Supplementary Material to Parametrization of two-center Lennard-Jones plus pointquadrupole force field models by multicriteria optimization

October 23, 2015

# Contents

1	Inve	estigation parameter correlation $\epsilon/Q$	3
<b>2</b>	Self-organizing patch plot		4
	2.1	Self-organizing map	4
	2.2	Training the self-organizing map	5
	2.3	Combination with Voronoi	7
3	3 Optimization data all studied fluids		9
	3.1	Data files $two Criteria Scenario.txt$ and $three Criteria Scenario.txt$ .	9
	3.2	Data in self-organizing patch plots	9
	3 3	Study literature models	19

# 1 Investigation parameter correlation $\epsilon/Q$

In Fig. 1 the Pareto sets of the two- and three-criteria scenarios for all studied fluids are plotted in a  $\epsilon$ -Q-diagram. As already mentioned in the main text, the inverse correlation between  $\epsilon$  and Q found for  $C_2H_2$  cannot be generalized and is only found for some studied fluids.

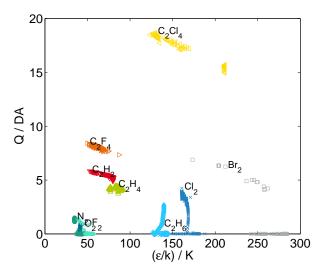


Figure 1: Pareto optimal model parameters  $\epsilon$  and Q for all studied fluids. The plot contains the Pareto solutions of the two- and three-criteria scenarios.

# 2 Self-organizing patch plot

The self-organizing patch plot SOPP is created with a combination of two existing algorithms, i.e. the self-organizing map SOM [1] and the construction of a Voronoi diagram [2]. Basically four steps are necessary to create a SOPP (cf. main text, Fig. 2):

- 1. Train SOM with Pareto solutions
- 2. Map every solution to best matching unit BMU
- 3. Create Voronoi diagram
- 4. Color patches

In this section, the concept of the SOM is shortly introduced. Then details how the SOPPs in Figs. 3 to 20 are created are given.

### 2.1 Self-organizing map

Self-organizing map SOM is a useful tool to visualize high-dimensional data in an e.g. two dimensional map [1]. Fig. 2 shows an example: The function values of four functions, i.e. the mean relative deviation in the liquid density, the vapor pressure, the surface tension and the relative deviation in the critical temperature (lower row) and their dependency of the parameters size  $\sigma$ , energy  $\epsilon$ , elongation L, and quadrupole moment Q (upper row) are shown for  $O_2$ . This SOM consists of  $60 \times 60$  units, also referred to as neurons. Each neuron is assigned a fixed location in the map, i.e. a row number m and column number m, and a feature vector. The feature vector has eight components, one for each of the eight properties shown

$$\mathbf{f}_{\mathrm{mn}} = \left(\sigma_{\mathrm{mn}}^{\star}, \epsilon_{\mathrm{mn}}^{\star}, L_{\mathrm{mn}}^{\star}, Q_{\mathrm{mn}}^{\star}, (\delta \rho')_{\mathrm{mn}}^{\star}, (\delta p^{S})_{\mathrm{mn}}^{\star}, (\delta \gamma)_{\mathrm{mn}}^{\star}, (\delta T_{\mathrm{C}})_{\mathrm{mn}}^{\star}\right) \in [0, 1]^{N_{\mathrm{features}}}$$
(1)

All properties are normalized, such that they take values between 0 and 1. This is marked by the star symbol. Each property p is assigned a diagram and the pixel in the map is colored according to the pth value of the feature vector assigned to the neuron. Thus Fig. 2 shows the function values of the mean relative deviation in the liquid density, the vapor pressure, the surface tension

and the relative deviation in the critical temperature for  $60 \times 60$  parameter combinations.

The advantage of the SOM is, that it can be trained to represent certain regions of an input space. Given a training data set representing the input space, a SOM attempts to map this into an output space, e.g. the two dimensional grid, by trying to preserve the topological relations in the input data. The map does not represent the exact training data, but is a best possible representation of the input space. More general information about SOM, which are also referred to as Kohonen maps can e.g. be found in [1].

## 2.2 Training the self-organizing map

To generate the SOM an implementation of Azzopardi [3] was used. The explanation how to train the SOM given here strongly alludes to this implementation and a rough sketch is given in Algorithm 1. Our SOM consists of  $60 \times 60$  neurons on a two dimensional grid. A neuron is located at row m and column n. Each neuron is assigned a feature vector  $\mathbf{f}_{mn} \in [0,1]^{N_{\text{features}}}$ . The SOM is trained with training vectors  $\mathbf{t}_i \in [0,1]^{N_{\text{features}}}$ .

The training vectors  $\mathbf{t}_i \in [0,1]^{N_{\text{features}}}$  are generated out of the Pareto optimal model parameters size  $\sigma$ , energy  $\epsilon$ , elongation L, and quadrupole moment Q and the corresponding mean relative deviations in the liquid density, the vapor pressure, the surface tension and the relative deviation in the critical temperature. All features are scaled such that the values for the corresponding features are between 0 and 1. E.g. the first component  $\sigma$  of the feature vector is transformed to  $\sigma^*$ :

$$\sigma^{\star} = \frac{\sigma - \sigma_{\min}}{\sigma_{\max} - \sigma_{\min}} \tag{2}$$

Where  $\sigma_{\min}$  and  $\sigma_{\max}$  are the minimal and maximal values, respectively, determined for  $\sigma$  in the set of Pareto optimal solutions. This is done analogously for the remaining properties and our training vector is

$$\mathbf{t}_{i} = (\sigma_{i}^{\star}, \epsilon_{i}^{\star}, L_{i}^{\star}, Q_{i}^{\star}, (\delta \rho')_{i}^{\star}, (\delta p^{S})_{i}^{\star}, (\delta \gamma)_{i}^{\star}, (\delta T_{\mathbf{C}})_{i}^{\star}) \in [0, 1]^{N_{\text{features}}}$$
(3)

with  $i = 1, ..., N_{\text{pareto}}$  and  $N_{\text{pareto}}$  the number of Pareto optimal solutions.

The algorithm starts with initializing each of the  $60 \times 60$  neurons with a random feature vector. Then the first training vector  $\mathbf{t}_i$  is taken and all distances to the feature vectors  $\mathbf{f}_{nm}$  in the SOM are calculated. The best matching unit BMU is the neuron with the smallest euclidean distance to the considered training

**Algorithm 1** Algorithm to create self-organizing map (Implementation of Azzopardi [3]).

```
zopardi [3]).

INPUT: N_{\text{features}} = 8, \text{ndim} = 60, N_{\text{epochs}} = 30, \text{trainingData}, \eta_0 = 0.1, \eta_{\text{d}} = 0.05, \kappa_0 = 30, \kappa_{\text{d}} = 0.05

OUTPUT: SOM

Initialize SOM with randomized vectors.

for all epochs do

Determine parameters for training rate and quality

Determine size of neighborhood

for all training vectors do

Find best matching unit (BMU)

Initialize Gauss function centered around BMU

Update neurons in neighborhood of BMU

end for

end for
```

vector. In the neighborhood of the BMU, the feature vectors  $\mathbf{f}_{mn}$  of the neurons are adjusted to better assimilate the training vector  $\mathbf{t}_i$  as

$$\mathbf{f}_{mn} = \mathbf{f}_{mn} + \eta \cdot g_{\text{BMU}}(m, n)(\mathbf{t}_i - \mathbf{f}_{mn}) \tag{4}$$

where  $\eta$  is the learning rate and  $g_{\rm BMU}(n,m)$  is a Gaussian function defined for the specific location of the BMU

$$g_{\text{BMU}}(m,n) = \exp\left(-\frac{(m - m_{BMU})^2 + (n - n_{BMU})^2}{2\kappa^2}\right)$$
 (5)

where  $m_{BMU}$  and  $n_{BMU}$  denote the row and the column of the BMU, respectively. The adjustment is such that the feature vectors  $\mathbf{f}_{nm}$  of the neurons closest to the BMU are adjusted most to assimilate the training vector  $\mathbf{t}_i$ . After all feature vectors in a neighborhood  $3\kappa$  are adjusted, the algorithm proceeds with the next training vector  $\mathbf{t}_{i+1}$ . Once all training vectors are processed, an epoch is over. The next epoch starts with a reduced learning rate  $\eta$  and a smaller neighborhood  $3\kappa$ . The learning rate  $\eta$  is calculated as

$$\eta = \eta_0 e^{-\eta_{\rm d} t_{\rm epoch}} \tag{6}$$

where  $\eta_0$  and  $\eta_d$  are input parameters for the algorithm. For our application they are set to  $\eta_0 = 0.1$  and  $\eta_d = 0.05$ , as in the example provided by Azzopardi

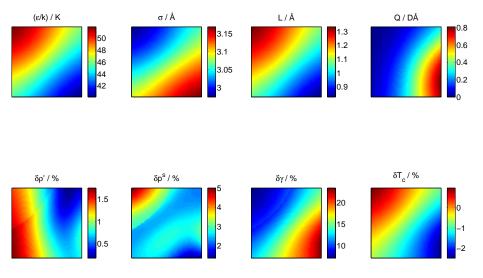


Figure 2:  $60 \times 60$  parameter combinations and corresponding function values for  $O_2$  displayed in a self-organizing map. Upper row: two-center Lennard-Jones plus pointquadrupole parameters energy  $\epsilon$ , size  $\sigma$ , elongation L, and quadrupole moment Q. Lower row: mean relative deviation for the liquid density  $\delta \rho'$ , the vapor pressure  $\delta p^S$ , the surface tension  $\delta \gamma$ , and the relative deviation for the critical temperature  $T_C$ .

[3]. The size of the neighborhood is  $3\kappa$  with

$$\kappa = \kappa_0 e^{-\kappa_{\rm d} t_{\rm epoch}} \tag{7}$$

where  $\kappa$  and  $\kappa_d$  are also input parameters for the algorithm. The values are set to  $\kappa=30$  and  $\kappa_d=0.05$ .

# 2.3 Combination with Voronoi

After training the SOM with the Pareto set, the neurons in the two dimensional grid represent the Pareto set. In Fig. 2 the SOM resulting from a training with the data of the multicriteria optimization for the liquid density and the vapor pressure of  $O_2$  is shown. Each considered feature is assigned a separate diagram and the pixels are colored according to the respective value in the assigned feature vector. In the upper row the values of the parameters and in the lower row the values in the goal functions are displayed. The conflict in the objectives is well represented, but the map does not represent the exact solutions of the multicriteria optimization task. The goal is to get a plot, in which only the

exact solutions gained in the multicriteria optimization are plotted. Thus we combine the self-organizing maps with the construction of a Voronoi diagram.

The input for constructing the Voronoi diagram is a set of points collocated in a plane also referred to as seeds. The Voronoi cells partition the plane into patches, such that each patch contains one seed. One patch consists of all points in the plane, that are closer to this seed, than to any other seed in the plane. With the help of the SOM the Pareto optimal solutions are arranged in a plane, such that similar solutions are close to each other. For each Pareto optimal solution the best matching unit, i.e. the neuron which represents the Pareto solution best regarding the Euclidean distance is identified and marked in the map. Each identified neuron is then assigned the feature vector exactly representing the Pareto optimal solution. So inversely, each Pareto solution is assigned a location in the map. At this point, the information gained from the SOM is no longer relevant. Only the information for the location of the Pareto solutions and their exact values are important for the next steps and are the input to create the Voronoi diagram. The seed  $s_i$  is the location of the BMU in the  $60 \times 60$  grid for the Pareto optimal solution i and a Voronoi diagram with  $N_{\text{pareto}}$  patches is created. To each patch with seed  $\mathbf{s}_i$  the corresponding Pareto optimal solution  $\mathbf{p}_i$  can be mapped. Due to the training of the selforganizing map based on the Pareto optimal solutions, patches close to each other represent similar Pareto optimal solutions. This Voronoi diagram, which was trained to represent Pareto optimal solutions with a self-organizing map, we call self-organizing patch plot SOPP. By generating copies of the SOPP and color coding the patches according to the features of the corresponding Pareto optimal solution, an intuitive representation of Pareto optimal solutions is generated.

# 3 Optimization data all studied fluids

The numerical data for the two- and three-criteria scenarios for all studied fluids is accessible in the data files twoCriteriaScenario.txt and threeCriteriaScenario.txt, respectively. Additionally SOPPs for all optimization scenarios are provided.

# 3.1 Data files two Criteria Scenario.txt and three Criteria Scenario.txt

The data for the two optimization scenarios can be found in the text files two Criteria Optimization.txt and three Criteria Optimization.txt, which are both included in the Supplementary Material. The data is organized as a table with nine columns. The columns contain the following data:

- 1. Substance
- 2. 2CLJQ parameter  $\sigma$ .
- 3. 2CLJQ parameter  $\epsilon$ .
- 4. 2CLJQ parameter L.
- 5. 2CLJQ parameter Q.
- 6. Mean relative deviation liquid density  $\delta \rho'$ .
- 7. Mean relative deviation vapor pressure  $\delta p^S$ .
- 8. Mean relative deviation surface tension  $\delta \gamma$ .
- 9. Relative deviation critical temperature  $T_{\rm C}$ .

The data is sorted in the order the sandwiching and hyperboxing algorithms identified the Pareto points in the Pareto set.

#### 3.2 Data in self-organizing patch plots

In Figs. 3 to 20 the results for all studied fluids are presented in self-organizing patch plots. The overview helps to get a first insight in the data of the Pareto optimal solutions provided with this work.

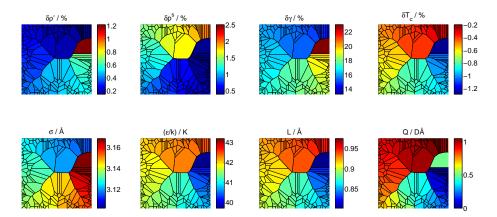


Figure 3: Pareto optimal solutions for the two-criteria scenario for  $O_2$  in a self-organizing patch plot. First row: mean relative deviation for the liquid density  $\delta \rho'$ , the vapor pressure  $\delta p^S$ , the surface tension  $\delta \gamma$ , and the relative deviation for the critical temperature  $\delta T_{\rm C}$ . Second row: two-center Lennard-Jones plus pointquadrupole parameters energy  $\epsilon$ , size  $\sigma$ , elongation L, and quadrupole moment Q.

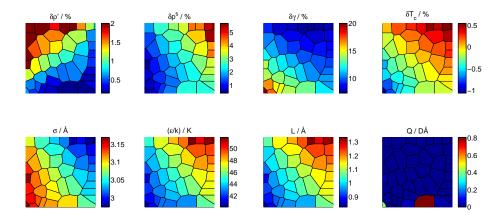


Figure 4: Pareto optimal solutions for the three-criteria scenario for  $O_2$  in a self-organizing patch plot. First row: mean relative deviation for the liquid density  $\delta \rho'$ , the vapor pressure  $\delta p^S$ , the surface tension  $\delta \gamma$ , and the relative deviation for the critical temperature  $\delta T_{\rm C}$ . Second row: two-center Lennard-Jones plus pointquadrupole parameters energy  $\epsilon$ , size  $\sigma$ , elongation L, and quadrupole moment Q.

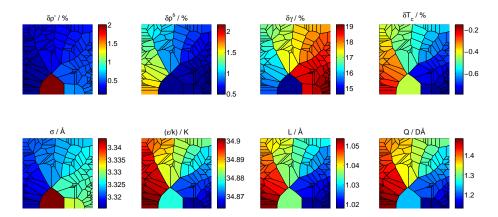


Figure 5: Pareto optimal solutions for the two-criteria scenario for  $N_2$  in a self-organizing patch plot. First row: mean relative deviation for the liquid density  $\delta \rho'$ , the vapor pressure  $\delta p^S$ , the surface tension  $\delta \gamma$ , and the relative deviation for the critical temperature  $\delta T_{\rm C}$ . Second row: two-center Lennard-Jones plus pointquadrupole parameters energy  $\epsilon$ , size  $\sigma$ , elongation L, and quadrupole moment Q.

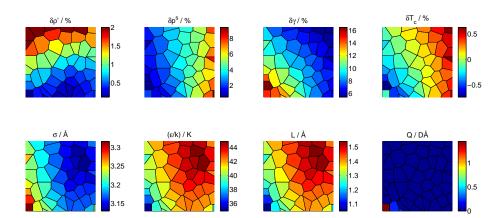


Figure 6: Pareto optimal solutions for the three-criteria scenario for N<sub>2</sub> in a self-organizing patch plot. First row: mean relative deviation for the liquid density  $\delta \rho'$ , the vapor pressure  $\delta p^S$ , the surface tension  $\delta \gamma$ , and the relative deviation for the critical temperature  $\delta T_{\rm C}$ . Second row: two-center Lennard-Jones plus point quadrupole parameters energy  $\epsilon$ , size  $\sigma$ , elongation L, and quadrupole moment Q.

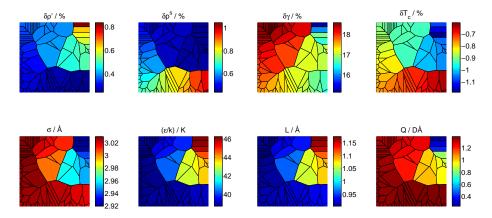


Figure 7: Pareto optimal solutions for the two-criteria scenario for  $F_2$  in a self-organizing patch plot. First row: mean relative deviation for the liquid density  $\delta \rho'$ , the vapor pressure  $\delta p^S$ , the surface tension  $\delta \gamma$ , and the relative deviation for the critical temperature  $\delta T_{\rm C}$ . Second row: two-center Lennard-Jones plus pointquadrupole parameters energy  $\epsilon$ , size  $\sigma$ , elongation L, and quadrupole moment Q.

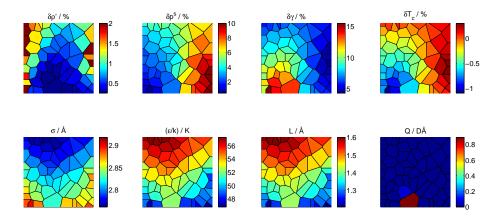


Figure 8: Pareto optimal solutions for the three-criteria scenario for  $F_2$  in a self-organizing patch plot. First row: mean relative deviation for the liquid density  $\delta \rho'$ , the vapor pressure  $\delta p^S$ , the surface tension  $\delta \gamma$ , and the relative deviation for the critical temperature  $\delta T_C$ . Second row: two-center Lennard-Jones plus pointquadrupole parameters energy  $\epsilon$ , size  $\sigma$ , elongation L, and quadrupole moment Q.

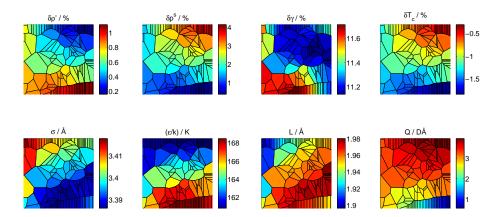


Figure 9: Pareto optimal solutions for the two-criteria scenario for  $\text{Cl}_2$  in a self-organizing patch plot. First row: mean relative deviation for the liquid density  $\delta \rho'$ , the vapor pressure  $\delta p^S$ , the surface tension  $\delta \gamma$ , and the relative deviation for the critical temperature  $\delta T_{\text{C}}$ . Second row: two-center Lennard-Jones plus pointquadrupole parameters energy  $\epsilon$ , size  $\sigma$ , elongation L, and quadrupole moment Q.

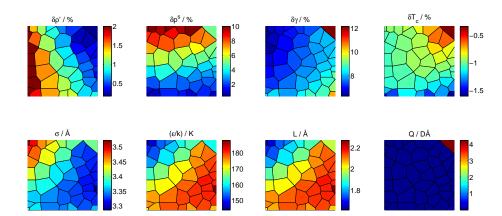


Figure 10: Pareto optimal solutions for the three-criteria scenario for  $\text{Cl}_2$  in a self-organizing patch plot. First row: mean relative deviation for the liquid density  $\delta \rho'$ , the vapor pressure  $\delta p^S$ , the surface tension  $\delta \gamma$ , and the relative deviation for the critical temperature  $\delta T_{\text{C}}$ . Second row: two-center Lennard-Jones plus pointquadrupole parameters energy  $\epsilon$ , size  $\sigma$ , elongation L, and quadrupole moment Q.

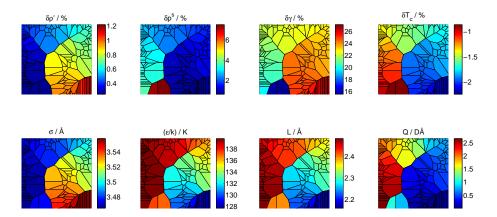


Figure 11: Pareto optimal solutions for the two-criteria scenario for  $C_2H_6$  in a self-organizing patch plot. First row: mean relative deviation for the liquid density  $\delta\rho'$ , the vapor pressure  $\delta p^S$ , the surface tension  $\delta\gamma$ , and the relative deviation for the critical temperature  $\delta T_C$ . Second row: two-center Lennard-Jones plus pointquadrupole parameters energy  $\epsilon$ , size  $\sigma$ , elongation L, and quadrupole moment Q.

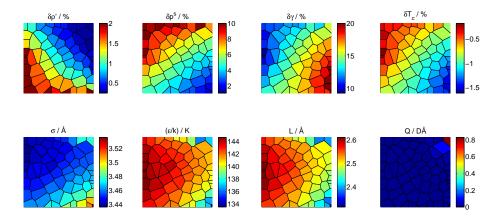


Figure 12: Pareto optimal solutions for the three-criteria scenario for  $C_2H_6$  in a self-organizing patch plot. First row: mean relative deviation for the liquid density  $\delta\rho'$ , the vapor pressure  $\delta p^S$ , the surface tension  $\delta\gamma$ , and the relative deviation for the critical temperature  $\delta T_C$ . Second row: two-center Lennard-Jones plus pointquadrupole parameters energy  $\epsilon$ , size  $\sigma$ , elongation L, and quadrupole moment Q.

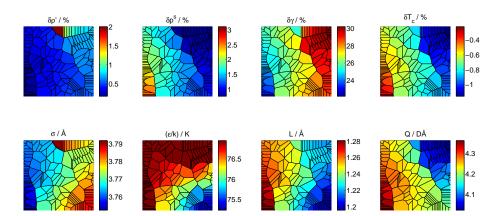


Figure 13: Pareto optimal solutions for the two-criteria scenario for  $C_2H_4$  in a self-organizing patch plot. First row: mean relative deviation for the liquid density  $\delta \rho'$ , the vapor pressure  $\delta p^S$ , the surface tension  $\delta \gamma$ , and the relative deviation for the critical temperature  $\delta T_C$ . Second row: two-center Lennard-Jones plus pointquadrupole parameters energy  $\epsilon$ , size  $\sigma$ , elongation L, and quadrupole moment Q.

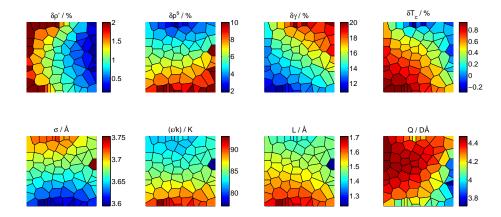


Figure 14: Pareto optimal solutions for the three-criteria scenario for  $C_2H_4$  in a self-organizing patch plot. First row: mean relative deviation for the liquid density  $\delta \rho'$ , the vapor pressure  $\delta p^S$ , the surface tension  $\delta \gamma$ , and the relative deviation for the critical temperature  $\delta T_{\rm C}$ . Second row: two-center Lennard-Jones plus pointquadrupole parameters energy  $\epsilon$ , size  $\sigma$ , elongation L, and quadrupole moment Q.

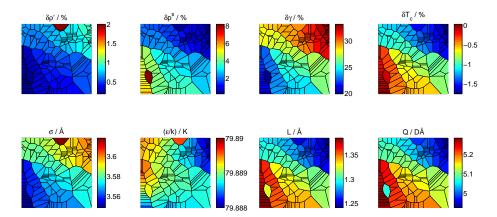


Figure 15: Pareto optimal solutions for the two-criteria scenario for  $C_2H_2$  in a self-organizing patch plot. First row: mean relative deviation for the liquid density  $\delta\rho'$ , the vapor pressure  $\delta p^S$ , the surface tension  $\delta\gamma$ , and the relative deviation for the critical temperature  $\delta T_C$ . Second row: two-center Lennard-Jones plus pointquadrupole parameters energy  $\epsilon$ , size  $\sigma$ , elongation L, and quadrupole moment Q.

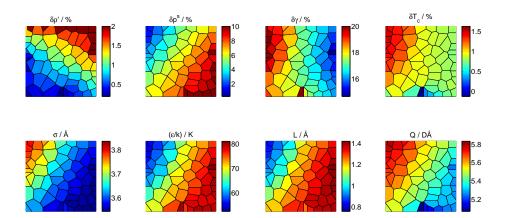


Figure 16: Pareto optimal solutions for the three-criteria scenario for  $C_2H_2$  in a self-organizing patch plot. First row: mean relative deviation for the liquid density  $\delta \rho'$ , the vapor pressure  $\delta p^S$ , the surface tension  $\delta \gamma$ , and the relative deviation for the critical temperature  $\delta T_C$ . Second row: two-center Lennard-Jones plus pointquadrupole parameters energy  $\epsilon$ , size  $\sigma$ , elongation L, and quadrupole moment Q.

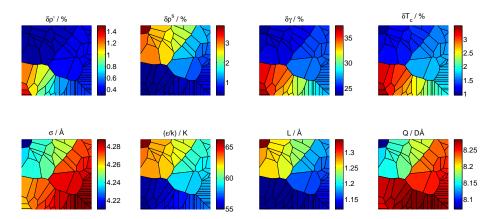


Figure 17: Pareto optimal solutions for the two-criteria scenario for  $C_2F_4$  in a self-organizing patch plot. First row: mean relative deviation for the liquid density  $\delta \rho'$ , the vapor pressure  $\delta p^S$ , the surface tension  $\delta \gamma$ , and the relative deviation for the critical temperature  $\delta T_C$ . Second row: two-center Lennard-Jones plus pointquadrupole parameters energy  $\epsilon$ , size  $\sigma$ , elongation L, and quadrupole moment Q.

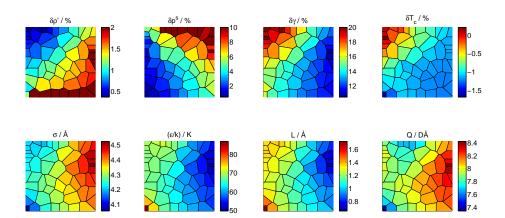


Figure 18: Pareto optimal solutions for the three-criteria scenario for  $C_2F_4$  in a self-organizing patch plot. First row: mean relative deviation for the liquid density  $\delta \rho'$ , the vapor pressure  $\delta p^S$ , the surface tension  $\delta \gamma$ , and the relative deviation for the critical temperature  $\delta T_C$ . Second row: two-center Lennard-Jones plus pointquadrupole parameters energy  $\epsilon$ , size  $\sigma$ , elongation L, and quadrupole moment Q.

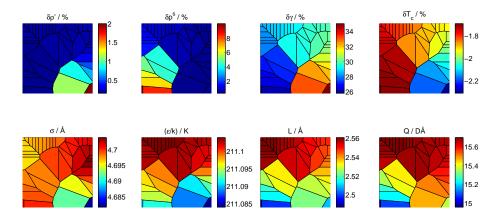


Figure 19: Pareto optimal solutions for the two-criteria scenario for  $C_2Cl_4$  in a self-organizing patch plot. First row: mean relative deviation for the liquid density  $\delta \rho'$ , the vapor pressure  $\delta p^S$ , the surface tension  $\delta \gamma$ , and the relative deviation for the critical temperature  $\delta T_{\rm C}$ . Second row: two-center Lennard-Jones plus pointquadrupole parameters energy  $\epsilon$ , size  $\sigma$ , elongation L, and quadrupole moment Q.

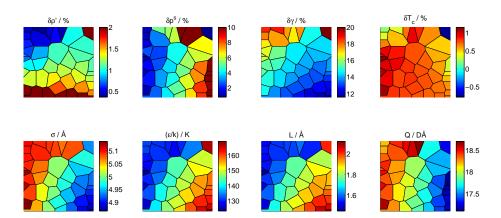


Figure 20: Pareto optimal solutions for the three-criteria scenario for  $C_2Cl_4$  in a self-organizing patch plot. First row: mean relative deviation for the liquid density  $\delta\rho'$ , the vapor pressure  $\delta p^S$ , the surface tension  $\delta\gamma$ , and the relative deviation for the critical temperature  $\delta T_C$ . Second row: two-center Lennard-Jones plus pointquadrupole parameters energy  $\epsilon$ , size  $\sigma$ , elongation L, and quadrupole moment Q.

#### 3.3 Study literature models

A comparison of various models from the literature can be found in the main text in Table 1. The deviations indicated for the liquid density, the vapor pressure, and the surface tension are obtained by evaluating the mean relative deviation in Eq. 5 and the relative deviation for the critical temperature in Eq. 8 for the specified model parameters under the same conditions as in the multicriteria optimization task from this work (cf. Section *Optimization task*). The literature study is briefly described here.

Singer et al. [4] calculated thermodynamic and structural features for liquid 2CLJ model systems for  $L^*$  between 0.5-0.8 and fitted parameters to experimental data of  $F_2$ ,  $Cl_2$ , and  $Br_2$ . Even though no quadrupole moment was considered, the model parameters for  $F_2$  and  $Cl_2$  yield results close to the Pareto set. The  $Br_2$  model parameters do not yield mean relative deviations in close proximity to the Pareto set. For  $F_2$ ,  $Cl_2$ , and  $Br_2$  the quadrupole moment of the Pareto optimal model parameters is in the range of Q=(0.07-0.9) DÅ, Q=(2.4-4.3) DÅ, and Q=(4.8-5.3) DÅ, respectively. As already concluded by Singer et al. [4] the simpler 2CLJ potential cannot account for the effect of a large quadrupole moment.

McGuigan et al. [5] published 2CLJ model parameters for  $N_2$ ,  $F_2$ , and  $Cl_2$ . The parameters were fitted to experimental data for the liquid density and the vapor pressure. Especially for the  $Cl_2$  model the results are close to the Pareto set with mean relative deviations in the liquid density of approximately 1 % and in the vapor pressure of 2.37 %.

Rivera et al. [6] parametrized  $N_2$  as a 2CLJ model by fitting simulated liquid and vapor coexisting densities to experimental data. In a subsequent step they also determined the surface tension. Their results for the mean relative deviation of the liquid density and vapor pressure are slightly outside the considered range of the diagrams in Fig. 3.

Cheung and Powles [7] compared model parameters for  $N_2$  as a 2CLJ and 2CLJQ potential for a variety of properties. For the two parameter sets, the quality in the liquid density is of the same order, but the mean relative deviation in the vapor pressure for the 2CLJQ model is better by 10 %.

Vrabec et al. [8] published models to all fluids studied in this paper. They carried out a single criteria optimization based on the correlations from Stoll et al. [9] considering the deviations between experimental and simulation data for

the liquid density, the vapor pressure and the critical temperature. The mean relative deviations achieved with their model parameters for  $O_2$ ,  $N_2$ ,  $Cl_2$ ,  $Br_2$ ,  $C_2H_6$ ,  $C_2H_4$ , and  $C_2H_2$  are very close to the Pareto set found in this work. For  $F_2$  it is possible to improve the quality of the vapor pressure by more than 1.5 % without having to decrease the quality of the liquid density. For  $C_2F_4$ , and  $C_2Cl_4$  it is possible to simultaneously improve the liquid density by about 0.5 % and the vapor pressure by more than 3 %.

Martin and Siepmann [10] published transferable potentials for phase equilibria (TraPPE) for n-alkanes and compared the performance with the models from Jorgensen et al. (OPLS) [11] and Siepmann et al. (SKS) [12], which are both based on methyl and methylene groups as building blocks for hydrocarbons. The adjustment for these models focused on reproducing energies and densities in the liquid phase (OPLS) and VLE data (SKS, TraPPE). The performances of the models to represent experimental data of  $C_2H_6$  indicate, that the task of getting generalized parameter sets for alkanes is demanding. Whereas the TraPPE [10] model yields an excellent quality for the liquid density, the mean relative deviation in the vapor pressure is about 23 %. For the SKS [12] and the OPLS [11] it is even 28 % and 70 %, respectively.

Nath et al. [13] also studied the VLE of alkanes and proposed slightly different Lennard-Jones parameters (NERD). They also conducted a comparison to the SKS and TraPPE force fields.

Wojcik et al. [14] developed correlations for the configuration energy and the liquid density for the 2CLJ potential model and demonstrated how to fit those to experimental data for liquid  $C_2H_6$ . The published model parameters are the base for Fincham et al. [15]. They altered the value in  $\sigma$  to improve the model performance for the vapor pressure in the VLE.

## References

- [1] T. Kohonen, *Self-organizing Maps*. Heidelberg: Springer Series in Information Sciences, Springer, 2001.
- [2] F. Aurenhammer, R. Klein, and D. Lee, *Voronoi diagrams and Delaunay triangulations*. Singapore: World Scientific, 2013.
- [3] G. Azzopardi, "Self-organizing map simple demonstration," http://www.mathworks.com/matlabcentral/fileexchange, 2013, file ID #39930.
- [4] K. Singer, A. Taylor, and J. Singer, Mol. Phys., vol. 33, no. 6, pp. 1757–1795, 1977.
- [5] D. B. McGuigan, M. Lupkowski, D. M. Paquet, and P. A. Monson, *Mol. Phys.*, vol. 67, no. 1, pp. 33–52, 1989.
- [6] J. L. Rivera, J. Alejandre, S. K. Nath, and J. J. de Pablo, Mol. Phys., vol. 98, no. 1, pp. 43–55, 2000.
- [7] P. S. Y. Cheung and J. G. Powles, Mol. Phys., vol. 32, no. 5, pp. 1383–1405, 1976.
- [8] J. Vrabec, J. Stoll, and H. Hasse, *J. Phys. Chem. B*, vol. 105, no. 48, pp. 12126–12133, 2001.
- [9] J. Stoll, J. Vrabec, H. Hasse, and J. Fischer, Fluid Phase Equilib., vol. 179, no. 1-2, pp. 339–362, 2001.
- [10] M. G. Martin and J. I. Siepmann, J. Phys. Chem. B, vol. 102, pp. 2569–2577, 1998.
- [11] W. L. Jorgensen, J. D. Madura, and C. J. Swenson, J. Am. Chem. Soc., vol. 106, no. 22, pp. 6638–6646, 1984.
- [12] J. I. Siepmann, S. Karaborni, and B. Smit, *Nature*, vol. 365, pp. 330–332, 1993.
- [13] S. K. Nath, F. A. Escobedo, and J. J. de Pablo, J. Chem. Phys., vol. 108, no. 23, pp. 9905–9911, 1998.
- [14] M. Wojcik, K. Gubbins, and J. Powles, Mol. Phys., vol. 45, no. 6, pp. 1209–1225, 1982.
- [15] D. Fincham, N. Quirke, and D. J. Tildesley, J. Chem. Phys., vol. 84, no. 8, pp. 4535–4546, 1986.