Supporting Information for Predictive Thermodynamic Modeling of Poorly Specified Mixtures and Applications in Conceptual Fluid Separation Process Design

Thomas Specht, Hans Hasse, and Fabian Jirasek*

Laboratory of Engineering Thermodynamics (LTD), RPTU Kaiserslautern, Erwin-Schrödinger-Straße 44, 67663 Kaiserslautern, Germany

E-mail: *fabian.jirasek@rptu.de

Calculation of Liquid-liquid Equilibria

The partitioning of all (pseudo-)components in thermodynamic equilibrium was calculated by solving the isoactivity criterion together with mass balances and summation conditions formulated in the Rachford-Rice equation:[1,2]

$$F(\theta) = \sum_{i=1}^{N}(x_i'' - x_i') = \sum_{i=1}^{N} \frac{x_i^{\mathrm{mix}} K_i}{1+\theta(K_i-1)} - \frac{x_i^{\mathrm{mix}}}{1+\theta(K_i-1)} \stackrel{!}{=} 0 \tag{S.1a}$$

$$K_i = \frac{\gamma_i'}{\gamma_i''}; \quad i = 1, ..., N \tag{S.1b}$$

where x_i' and x_i'' are the mole fraction of (pseudo-)component i in the raffinate (water-rich) and the extract (extracting agent-rich) phase, respectively. γ_i' and γ_i'' are the activity coefficients in the raffinate and extract phase, respectively, and $K_i = x_i''/x_i'$ is the partition

coefficient of (pseudo-)component i. x_i^{mix} is the lumped mole fraction of i after mixing the feed with the extracting agent E (irrespective of the phase separation), and $\theta = \frac{n''}{n'+n''}$ is the molar extract phase fraction.

Eqs. (S.1a) and (S.1b) were solved in a double-loop approach[2–4] for a constant temperature of $T = 298.15$ K. Therefore, first, the values of K_i for all (pseudo-)components were calculated using UNIFAC[5,6] for modeling the activity coefficients in both coexisting phases. Then, Eq. (S.1a) was solved with respect to θ using the MATLAB[7] 2021b function 'fsolve', from which updated composition of the extract and raffinate phase and, consequently, updated values for K_i were obtained. This process was repeated until convergence of θ was observed, specifically, until θ changed by less than 10^{-12} between two successive iterations.

Prediction of Residue Curves

Residue curves were modeled by the Rayleigh-equation:[8,9]

$$\frac{dx_i}{dn^{\mathrm{L}}} = \frac{y_i - x_i}{n^{\mathrm{L}}}; \quad i = 1...N-1 \tag{S.2a}$$

$$x_{\mathrm{W}} = 1 - \sum_{i=1}^{N-1} x_i \tag{S.2b}$$

where n^{L} is the total mole number in the liquid phase (which decreases with evaporation over time), x_i and y_i are the mole fractions of all components i except for the solvent water (W) in the liquid and the vapor phase, respectively, in thermodynamic equilibrium. The mole fraction of water in the liquid phase x_{W} was calculated by the summation condition, cf. Eq. (S.2b).

The system of ordinary differential equations defined by Eq. (S.2a) was solved in MATLAB[7] using the composition of the feed as starting composition, i.e., $\boldsymbol{x}(n^{\mathrm{L}} = n^{\mathrm{L},0}) = \boldsymbol{x}_{\mathrm{Feed}}$. The calculation of the residue curves was stopped if the number of moles n^{L} approaches the total amount of non-volatile components (to avoid numerical issues, a small tolerance margin

was added).

The vapor-liquid equilibria were modeled by Raoult's law, cf. Eq. (2) in the manuscript, whereby the pressure p and all mole fractions x_i in the liquid phase for the respective time step were specified, and the activity coefficients γ_i in the liquid phase were calculated with UNIFAC.[5,6] Note that for numerical reasons, first, the boiling temperature of the mixture T was calculated by adjusting T until the sum of the partial pressures p_i in Eq. (2) over all N components was equal to the total (specified) pressure p. Subsequently, the mole fractions y_i in the gas phase were calculated.

For the fully specified mixtures, which were considered for comparison only, the vapor pressure of each pure component was calculated with the Antoine equation:

$$\log_{10}\left(750.062\frac{p_i^{\mathrm{S}}}{\mathrm{bar}}\right) = A_i - \frac{B_i}{\left(\frac{T}{\mathrm{K}} - 273.15\right) + C_i} \tag{S.3}$$

whereby the component-specific constants (A_i, B_i, C_i) were taken from the Dortmund Data Bank (DDB).[10] For most of the studied components, the DDB reports parameter sets for two different temperature ranges. If only one parameter set was labeled as valid for the temperature of interest here, this one was used; if both parameter sets were labeled as suitable, both sets were used for calculating the vapor pressure p_i^{S}, and the mean value of the two p_i^{S} was subsequently used. In some cases, the vapor pressure calculation was also performed outside the given ranges.

The vapor pressures of the pseudo-components k in the poorly specified mixtures were estimated using the group-contribution methods from Refs.:[11,12]

$$\log_{10}\left(\frac{p_k^{\mathrm{S}}}{1.01325\ \mathrm{bar}}\right) = (4.1012 + dB_k)\left(\frac{\frac{T/\mathrm{K}}{T_{\mathrm{b},k}/\mathrm{K}} - 1}{\frac{T/\mathrm{K}}{T_{\mathrm{b},k}/\mathrm{K}} - 0.125}\right) \tag{S.4}$$

where $T_{\mathrm{b},k}$ is the normal boiling point and dB_k is the so-called "slope term"[12] of pseudo-component k. For calculating $T_{\mathrm{b},k}$, the approach from Ref.[11] was used; for calculating dB_k,

the approach from Ref.[12] was used. At the heart of both group-contribution approaches, the absolute number of groups in each (pseudo-)component is required, which was directly obtained by the NMR fingerprinting and pseudo-component method described in Ref.[13] and is given in Tables S.4-S.6. The considered structural groups, as well as the assignment procedure, are described in Table 1 in the manuscript.

In the group-contribution methods of Refs.,[11,12] a so-called group-interaction contribution has to be considered for strongly interacting groups,[11,12] which holds for the hydroxyl ('OH'), carboxylic acid ('COOH'), ketone ('CO') and aldehyde ('CHO') groups here. The calculation of the group-interaction contribution requires that the absolute number of each interactive structural group is an integer in each pseudo-component. Therefore, the absolute number of the interactive groups in each pseudo-component was rounded to the nearest integer only for the calculation of the group-interaction contribution.

Prediction of Feed Composition

In the following, the calculation of the feed composition on the basis of the NMR fingerprinting and pseudo-component method is described. The feed composition is the basis for the thermodynamic modeling of the mixtures. For predicting the composition of a poorly specified feed with the proposed methodology, the following information was assumed to be available and used:

- the total mass of the poorly specified feed m;
- the nature of the solvent (here: always water W);
- the nature and mass fraction $x_{\mathrm{T}}^{(\mathrm{m})}$ of the target component T, in the feed.

From the assumptions, the amount of the target component was calculated:

$$n_{\mathrm{T}} = \frac{x_{\mathrm{T}}^{(\mathrm{m})} m}{M_{\mathrm{T}}} \tag{S.5}$$

Additionally, a mean area $\bar{A}_{\mathrm{T}}$ for the target component was calculated:

$$\bar{A}_{\mathrm{T}} = \frac{A_{\mathrm{T}}^{\mathrm{total}}}{z_{\mathrm{T}}^{\mathrm{total}}} \tag{S.6}$$

where $z_{\mathrm{T}}^{\mathrm{total}}$ is the total number of NMR-active nuclei in the target component and $A_{\mathrm{T}}^{\mathrm{total}}$ is the total area of the target component in the corresponding spectrum. Furthermore, we defined a mean area of a structural group g, taken from the UNIFAC table,[5,6] cf. Table 1 in the manuscript, in pseudo-component k:

$$\bar{A}_{g,k} = \frac{A_{g,k}}{z_g} \tag{S.7}$$

where z_g is the number of NMR-active nuclei in group g that are expected in the same chemical shift region and $A_{g,k}$ was the total area of group g assigned to pseudo-component k. Note that for the CH3CO/CH2CO group in UNIFAC, one typically expects two peaks in different chemical shift regions. The CH3CO/CH2CO group was therefore only identified if in both regions of the ^{13}C NMR spectrum "0-90 ppm" and ">180 ppm" peaks were observed *and* assigned to the same pseudo-component; as a consequence, 'CH3/CH2' groups were only assigned to the peaks in the region "0-90 ppm" that exceeded the peaks in the region ">180 ppm" for each pseudo-component. From this, the mole number of each structural group g in pseudo-component k was calculated:

$$n_{g,k} = \frac{\bar{A}_{g,k}}{\bar{A}_{\mathrm{T}}} n_{\mathrm{T}} \tag{S.8}$$

This yielded the total mass of all groups in pseudo-component k:

$$m_k = \sum_{g=1}^{G} n_{g,k} M_g \tag{S.9}$$

where G is the total number of considered structural groups and M_g the known molar mass of group g. With the known total mass of the pseudo-components, the mass of the target component, and the total mass of the mixture, the mass of the solvent W can be calculated:

$$m_{\mathrm{W}} = m - m_{\mathrm{T}} - \sum_{k=1}^{K} m_k \tag{S.10}$$

Using the information on the molar masses of the respective (pseudo-)components, the mole fraction was calculated as follows:

$$x_i = \frac{n_i}{\sum_{i=1}^{N} n_i} \tag{S.11}$$

Predicted Composition of Pseudo-components

In Tables S.1-S.6, information on the stoichiometry of the pseudo-component as predicted with the approach of the present work for all studied mixtures is summarized. For this purpose, the absolute numbers of the structural groups g in each molecule of the defined pseudo-components k, denoted by $\nu_{g,k}$, in the three test mixtures are compiled. Tables S.1-S.3 thereby consider groups from the UNIFAC[5,6] table, whereas Tables S.4-S.6 consider groups from the table of Nannonal et al.[11,12]

Table S.1: Absolute numbers $\nu_{g,k}$ of structural groups g in pseudo-components k for test mixture I, cf. Table 1 and 2 in the manuscript, according to UNIFAC.[5,6,14] The numbers in parentheses are the identifiers for the UNIFAC sub-groups.

UNIFAC label	Stoichiometry	$\tilde{U}_1$	$\tilde{U}_2$	$\tilde{U}_3$
CH3 (1)	$\nu_{CH3,k}$	0.782	-	1.147
CH2 (2)	$\nu_{CH2,k}$	-	6.464	-
CH (3)	$\nu_{CH,k}$	-	-	-
C (4)	$\nu_{C,k}$	-	-	-
OH (14)	$\nu_{OH,k}$	-	3.206	-
CH=CH (6)	$\nu_{CH=CH,k}$	-	-	-
C=C (70)	$\nu_{C=C,k}$	-	-	-
COOH (42)	$\nu_{COOH,k}$	-	-	1.219
CHO (20)	$\nu_{CHO,k}$	-	-	-
CH3CO (18)	$\nu_{CH3CO,k}$	0.865	-	-
CH2CO (19)	$\nu_{CH2CO,k}$	-	-	-

Table S.2: Absolute numbers $\nu_{g,k}$ of structural groups g in pseudo-components k for test mixture II, cf. Table 1 and 2 in the manuscript, according to UNIFAC.[5,6,14] The numbers in parentheses are the identifiers for the UNIFAC sub-groups.

UNIFAC label	Stoichiometry	$\tilde{U}_1$	$\tilde{U}_2$
CH3 (1)	$\nu_{CH3,k}$	-	-
CH2 (2)	$\nu_{CH2,k}$	3.879	1.898
CH (3)	$\nu_{CH,k}$	-	1.680
C (4)	$\nu_{C,k}$	-	0.759
OH (14)	$\nu_{OH,k}$	-	2.846
CH=CH (6)	$\nu_{CH=CH,k}$	-	0.203
C=C (70)	$\nu_{C=C,k}$	-	-
COOH (42)	$\nu_{COOH,k}$	-	2.238
CHO (20)	$\nu_{CHO,k}$	-	-
CH3CO (18)	$\nu_{CH3CO,k}$	-	-
CH2CO (19)	$\nu_{CH2CO,k}$	0.953	-

Table S.3: Absolute numbers $\nu_{g,k}$ of structural groups g in pseudo-components k for test mixture III, cf. Table 1 and 2 in the manuscript, according to UNIFAC.[5,6,14] The numbers in parentheses are the identifiers for the UNIFAC sub-groups.

UNIFAC label	Stoichiometry	$\tilde{U}_1$	$\tilde{U}_2$	$\tilde{U}_3$	$\tilde{U}_4$	$\tilde{U}_5$	$\tilde{U}_6$	$\tilde{U}_7$
CH3 (1)	$\nu_{\mathrm{CH3},k}$	0.853	1.111	2.111	-	-	-	-
CH2 (2)	$\nu_{\mathrm{CH2},k}$	-	-	1.409	4.688	6.625	1.812	1.617
CH (3)	$\nu_{\mathrm{CH},k}$	-	-	0.712	-	-	3.681	3.135
C (4)	$\nu_{\mathrm{C},k}$	-	-	-	-	-	-	-
OH (14)	$\nu_{\mathrm{OH},k}$	-	-	1.440	-	3.291	4.577	4.752
CH=CH (6)	$\nu_{\mathrm{CH=CH},k}$	-	-	-	-	-	0.435	-
C=C (70)	$\nu_{\mathrm{C=C},k}$	-	-	-	-	-	-	0.858
COOH (42)	$\nu_{\mathrm{COOH},k}$	-	1.208	-	-	-	1.907	3.308
CHO (20)	$\nu_{\mathrm{CHO},k}$	-	-	-	-	-	-	-
CH3CO (18)	$\nu_{\mathrm{CH3CO},k}$	0.903	-	-	-	-	-	-
CH2CO (19)	$\nu_{\mathrm{CH2CO},k}$	-	-	-	1.194	-	-	-

Table S.4: Absolute numbers $\nu_{g,k}$ of structural groups g in pseudo-components k for test mixture I, cf. Table 1 and 2 in the manuscript, according to Refs.[11,12]

Nannonal label	Stoichiometry	$\tilde{U}_1$	$\tilde{U}_2$	$\tilde{U}_3$
CH3 (1)	$\nu_{\mathrm{CH3},k}$	1.647	-	1.147
CH2 (4)	$\nu_{\mathrm{CH2},k}$	-	3.258	-
CH (5)	$\nu_{\mathrm{CH},k}$	-	-	-
C (6)	$\nu_{\mathrm{C},k}$	-	-	-
CH2 (7)	$\nu_{\mathrm{CH2},k}$	-	3.206	-
CH (7)	$\nu_{\mathrm{CH},k}$	-	-	-
C (7)	$\nu_{\mathrm{C},k}$	-	-	-
OH(P) (35)	$\nu_{\mathrm{OH(P)},k}$	-	3.206	-
OH(P) (36)	$\nu_{\mathrm{OH(P)},k}$	-	-	-
OH(S) (34)	$\nu_{\mathrm{OH(S)},k}$	-	-	-
OH(T) (33)	$\nu_{\mathrm{OH(T)},k}$	-	-	-
CH=CH (58)	$\nu_{\mathrm{CH=CH},k}$	-	-	-
C=C (58)	$\nu_{\mathrm{C=C},k}$	-	-	-
COOH (44)	$\nu_{\mathrm{COOH},k}$	-	-	1.219
CHO (52)	$\nu_{\mathrm{CHO},k}$	-	-	-
CO (51)	$\nu_{\mathrm{CO},k}$	0.865	-	-

Table S.5: Absolute numbers $\nu_{g,k}$ of structural groups g in pseudo-components k for test mixture II, cf. Table 1 and 2 in the manuscript, according to Refs.[11,12]

Nannonal label	Stoichiometry	$\tilde{U}_1$	$\tilde{U}_2$
CH3 (1)	$\nu_{\mathrm{CH3},k}$	-	-
CH2 (4)	$\nu_{\mathrm{CH2},k}$	4.832	1.491
CH (5)	$\nu_{\mathrm{CH},k}$	-	-
C (6)	$\nu_{\mathrm{C},k}$	-	-
CH2 (7)	$\nu_{\mathrm{CH2},k}$	-	0.407
CH (7)	$\nu_{\mathrm{CH},k}$	-	1.680
C (7)	$\nu_{\mathrm{C},k}$	-	0.759
OH(P) (35)	$\nu_{\mathrm{OH(P)},k}$	-	0.407
OH(P) (36)	$\nu_{\mathrm{OH(P)},k}$	-	-
OH(S) (34)	$\nu_{\mathrm{OH(S)},k}$	-	1.680
OH(T) (33)	$\nu_{\mathrm{OH(T)},k}$	-	0.759
CH=CH (58)	$\nu_{\mathrm{CH=CH},k}$	-	-
C=C (58)	$\nu_{\mathrm{C=C},k}$	-	-
COOH (44)	$\nu_{\mathrm{COOH},k}$	-	2.238
CHO (52)	$\nu_{\mathrm{CHO},k}$	-	-
CO (51)	$\nu_{\mathrm{CO},k}$	0.953	-

Table S.6: Absolute numbers $\nu_{g,k}$ of structural groups g in pseudo-components k for test mixture III, cf. Table 1 and 2 in the manuscript, according to Refs.[11,12]

Nannonal label	Stoichiometry	$\tilde{U}_1$	$\tilde{U}_2$	$\tilde{U}_3$	$\tilde{U}_4$	$\tilde{U}_5$	$\tilde{U}_6$	$\tilde{U}_7$
CH3 (1)	$\nu_{\mathrm{CH3},k}$	1.757	1.111	2.111	-	-	-	-
CH2 (4)	$\nu_{\mathrm{CH2},k}$	-	-	0.682	5.883	3.334	0.916	-
CH (5)	$\nu_{\mathrm{CH},k}$	-	-	-	-	-	-	-
C (6)	$\nu_{\mathrm{C},k}$	-	-	-	-	-	-	-
CH2 (7)	$\nu_{\mathrm{CH2},k}$	-	-	0.728		3.291	0.895	1.617
CH (7)	$\nu_{\mathrm{CH},k}$	-	-	0.712	-	-	3.681	3.135
C (7)	$\nu_{\mathrm{C},k}$	-	-	-	-	-	-	-
OH(P) (35)	$\nu_{\mathrm{OH(P)},k}$	-	-	-	-	3.291	0.895	1.617
OH(P) (36)	$\nu_{\mathrm{OH(P)},k}$	-	-	0.728	-	-	-	-
OH(S) (34)	$\nu_{\mathrm{OH(S)},k}$	-	-	0.712	-	-	3.681	3.135
OH(T) (33)	$\nu_{\mathrm{OH(T)},k}$	-	-	-	-	-	-	-
CH=CH (58)	$\nu_{\mathrm{CH=CH},k}$	-	-	-	-	-	0.435	-
C=C (58)	$\nu_{\mathrm{C=C},k}$	-	-	-	-	-	-	0.858
COOH (44)	$\nu_{\mathrm{COOH},k}$	-	1.208	-	-	-	1.907	3.308
CHO (52)	$\nu_{\mathrm{CHO},k}$	-	-	-	-	-	-	-
CO (51)	$\nu_{\mathrm{CO},k}$	0.903	-	-	1.194	-	-	-

Table S.7 shows the composition of all components in the test mixtures regarding the groups of UNIFAC.[5,6,14] Table S.8 shows the composition of all extracting agents regarding the groups of UNIFAC.[5,6,14] Note that the UNIFAC nomenclature uses 'THF,'[14] as an abbreviation for cyclic ether groups, but since we found this misleading, we used 'cy-CH2O' instead.

Table S.7: Components considered in this work in the test mixtures, cf. Table 2 in the manuscript, and their composition regarding groups from the UNIFAC table.[5,6,14] The numbers in parentheses are the identifiers for the UNIFAC sub-groups.

Component	UNIFAC groups
acetone	1 x 'CH3' (1) 1 x 'CH3CO' (18)
acetic acid	1 x 'CH3' (1) 1 x 'COOH' (42)
acetonitrile	1 x 'CH3CN' (40)
ascorbic acid	1 x 'CH2' (2) 2 x 'CH' (3) 4 x 'OH' (14) 1 x 'C=C' (70) 1 x 'COO' (77)
1,4-butanediol	4 x 'CH2' (2) 2 x 'OH' (14)
citric acid	2 x 'CH2' (2) 1 x 'C' (4) 1 x 'OH' (14) 3 x 'COOH' (42)
cyclohexanone	4 x 'CH2' (2) 1 x 'CH2CO' (19)
1,4-dioxane	2 x 'CH2' (2) 2 x 'cy-CH2O' (27)
glucose	1 x 'CH2' (2) 4 x 'CH' (3) 5 x 'OH' (14) 1 x 'CHO' (26)
malic acid	1 x 'CH2' (2) 1 x 'CH' (3) 1 x 'OH' (14) 2 x 'COOH' (42)
1-propanol	1 x 'CH3' (1) 2 x 'CH2' (2) 1 x 'OH' (14)
2-propanol	2 x 'CH3' (1) 1 x 'CH' (3) 1 x 'OH' (14)
water	1 x 'H2O' (16)
xylose	4 x 'CH' (3) 4 x 'OH' (14) 1 x 'cy-CH2O' (27)

Table S.8: Extracting agents considered in this work and their composition regarding groups from the UNIFAC table.[5,6,14] The numbers in parentheses are the identifiers for the UNIFAC sub-groups.

Component	UNIFAC groups
1-decanol	1 x 'CH3' (1) 9 x 'CH2' (2) 1 x 'OH' (14)
decane	2 x 'CH3' (1) 8 x 'CH2' (2)
dipropyl ether	2 x 'CH3' (1) 3 x 'CH2' (2) 1 x 'CH2O' (25)
ethyl propionate	2 x 'CH3' (1) 1 x 'CH2' (2) 1 x 'CH2COO' (22)
hexane	2 x 'CH3' (1) 4 x 'CH2' (2)
toluene	5 x 'ACH' (9) 1 x 'ACCH3' (11)
1-octanol	1 x 'CH3' (1) 7 x 'CH2' (2) 1 x 'OH' (14)
3-octanone	2 x 'CH3' (1) 4 x 'CH2' (2) 1 x 'CH2CO' (19)

Results for Pure-component Vapor Pressures

Figures S.1 to S.3 shows the results for the vapor pressures of the (pseudo-)components over the boiling temperature T. For the true components (solid/dotted lines), the results were obtained with the Antoine equation, cf. Eq. (S.3), for the respective pseudo-components (dashed lines), the group-contribution method of Nannonal et al.,[12] cf. Eq. (S.4), was used.

For five of the components used in this study, namely glucose, xylose, citric acid, ascorbic acid, and malic acid, no Antoine-parameters were available in the DDB. However, we can assume that the vapor pressure of these components in the considered temperature range is

negligible, in particular in comparison to the other considered components; therefore, these components were assumed to remain completely in the liquid phase throughout ($y_i = 0$), and Eq. (2) in the manuscript was not considered here. For (pseudo-)components, for which the vapor pressure was assumed negligible ($p_i^{\mathrm{S}} = 0$), no results are shown here.

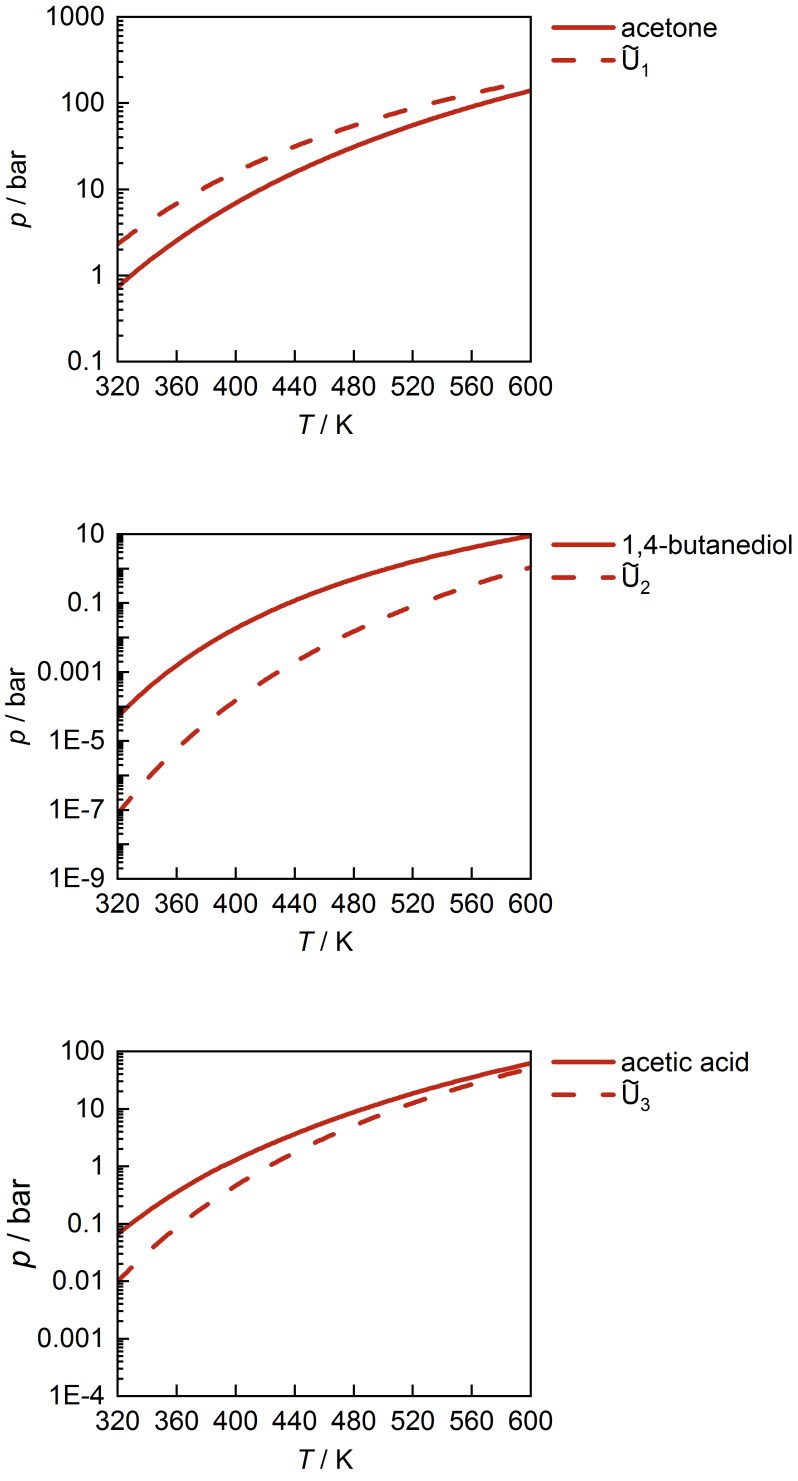

Figure S.1: Vapor pressures p^{S} of pure (pseudo-)components in mixture I. Solid lines: true components, calculated with the Antoine equation, cf. Eq. (S.3). Dashed lines: pseudo-components, estimated with the group-contribution method of Ref.,[12] cf. Eq. (S.4).

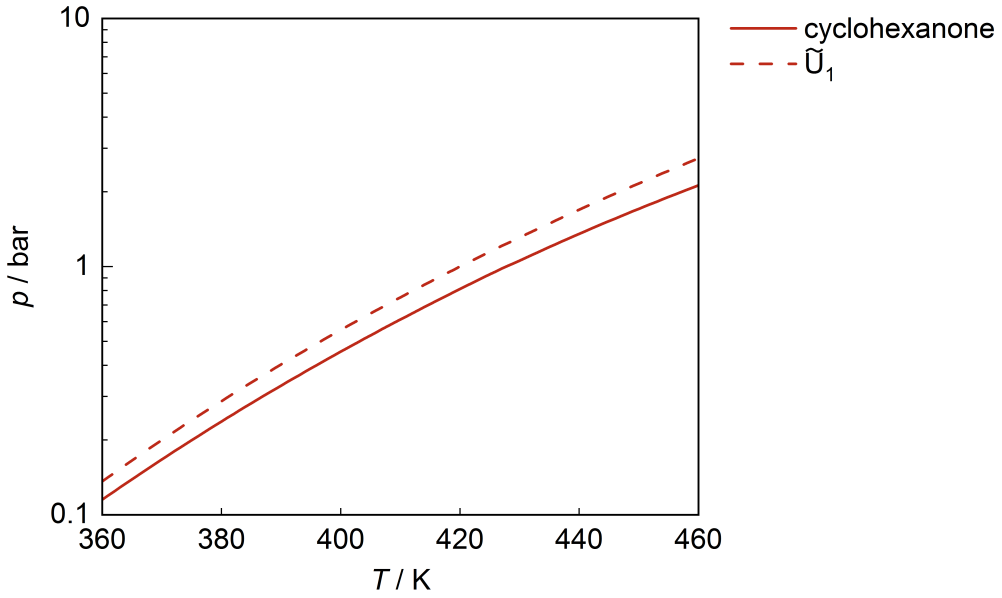

Figure S.2: Vapor pressures p^{S} of pure (pseudo-)components in mixture II. Solid lines: true component, calculated with the Antoine equation, cf. Eq. (S.3). Dashed lines: pseudo-component, estimated with the group-contribution method of Ref.,[12] cf. Eq. (S.4).

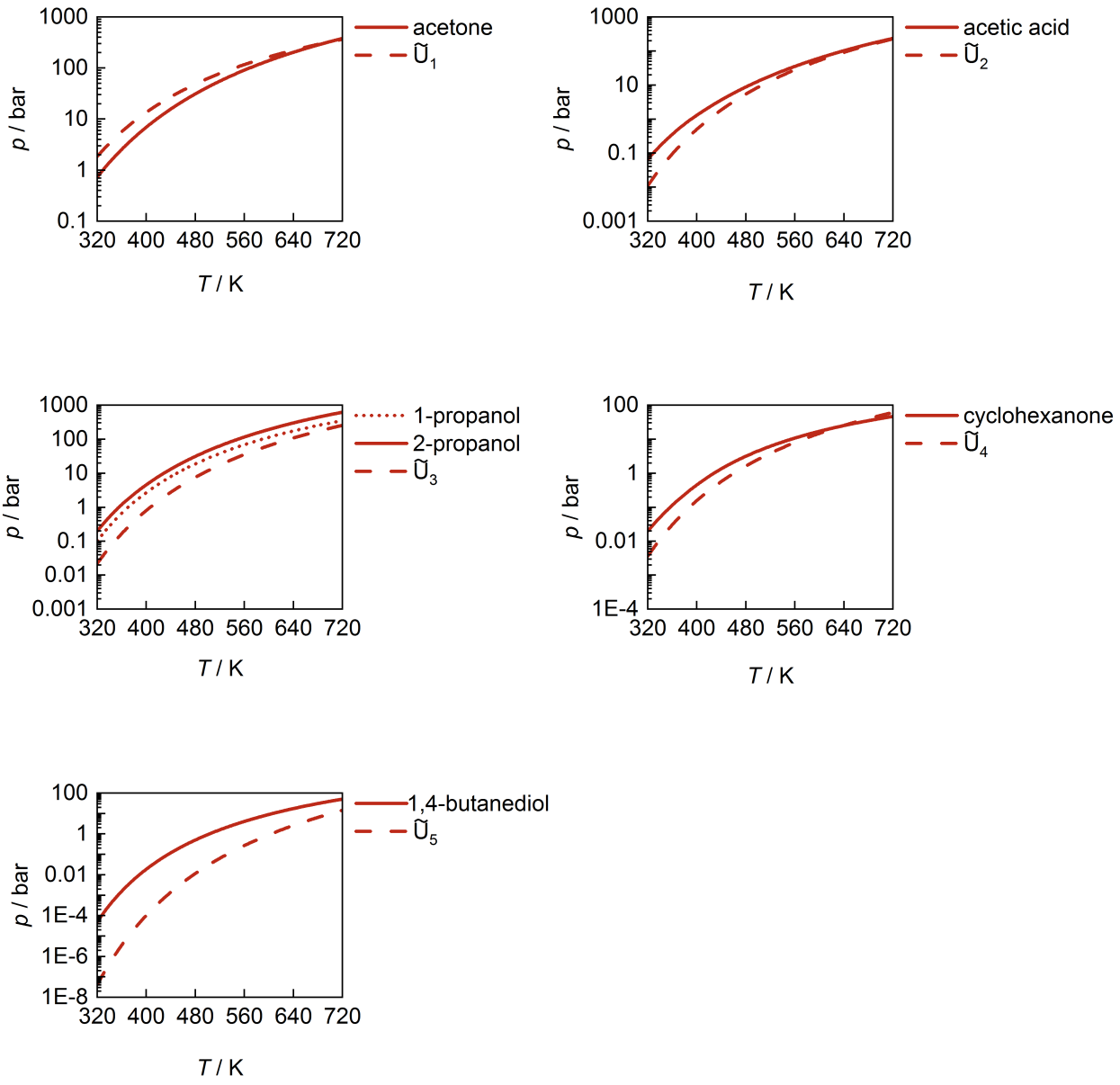

Figure S.3: Vapor pressures p^{S} of pure (pseudo-)components in mixture III. Solid and dotted lines: true components, calculated with the Antoine equation, cf. Eq. (S.3). Dashed lines: pseudo-components, estimated with the group-contribution method of Ref.,[12] cf. Eq. (S.4).

Additional Results for the Prediction of Residue Curves

In Figures S.4 and S.5, additional results for the residue curves of test mixtures II and III, cf. Table 2 in the manuscript, are shown. In contrast to the results shown in the manuscript, the concentration of all true components is presented separately here, even if they were lumped into pseudo-components by our algorithms.

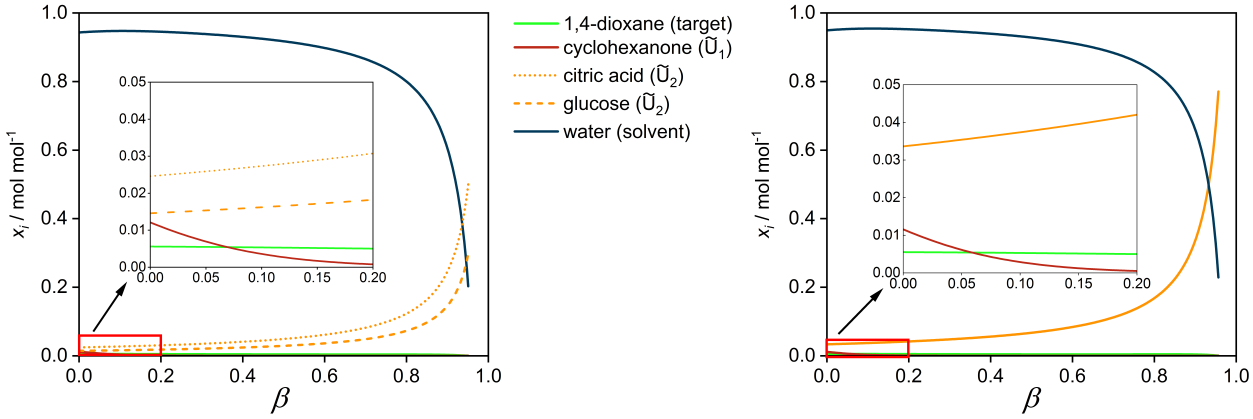

Figure S.4: Residue curves showing the liquid-phase mole fractions as a function of the evaporation ratio β for mixture II, cf. Table 2 in the manuscript, at $p = 1$ bar. Left: results obtained using the full speciation. Right: predictions for the poorly specified feed based on NMR fingerprinting and the pseudo-component method.

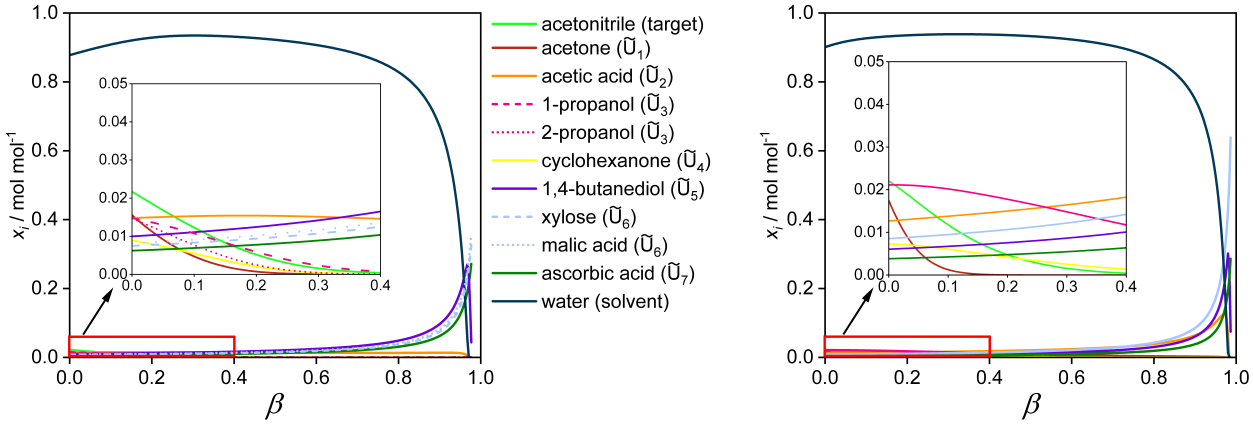

Figure S.5: Residue curves showing the liquid-phase mole fractions as a function of the evaporation ratio β for mixture III, cf. Table 2 in the manuscript, at $p = 1$ bar. Left: results obtained using the full speciation. Right: predictions for the poorly specified feed based on NMR fingerprinting and the pseudo-component method.

References

(1) Rachford, J., H.H.; Rice, J. Procedure for Use of Electronic Digital Computers in Calculating Flash Vaporization Hydrocarbon Equilibrium. *Journal of Petroleum Technology* **1952**, *4*, 19–3, DOI: 10.2118/952327-G.

(2) Teh, Y.; Rangaiah, G. A Study of Equation-Solving and Gibbs Free Energy Minimization Methods for Phase Equilibrium Calculations. *Chemical Engineering Research and Design* **2002**, *80*, 745–759, DOI: 10.1205/026387602320776821.

(3) Ohanomah, M.; Thompson, D. Computation of Multicomponent Phase Equilibria-Part I. Vapour-liquid Equilibria. *Computers & Chemical Engineering* **1984**, *8*, 147–156, DOI: 10.1016/0098-1354(84)87001-5.

(4) Nichita, D. V.; Leibovici, C. F. A Rapid and Robust Method for Solving the Rachford–Rice Equation Using Convex Transformations. *Fluid Phase Equilibria* **2013**, *353*, 38–49, DOI: 10.1016/j.fluid.2013.05.030.

(5) Fredenslund, A.; Jones, R. L.; Prausnitz, J. M. Group-Contribution Estimation of Activity Coefficients in Nonideal Liquid Mixtures. *AIChE Journal* **1975**, *21*, 1086–1099, DOI: 10.1002/aic.690210607.

(6) Wittig, R.; Lohmann, J.; Gmehling, J. Vapor-Liquid Equilibria by UNIFAC Group Contribution. 6. Revision and Extension. *Industrial & Engineering Chemistry Research* **2003**, *42*, 183–188, DOI: 10.1021/ie0205061.

(7) MATLAB, *version 9.11.0 (R2021b)*; The MathWorks Inc.: Natick, Massachusetts, 2021.

(8) Rayleigh, L. LIX. On the Distillation of Binary Mixtures. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **1902**, *4*, 521–537, DOI: 10.1080/14786440209462876.

(9) Stichlmair, J.; Klein, H.; Rehfeldt, S. *Distillation*, 2nd ed.; John Wiley & Sons, Inc.: Hoboken (New Jersey), 2021; DOI: 10.1002/9781119414674.

(10) Dortmund Data Bank. www.ddbst.com, 2021.

(11) Nannoolal, Y.; Rarey, J.; Ramjugernath, D.; Cordes, W. Estimation of Pure Component Properties: Part 1. Estimation of the Normal Boiling Point of Non-electrolyte Organic Compounds via Group Contributions and Group Interactions. *Fluid Phase Equilibria* **2004**, *226*, 45–63, DOI: 10.1016/j.fluid.2004.09.001.

(12) Nannoolal, Y.; Rarey, J.; Ramjugernath, D. Estimation of Pure Component Properties: Part 3. Estimation of the Vapor Pressure of Non-Electrolyte Organic Compounds via Group Contributions and Group Interactions. *Fluid Phase Equilib.* **2008**, *269*, 117–133, DOI: 10.1016/j.fluid.2008.04.020.

(13) Specht, T.; Münnemann, K.; Hasse, H.; Jirasek, F. Rational Method for Defining and Quantifying Pseudo-components Based on NMR Spectroscopy. *Physical Chemistry Chemical Physics* **2023**, *25*, 10288–10300, DOI: 10.1039/d3cp00509g.

(14) Online Group Assignment for UNIFAC. http://www.ddbst.com/unifacga.html (Last accessed: 30.03.2023).