}, C_8H_{18}, C_9H_{20}	293.15 315.15 338.15	10x 2 ns	NVT	η
C_6H_{14}, C_7H_{16}, C_8H_{18}, C_9H_{20}	293.15 315.15 338.15	30 ns	NPT	ρ

2.4. Simulation tools and setup

The optimization workflow used herein is called Force FieLd Optimization Workflow (FFLOW), which is a modular optimization toolbox containing structures and methods introduced by Gradient-based Optimization Workflow (GrOW) [27]. The FFLOW toolbox is designed such that optimization objectives and targets can easily be altered or added. This allows the user to optimize FFParams for their molecules of interest, based on their individual selection of optimization targets. The FFLOW software files, scripts for parameter conversion (Gromacs to Amber), steering scrips that manage, for example, the queuing system, used input files, and all simulation configurations are available on GitHub via [28]. The density calculations were carried out using Gromacs [29,30], while the RCE were computed by Amber [2].

The following fourteen inter-domain weight distribution setups are investigated, each applied to the four intra-domain balancing approaches: “0-100”, “05-95”, “10-90”, “15-85”, “20-80”, “25-75”, “40-60”, “50-50”, “60-40”, “80-20”, “85-15”, “90-10”, “95-05” and “100-0”. For all of those (i.e., $14 \cdot 4$) setups two different loss functions are studied, the original and the NLP loss function (see Section 2.2). In total, 112 (i.e., $14 \cdot 4 \cdot 2$) different setups are investigated.

The surface tension, viscosity and density for the transferability test (see Section 3.5) are calculated using Gromacs [29,30]. Table 1 shows the combination of substances, properties and temperatures that are calculated, including the ensembles used. The input files and simulation configuration are also available on GitHub via [28].

Note, that the MM calculations for reproducing the RCE are deterministic, and for the MD simulations of the density, equilibrated systems are used that were initialized with the identical “random” seeds. Thus, all calculation results have a very low variance, as shown in previous work [18], and thus are not statistically repeated or analyzed.

3. Results

In this section, the results for the original loss function (Eq. (1)), followed by the results using the NLP loss function (Eq. (7)) are presented, including every combination of inter- and intra-domain balancing used. Finally, the best overall optimized parameter set is presented.

3.1. Original loss function

The errors of the optimization objectives (i.e. density and RCE) using the original loss function (Eq. (1)) are visualized in Fig. 4. The errors for the density (blue line) and RCE (orange line) are shown depending on the overall weight that is assigned to the density weighting factor (i.e. the inter-domain balancing). Consequently, an x-axis value

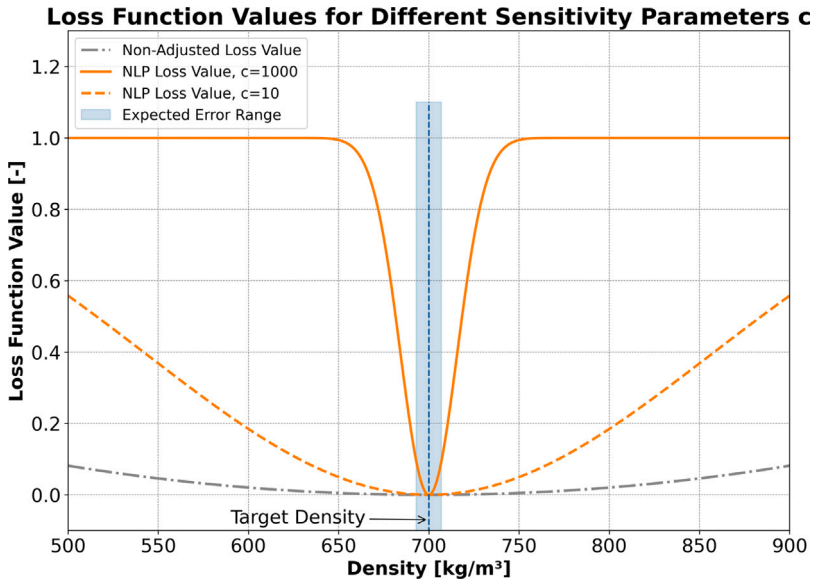

Fig. 2. Impact of sensitivity hyperparameter c_{density} on the NLP loss function value (solid and dashed orange lines). The vertical blue line marks the target density and the blue area is the acceptable error for the optimization target (here: 1%). The original loss function value is visualized by the gray graph. Note, that the latter can exceed 1 (not shown in this figure). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

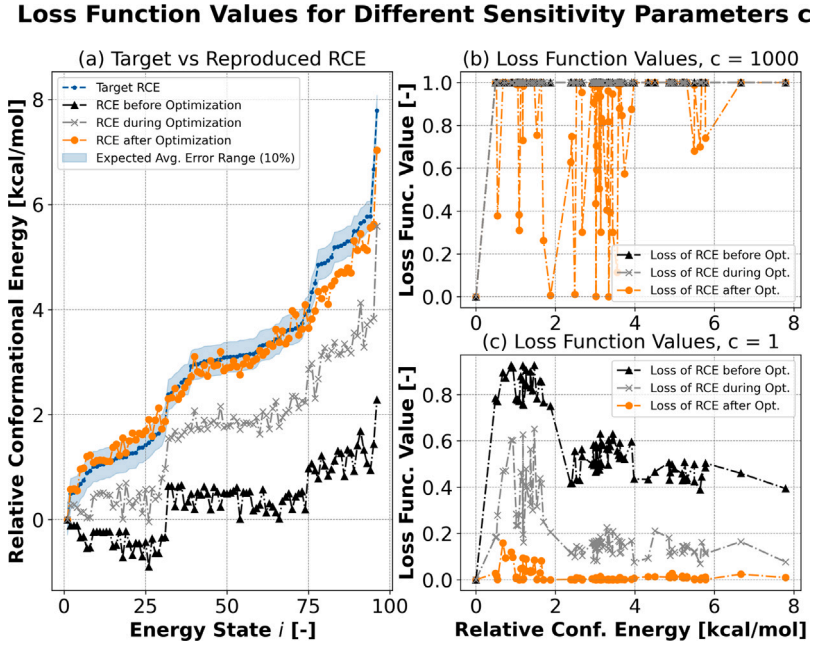

Fig. 3. Subfigure (a): the reproduced RCE before (black graph), for an intermediate iteration of the optimization (gray graph) and after (orange graph) the optimization compared to their target values (blue graph), including a 10% error range (blue area).
Subfigure (b): Loss function values corresponding to the RCE before, for an intermediate iteration of the optimization and after the optimization (black, gray and orange graphs, respectively) for $c_{\text{energy}} = 1000$. Note, that the gray and black graphs overlap.
Subfigure (c): Loss function values corresponding to the RCE before, for an intermediate iteration of the optimization and after the optimization (black, gray and orange graphs, respectively) for $c_{\text{energy}} = 1$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

of 20% stands for the "20-80" balancing where 80% of the total weight is distributed among the RCE weighting factors and 20% is assigned to the density weighting factor. The subfigures show the data for the four different intra-domain weighting factor balancing strategies.

For every setup, the density and RCE error curves intersect in the region where approximately 95% to 99% of the weight is assigned to the density. The location of the intersection is an interesting characteristic because it is an indicator for a well-balanced weighting factor distribution. Note that this does not necessarily correspond to the lowest combined error (i.e. the sum of the density and RCE errors). The intersections around "40-60" in Fig. 4(a) are caused by the optimization algorithm being stuck in a local minimum, which is a general problem for local optimization algorithms. The corresponding resulting parameter set also differs from the other resulting parameter sets (see tables 3, 4, 5 and 6 in section 5). Generally, the resulting parameter sets do not show a clearly favorable parameter combination (see Fig. 5). Some combinations of the force-field parameters σ_{H}, ϵ_{C} and partially σ_{C} occur more frequently than others, but other combinations can result in similar or even lower loss function values.

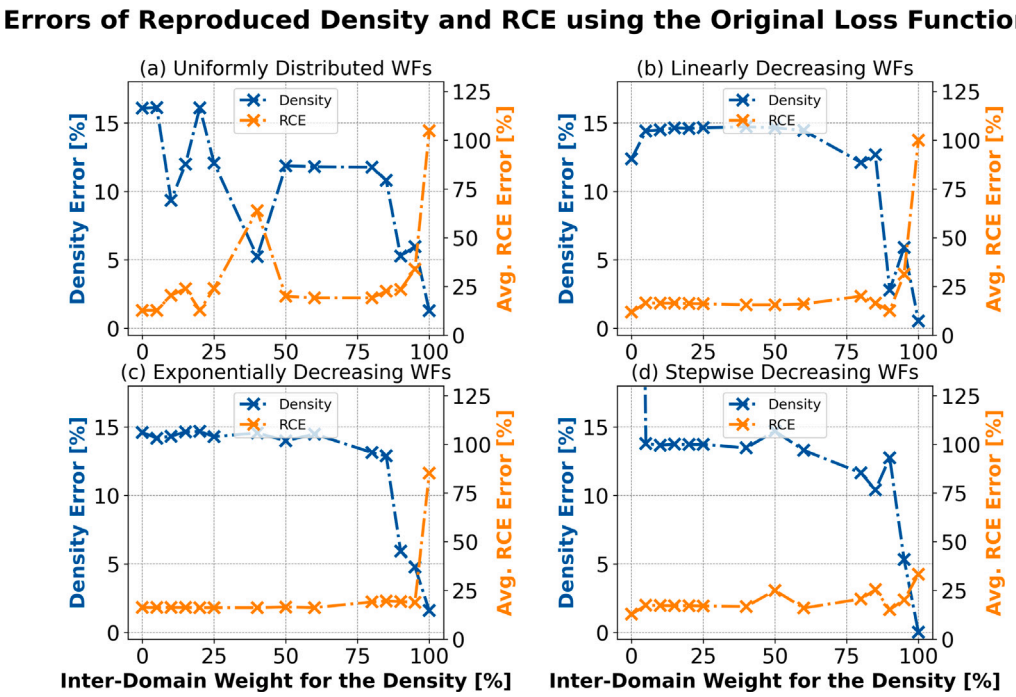

Fig. 4. Errors for the reproduction of the density (blue) and RCE (orange). The FFParams optimizations are performed using the original loss function and all combinations of inter- and intra-domain weighting factor balancing approaches (figures (a) – (d)). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

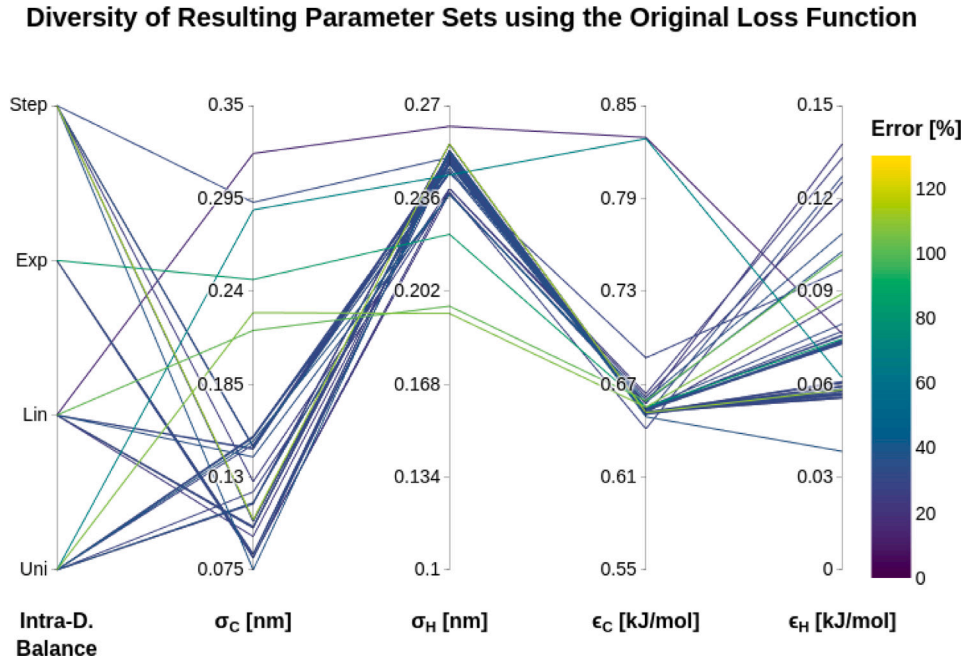

Fig. 5. Resulting Lennard-Jones parameters (i.e. 56 optimized parameter sets) using the original loss function and all combinations of inter- and intra-domain balancing approaches.

3.2. NLP loss function

Optimizing the Lennard-Jones parameters using the NLP loss function (Eq. (7)) shifts the intersection points of the graphs to a balancing where approximately 25% to 30% of the overall weight is assigned to the density weighting factor (see Fig. 6). All resulting parameter sets and their corresponding errors are listed in tables 7, 8, 9 and 10 in section 5.

On average, the combined errors (i.e. the sum of the density and RCE errors) using the NLP loss function are higher than the combined errors using the original loss function. However, the optimizations resulting in the lowest combined error are predominately achieved when applying the NLP to the loss function (see Fig. 7). Thus, we recommend to initially perform the optimization without the use of a NLP, and subsequently implement it if the optimization results do not meet the desired quality.

Similar to the resulting parameters using the original loss function, there is an even larger diversity of the optimization results (see Fig. 8). This is more evidence that the loss landscape is highly multi-modal

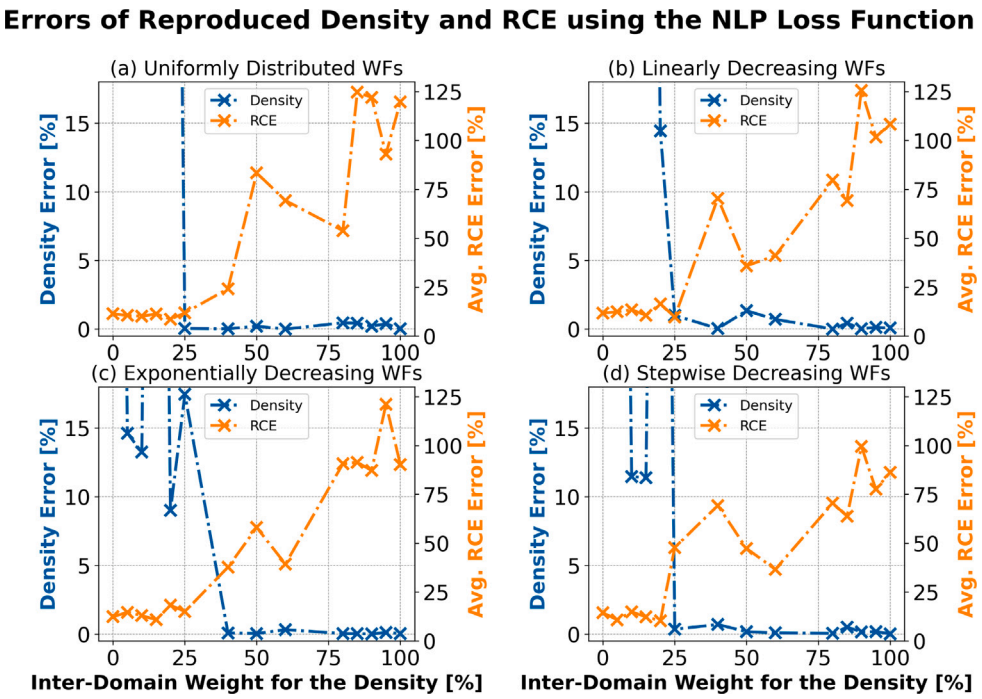

Fig. 6. Errors for the reproduction of the density (blue) and RCE (orange). The FFParams optimizations are performed using the NLP loss function and all combinations of inter- and intra-domain weighting factor balancing approaches (figures (a) – (d)). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

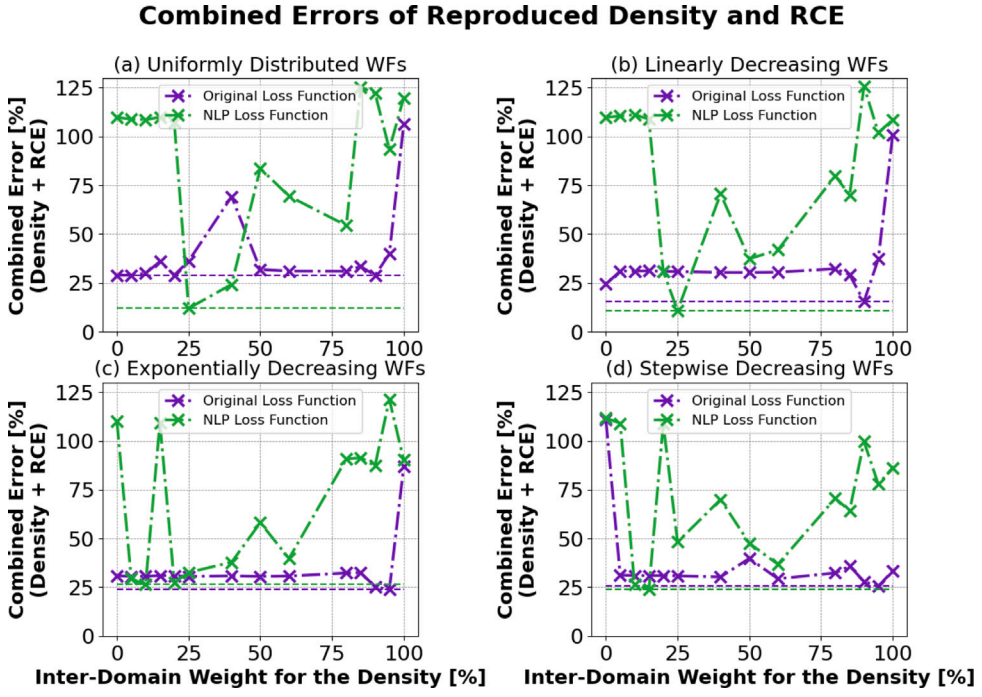

Fig. 7. Comparison of the combined errors using the original loss function (purple) and the NLP loss function (green). All combinations of inter- (on the x-axis) and intra-domain (figures (a) – (d)) weighting factor balancing approaches are used. The lowest combined errors for the setups using the original and NLP loss function are marked with a horizontal dashed line. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

and gradient-based optimization can get stuck in wildly different local optima even depending on the weighting strategy in the loss function.

3.3. Intra-domain balancing

Comparing the performances of the intra-domain balancing approaches shows a difference between the "uniformly distributed" and the other approaches for the original loss function (see Fig. 4), while there is less difference between them for the NLP loss function (see Fig. 6). We conclude that the re-balancing of the intra-domain weighting factors has a positive impact on the optimization results. However, the effect of the NLP of the loss function is stronger, softening the effect of the intra-domain weighting.

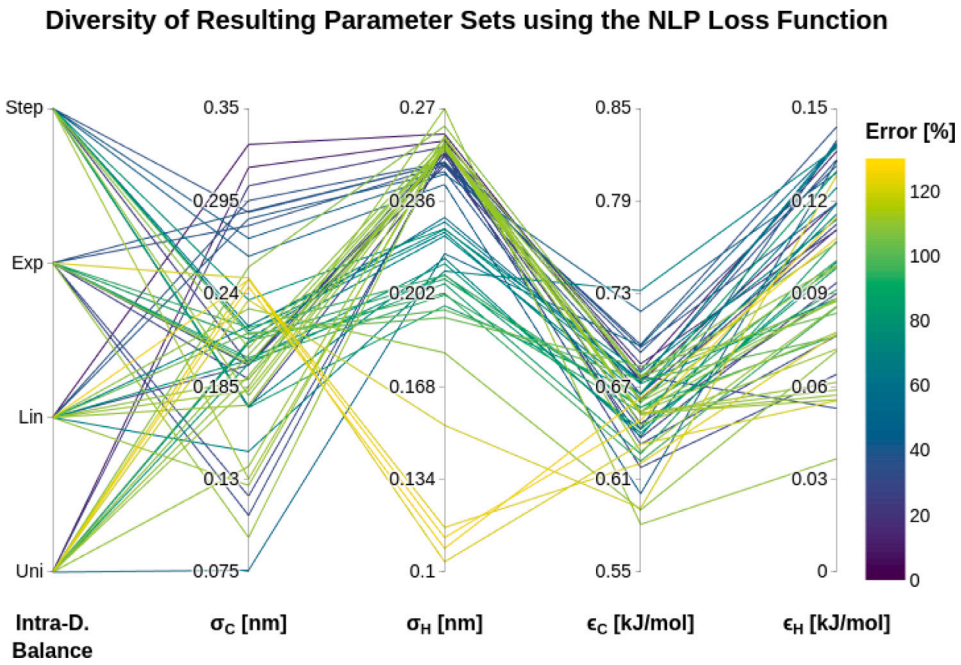

Fig. 8. Resulting Lennard-Jones Parameters (i.e. 56 optimized parameter sets) using the NLP loss function and all combinations of inter- and intra-domain balancing approaches.

Table 2
Resulting optimized parameter set with the lowest combined error. The setup optimization used for this result combines the NLP loss function, linearly decreasing WFs for the RCE and an inter-domain balancing of 25% for the density and 75% for the RCE.

σ_C [nm]	σ_H [nm]	ϵ_C [kJ/mol]	ϵ_H [kJ/mol]	err_{dens}	err_{ener}
0.3286	0.2606	0.6730	0.1194	1.0%	9.75%

3.4. Best optimized parameter set

The best overall result for the combined error of density and RCE yields the optimization using the NLP loss function, linearly decreasing WFs for the RCE and an overall balancing between density and RCE WFs of 25% to 75% (i.e. the "25-75" setup), respectively. The resulting FFParams and their corresponding errors, err_{dens} for the density and err_{ener} for the RCE, is shown in Table 2.

The reproduced density is $(707.0 \pm 4.7)\frac{kg}{m^3}$, leading to an error of $1.0\% \pm 0.67\%$. For the herein used target density of $0.7\frac{g}{cm^3}$ [26], that is converted to $700\frac{kg}{m^3}$, no error range is provided in the data's source [31]. However, other literature, for example the Chemistry Webbook [32] of the National Institute of Standards and Technology (NIST), report an estimated error of 0.1% to 0.3% for their density values [33,34]. Therefore, the herein used target density is assumed to have an error less or equal to 0.3%.

The average error of the RCE is 9.75% (for the individual RCE values see table 11 and figure 12 in section 5). The average error for the RCE is calculated by adding all individual errors and dividing the sum by the amount of RCE. If the (absolute) error for the target RCE based on MP2.5/AVTZ's reproduction of CCSD(T) results (Table 2 of Ref. [25]) is estimated, the upper limit was $-0.012\,kcal\,mol^{-1}$ as computed for *n*-hexane (i.e., 12 conformers). This results in an average relative error of -0.91% and the herein achieved results and error ranges are not as good as those results and error ranges. However, considering the different theories (i.e. MM vs. CCSD(T) or MP2.5/AVTZ calculations) and the required time to compute the results, the average relative error range of 9.75% achieved herein is acceptable. It is worth noting that this level of error is similar when comparing 96 Density Functional Theory (DFT) (B2K-PLYP-D) computed octane values to MP2.5/AVTZ values, which has an error of $+0.19 kcal\,mol^{-1}$ (i.e., 9.43%) [25].

In total, the previously achieved results [18] are improved. The average error of the RCE increased from 7.85% to 9.75%, but the error for the density is decreased substantially from $-4.14\% \pm 0.04\%$ to $1\% \pm 0.67\%$.

3.5. Transferability

The transferability is tested by using the resulting parameter set (see Table 2) for the reproduction of the properties, systems and temperatures that are also used for the transferability test in Ref. [18], allowing for a direct comparison to previous work. Specifically, the properties are the surface tension, viscosity and density for the substances *n*-hexane, *n*-heptane, *n*-octane and *n*-nonane at the temperatures 293.15, 315.15 and 338.15 K. The MD setups are listed in Section 2.4 and all input files are available on GitHub via [28].

For the surface tension there is no noticeable change in error for its reproduction perceivable, considering the calculations' uncertainties (see Fig. 9 or table 12 in Supplementary Information, for more details). Only the error for *n*-heptane at 293.15 K slightly increased. Even though many errors improved, the uncertainty is too high to claim that the transferability with respect to the surface tension improved. However, it is also not worse than before.

The reproducibility of the viscosity improved for almost all systems (see Fig. 10 or table 13 in Supplementary Information, for more details). However, the remaining error for the re-optimized parameters is still too large to provide good transferability for the calculation of the viscosity.

For the reproduction of the densities almost all calculations improved noticeably in accuracy (see Fig. 11 or table 14 in Supplementary Information, for more details), being much closer to the target values. Only the error for *n*-nonane at 293.15 K increased compared to the previous results.

4. Discussion

For the original loss function, the graphs in Fig. 4 show that the most balanced inter-domain weighting factor distribution is reached when most of the weight (~ 95%) is assigned to the density. However, for this setup, the errors for the density reproduction are large with

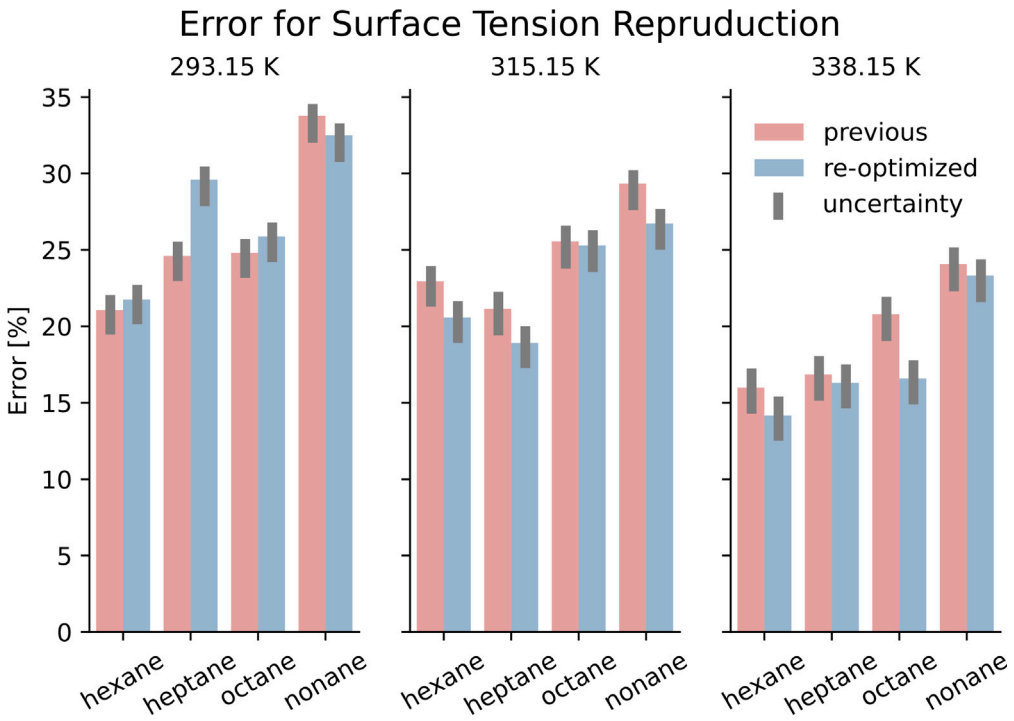

Fig. 9. Errors and uncertainties for the reproduction of the surface tension using the previously optimized [18] (red bars) and herein re-optimized (blue bars) parameter sets ($x_{p,prev}$ and $x_{p,opt}$, respectively). Note that the absolute error values are shown without a sign. The target values the errors refer to can be found in [35]. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

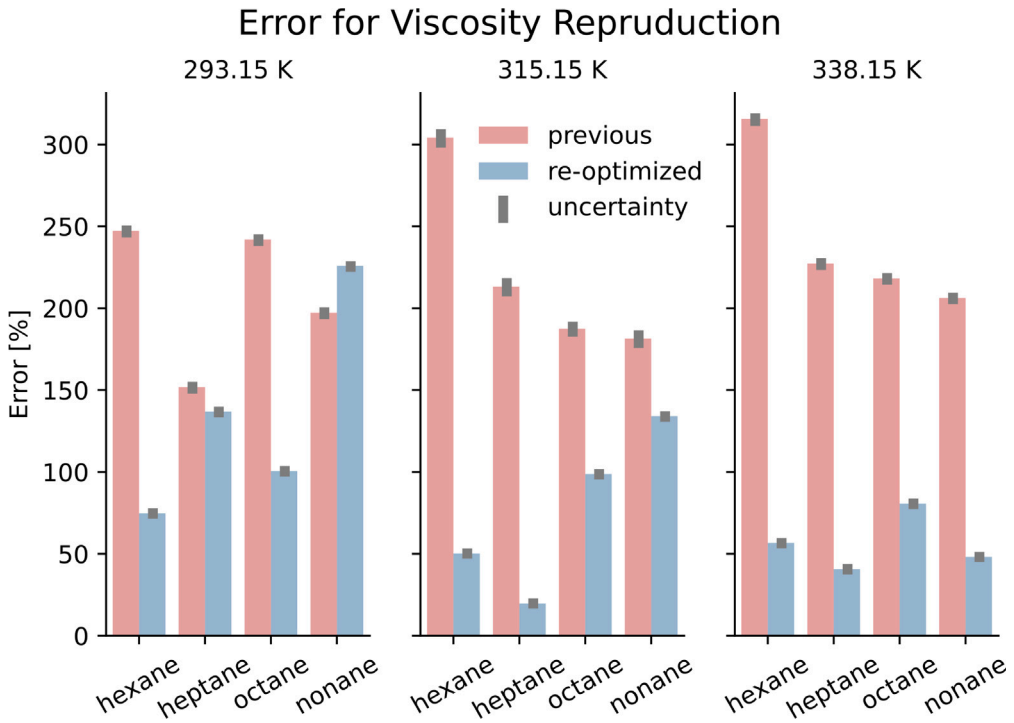

Fig. 10. Errors and uncertainties for the reproduction of the viscosity using the previously optimized [18] (red bars) and herein re-optimized (blue bars) parameter sets ($x_{p,prev}$ and $x_{p,opt}$, respectively). Note that the absolute error values are shown without a sign. The target values the errors refer to can be found in [36–39]. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the best values lying between 2.80% to 5.90% (see tables 3, 4, 5 and 6 in section 5). Further analyses of the loss function suggest that applying a non-linear projection (NLP) to the individual parts of the loss function is beneficial, increasing the comparability of the particular optimization objectives. The NLP ensures that the loss function values for the particular objectives have the same order of magnitude, once the sensitivity hyperparameter c is configured appropriately.

Using the NLP loss function shifts the most balanced inter-domain weighting factor distribution to a setup with $\sim$ 25% of the overall weight assigned to the density (see Figs. 6 and 7 as well as tables 7,

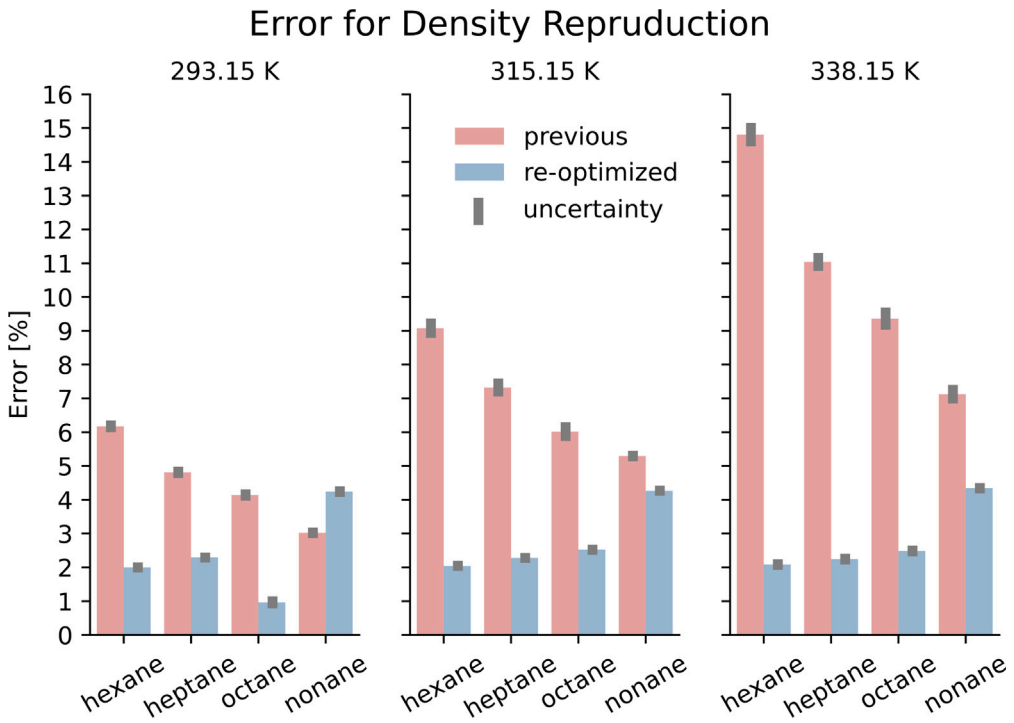

Fig. 11. Errors and uncertainties for the reproduction of the density using the previously optimized [18] (red bars) and herein re-optimized (blue bars) parameter sets ($x_{p,prev}$ and $x_{p,opt}$, respectively). Note that the absolute error values are shown without a sign. The target values the errors refer to can be found in [36–40]. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

8, 9 and 10 in section 5). The NLP works as expected and improves the comparability of the errors (i.e. the loss function values) of the different target property domains. Furthermore, the best optimization result is achieved by using the NLP loss function. However, this method introduces an additional layer of complexity represented by the sensitivity hyperparameter c. It needs to be determined and configured for every optimization objective separately and cannot be generalized. This requires a deeper insight into that specific optimization objective, which might not be a problem for experienced users but is oppositional to a user-friendly, fast and automatized force field optimization workflow.

The transferability tests showed no change in the error for the reproduction of the surface tension. The error for the viscosity calculations decreased, but is still too high to be called transferable for that property. Note, that both of the aforementioned properties were not among the optimization targets and are a different type of property than the herein used optimization targets (i.e. transport properties vs. structural and thermodynamic properties, respectively). The error for three properties at 293.15 K slightly increased and are better with the previously optimized parameters. For all other systems the errors of the reproduced properties are similar or improved. Here, the overall transferability of the FFParams is improved, especially for the property that is among the optimization targets (i.e. the density). By including other types of properties (e.g. transport properties like the viscosity) to the current thermodynamic and molecular properties and adding more and longer molecules (e.g. nonane, decane, ..) to the targeted systems the transferability could probably be further increased.

Furthermore, the diversity of the optimization results (see Figs. 5 and 8) shows the multi-modal nature of the loss function's landscape. Almost every optimization run results in a different FFParams set, depending on the optimization setup. A major disadvantage of this local optimization algorithm is that it can get stuck in a local minimum of the optimization landscape, unable to discover other minima that might be representatives of more suitable parameter sets. A consequence is that the current optimization workflow is very dependent on the initial parameter set. This might be a small issue when an already well-performing force field is re-optimized, but can be challenging for newly investigated models of molecules. Note, that herein we focused on improving the previously achieved results where different initial guesses were already explored [18].

5. Conclusion

This study's goal is to reduce the reproduction error of the density, arisen in a previous study [18], while maintaining the accuracy of the RCE reproduction. This goal is accomplished by adapting the particular WFs and the objective function used for this multiscale optimization process. The error for the density reproduction is decreased by a considerable amount, while the error for the reproduction of the RCE remains at approximately the same level.

Compared to the previously achieved results, the error regarding the bulk-phase density is improved from $(-4.14 \pm 0.04)\%$ to $(1.0 \pm 0.67)\%$, which is an acceptable error considering experimental uncertainties of 0.1% to 0.3%, keeping in mind that MD and MM calculations are mathematical models based on several assumptions and simplifications. The average error for the RCE reproduction, however, slightly increased from 7.85% to 9.75%. This is still approximately the same quality of results. This error range is not as good as the target data's average error of -0.91%, however, considering the necessary time to compute the RCE using the different modeling theories, an average relative error of 9.75% is acceptable, especially compared to their average reproduction error of $\sim 100\%$ prior to the multiscale optimization, this is a substantial improvement. It is also worth noting that this error level is similar to that seen when comparing DFT (B2K-PLYP-D) to MP2.5/ATZ computed values. Additionally, the transferability of the FFParams increased when using them for the density (i.e. a property that is part of the optimization targets) but similar systems and different temperatures.

To tackle the challenge of the loss function's multi-modality and diversity of optimization results for future development in automated custom force-field parameter optimization, we propose to use a global optimization approach. Additionally, the current optimization scheme

yields a single resulting parameter set. However, the different optimization setups show that there are different parameter sets that might be promising, depending on the intended use case. Furthermore, the two optimization objectives used herein seem to be contrary to one another because, if there is an improvement for one property, often the error of the other increases. Thus, for further development, we additionally plan to use a multi-objective optimization with multiple solutions. Prospectively designing the sensitivity parameter c to be adaptive and change during the optimization according to the shape of the loss landscape and acceptable errors.

CRediT authorship contribution statement

Robin Strickstrock: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Alexander Hagg:** Writing – review & editing, Visualization, Formal analysis. **Marco Hülsmann:** Writing – review & editing, Visualization, Formal analysis. **Karl N. Kirschner:** Writing – review & editing, Visualization, Supervision, Resources, Funding acquisition, Conceptualization. **Dirk Reith:** Writing – review & editing, Visualization, Supervision, Resources, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors thank Rudolf Berrendorf and Javed Razzaq for their continuous development and support of the BRSU's computing cluster. All of the calculations were performed on this high-performance computing cluster of the Bonn-Rhein-Sieg University of Applied Sciences, located in Sankt Augustin, Germany. The BRSU's computer hardware was supported by the Federal Ministry for Education and Research and by the Ministry for Innovation, Science, Research, and Technology of the state Northrhine-Westfalia (research grant 13FH156IN6). This work was partially funded by the Ministry of Culture and Science of the State of North Rhine-Westphalia, grant FKZ 005-2302-0023. Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) Project-ID 528562393 - FIP 26. The authors thank Prof. Dr. Wolfgang Imhof for the fruitful discussions and Yoorim Gye for her support of the visual data analysis.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jmgm.2025.109035.

Data availability

We have shared the link to the GitHub repository containing Software and Data in the Manuscript.

References

[1] J. Wang, R. Wolf, J. Caldwell, P. Kollman, D. Case, Development and testing of a general amber force field, J. Comput. Chem. 25 (2004) 1157–1174, https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.20035.

[2] D. Case, T. Cheatham, T. Darden, H. Gohlke, R. Luo, K. Merz, A. Onufriev, C. Simmerling, B. Wang, R. Woods, The Amber biomolecular simulation programs, J. Comput. Chem. 26 (2005) 1668–1688, http://dx.doi.org/10.1002/jcc.20290.

[3] B. Xue, Q. Yang, Q. Zhang, X. Wan, D. Fang, X. Lin, G. Sun, G. Gobbo, F. Cao, A. Mathiowetz, B. Burke, R. Kumpf, B. Rai, G. Wood, F. Pickard, J. Wang, P. Zhang, J. Ma, Y. Jiang, S. Wen, X. Hou, J. Zou, M. Yang, Development and comprehensive benchmark of a high-quality AMBER-consistent small molecule force field with broad chemical space coverage for molecular modeling and free energy calculation, J. Chem. Theory Comput. (2024) 799–818, http://dx.doi.org/10.1021/acs.jctc.3c00920, PMID: 38157475.

[4] K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, A. Mackerell Jr., CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields, J. Comput. Chem. 31 (2010) 671–690, https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.21367.

[5] B. Brooks, R. Bruccoleri, B. Olafson, D. States, S. Swaminathan, M. Karplus, CHARMM: A program for macromolecular energy, minimization, and dynamics calculations, J. Comput. Chem. 4 (1983) 187–217, http://dx.doi.org/10.1002/jcc.540040211.

[6] B. Brooks, C. Brooks III, A. Mackerell Jr., L. Nilsson, R. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. Pastor, C. Post, J. Pu, M. Schaefer, B. Tidor, R. Venable, H. Woodcock, X. Wu, W. Yang, D. York, M. Karplus, CHARMM: The biomolecular simulation program, J. Comput. Chem. 30 (2009) 1545–1614, https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.21287.

[7] W. Jorgensen, J. Tirado-Rives, The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin, J. Am. Chem. Soc. 110 (1988) 1657–1666, PMID: 27557051.

[8] W. Jorgensen, D. Maxwell, J. Tirado-Rives, Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids, J. Am. Chem. Soc. 118 (1996) 11225–11236.

[9] L. Wang, T. Martinez, V. Pande, Building force fields: an automatic, systematic, and reproducible approach, J. Phys. Chem. Lett. 5 (2014) 1885–1891.

[10] Y. Qiu, D. Smith, C. Stern, M. Feng, H. Jang, L. Wang, Driving torsion scans with wavefront propagation, J. Chem. Phys. 152 (2020) 244116, http://dx.doi.org/10.1063/5.0009232.

[11] J. Horton, S. Boothroyd, J. Wagner, J. Mitchell, T. Gokey, D. Dotson, P. Behara, V. Ramaswamy, M. Mackey, J. Chodera, J. Anwar, D. Mobley, D. Cole, Open force field BespokeFit: Automating bespoke torsion parametrization at scale, J. Chem. Inf. Model. 62 (2022) 5622–5633, PMID: 36351167.

[12] D. Mobley, C. Bannan, A. Rizzi, C. Bayly, J. Chodera, V. Lim, N. Lim, K. Beauchamp, M. Shirts, M. Gilson, P. Eastman, Open force field consortium: Escaping atom types using direct chemical perception with SMIRNOFF v0.1, 2018, BioRxiv. https://www.biorxiv.org/content/early/2018/07/13/286542.

[13] Peter Eastman, Jason Swails, John D. Chodera, Robert T. McGibbon, Yutong Zhao, Kyle A. Beauchamp, Lee-Ping Wang, Andrew C. Simmonett, Matthew P. Harrigan, Chaya D. Stern, Rafal P. Wiewiora, Bernard R. Brooks, Vijay S. Pande, OpenMM 7: Rapid development of high performance algorithms for molecular dynamics, PLoS Comput. Biol. 13 (2017) (2017) 1–17, http://dx.doi.org/10.1371/journal.pcbi.1005659.

[14] X. He, B. Walker, V. Man, P. Ren, J. Wang, Recent progress in general force fields of small molecules, Curr. Opin. Struct. Biology 72 (2022) 187–193, https://www.sciencedirect.com/science/article/pii/S0959440X21001603.

[15] C. Bayly, P. Cieplak, W. Cornell, P. Kollman, A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model, J. Phys. Chem. 97 (1993) 10269–10280.

[16] A. Jakalian, D. Jack, C. Bayly, Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation, J. Comput. Chem. 23 (2002) 1623–1641, https://onlinelibrary.wiley.com/doi/abs/10.1002/jcc.10128.

[17] P. Li, L. Song, K. Merz, Parameterization of highly charged metal ions using the 12-6-4 LJ-type nonbonded model in explicit water, J. Phys. Chem. B 119 (2015) 883–895, PMID: 25145273.

[18] R. Strickstrock, M. Hülsmann, D. Reith, K. Kirschner, Optimizing lennard-jones parameters by coupling single molecule and ensemble target data, Comput. Phys. Commun. 274 (2022) 108285, https://www.sciencedirect.com/science/article/pii/S0010465522000030.

[19] Y. Miao, W. Sinko, L. Pierce, D. Bucher, R. Walker, J. McCammon, Improved reweighting of accelerated molecular dynamics simulations for free energy calculation, J. Chem. Theory Comput. 10 (2014) 2677–2689, http://dx.doi.org/10.1021/ct500090q.

[20] J. Köfinger, L. Stelzl, K. Reuter, C. Allande, K. Reichel, G. Hummer, Efficient ensemble refinement by reweighting, J. Chem. Theory Comput. 15 (2019) 3390–3401, http://dx.doi.org/10.1021/acs.jctc.8b01231.

[21] P.G. Bolhuis, Z.F. Brotzakis, B.G. Keller, Force field optimization by imposing kinetic constraints with path reweighting, 2022, http://dx.doi.org/10.48550/arXiv.2207.04558, arXiv:2207.04558.
[22] X. Cui, H. Liu, H. Chen, Polarizable force field of intrinsically disordered proteins with CMAP and reweighting optimization, J. Chem. Inf. Model. 62 (2022) 4970–4982, http://dx.doi.org/10.1021/acs.jcim.2c00835.
[23] J. Chen, H. Liu, X. Cui, Z. Li, H. Chen, RNA-specific force field optimization with CMAP and reweighting, J. Chem. Inf. Model. 62 (2022) 372–385, http://dx.doi.org/10.1021/acs.jcim.1c01148.
[24] X. Ding, Optimizing force fields with experimental data using ensemble reweighting and potential contrasting, J. Phys. Chem. B. 128 (2024) 6760–6769, http://dx.doi.org/10.1021/acs.jpcb.4c02147.
[25] K. Kirschner, W. Heiden, D. Reith, Relative electronic and free energies of octane's unique conformations, Mol. Phys. 115 (2017) 1155–1165.
[26] Gestis-Stoffdatenbank Properties of Octan, 2024, https://gestis.dguv.de/data?name=013810. (Accessed: 19 January 2024).
[27] M. Hülsmann, T. Köddermann, J. Vrabec, D. Reith, GROW: A gradient-based optimization workflow for the automated development of molecular models, Comput. Phys. Commun. 181 (2010) 499–513, http://www.sciencedirect.com/science/article/pii/S0010465509003452.
[28] Strickstrock, R. FFLOW Github Repository, https://github.com/rstrickstrock/tuning_weighting-factors_objective-function.git.
[29] H. Berendsen, D. Spoel, R. Drunen, GROMACS: A message-passing parallel molecular dynamics implementation, Comput. Phys. Comm. 91 (1995) 43–56, http://dx.doi.org/10.1016/0010-4655(95)00042-E.
[30] E. Lindahl, B. Hess, D. Spoel, GROMACS 3.0: a package for molecular simulation and trajectory analysis, Mol. Model. Annu. 7 (2001) 306–317, http://dx.doi.org/10.1007/s008940100045.
[31] K. Nabert, G. Schön, Sicherheitstechnische Kennzahlen brennbarer Gase und Dämpfe (mit 6. Nachtrag), Deutscher Eichverlag, 1963.
[32] NIST-chemistry-webbook isobaric properties for octane, 2024, https://webbook.nist.gov/cgi/fluid.cgi?P=1&TLow=290.15&THigh=300.15&TInc=1&Digits=5&ID=C111659&Action=Load&Type=IsoBar&TUnit=K&PUnit=bar&DUnit=kg%252Fm3&HUnit=kcal%252Fmol&WUnit=m%252Fs&VisUnit=uPa*s&STUnit=N%252Fm&RefState=DEF. (Accessed: 20 February 2024).
[33] O. Kunz, W. Wagner, Wide-range equation of state for natural gases and other mixtures: An expansion of GERG-2004, J. Chem. Eng. Data 57 (2012) 3032–3091.
[34] R. Beckmüller, M. Thol, I. Bell, E. Lemmon, R. Span, New equations of state for binary hydrogen mixtures containing methane, nitrogen, carbon monoxide, and carbon dioxide, J. Phys. Chem. Ref. Data 50 (2021) 013102, http://dx.doi.org/10.1063/5.0040533.
[35] S. Zeppieri, J. Rodríguez, A. Ramos, Interfacial tension of alkane + water systems, J. Chem. Eng. Data 46 (2001) 1086–1088.
[36] NIST-database fluid properties of octane, 2020, https://webbook.nist.gov/cgi/fluid.cgi?P=1&TLow=217.15&THigh=355.15&TInc=1.0&Applet=on&Digits=5&ID=C111659&Action=Load&Type=IsoBar&TUnit=K&PUnit=bar&DUnit=kg%252Fm3&HUnit=kJ%252Fmol&WUnit=m%252Fs&VisUnit=uPa*s&STUnit=N%252Fm&RefState=DEF. (Accessed: 15 May 2020).
[37] NIST-database fluid properties of hexane, 2020, https://webbook.nist.gov/cgi/fluid.cgi?P=1&TLow=217.15&THigh=355.15&TInc=1.0&Applet=on&Digits=5&ID=C110543&Action=Load&Type=IsoBar&TUnit=K&PUnit=bar&DUnit=kg%252Fm3&HUnit=kJ%252Fmol&WUnit=m%252Fs&VisUnit=uPa*s&STUnit=N%252Fm&RefState=DEF. (Accessed: 15 May 2020).
[38] NIST-database fluid properties of heptane, 2020, https://webbook.nist.gov/cgi/fluid.cgi?P=1&TLow=217.15&THigh=355.15&TInc=1.0&Applet=on&Digits=5&ID=C142825&Action=Load&Type=IsoBar&TUnit=K&PUnit=bar&DUnit=kg%252Fm3&HUnit=kJ%252Fmol&WUnit=m%252Fs&VisUnit=uPa*s&STUnit=N%252Fm&RefState=DEF. (Accessed: 15 May 2020).
[39] NIST-database fluid properties of nonane, 2020, https://webbook.nist.gov/cgi/fluid.cgi?P=1&TLow=240.15&THigh=355.15&TInc=1.0&Applet=on&Digits=5&ID=C111842&Action=Load&Type=IsoBar&TUnit=K&PUnit=bar&DUnit=kg%252Fm3&HUnit=kJ%252Fmol&WUnit=m%252Fs&VisUnit=uPa*s&STUnit=N%252Fm&RefState=DEF. (Accessed: 15 May 2020).
[40] Gestis-Stoffdatenbank Properties of Octan, 2020, http://gestis.itrust.de/nxt/gateway.dll/gestis_de/013810.xml?f=templatesfn=document-frame.htm3.0GLOBAL=G_&G_DIEXSL=gestis.xslq=%5Bfield,acstoff%3Aoctan%5D%20uq=x=serverup=1#0604. (Accessed: 15 May 2020).