Supporting Information for NEAT - NMR Spectroscopy for the Estimation of Activity Coefficients of Target Components in Poorly Specified Mixtures

Fabian Jirasek,[*,†] Jakob Burger,[‡] and Hans Hasse[¶]

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

‡*Chair of Chemical Process Engineering, Technical University of Munich (TUM), Campus Straubing for Biotechnology and Sustainability, Schulgasse 16, 94315 Straubing, Germany*

¶*Laboratory of Engineering Thermodynamics (LTD), Technische Universität Kaiserslautern (TUK), Erwin-Schrödinger-Straße 44, 67663 Kaiserslautern, Germany*

E-mail: fabian.jirasek@mv.uni-kl.de

Flowchart of NEAT procedure

Figure S.1 shows the procedure of the application of NEAT for estimating the activity coefficients of target components in poorly specified aqueous mixtures as used in this work based on the assignment in Table 3 in the main text. The procedure for the application to nonaqueous mixtures is analog with small differences as described in the main text.

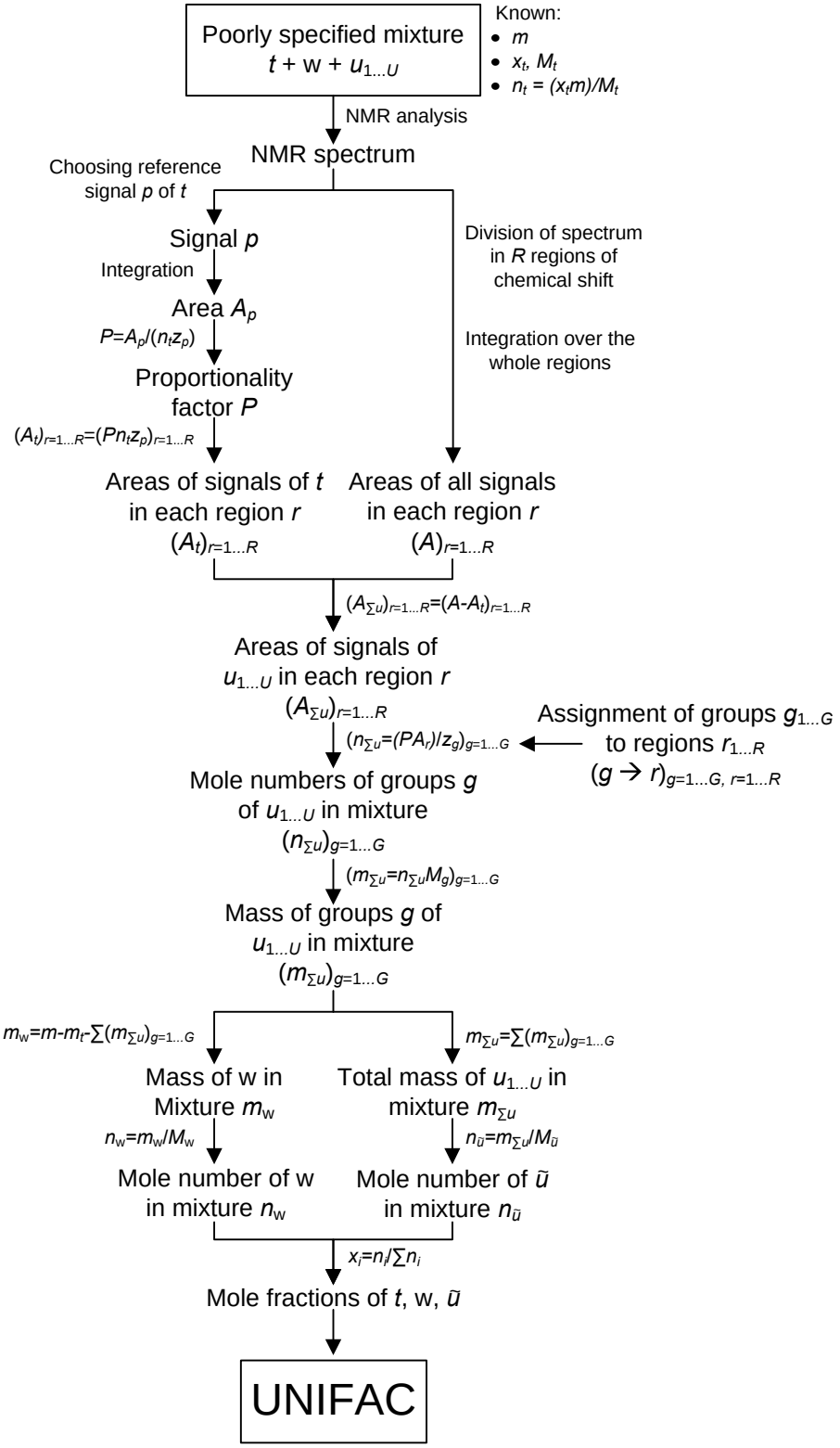

Figure S.1: Schematic representation of the procedure of NEAT for the application to aqueous poorly specified mixtures in this work. The assignment of groups to regions ($g \rightarrow r$) includes the consideration of multiple signals per group in different regions depending on the variant of NEAT.

Composition of formaldehyde-containing mixtures

In Tables S.1 and S.2, the overall compositions of the formaldehyde-containing mixtures from this work are summarized. Formaldehyde is reactive in these mixtures and forms polyoxymethylene glycols (MG_x) with water and polyoxymethylene hemiformals (HF_x) with methanol. In Tables S.3 and S.4, the true compositions of these mixtures are given as calculated with the activity-based chemical equilibrium model from Kuhnert et al.[1]

Table S.1: Overall compositions of the mixtures from Figure 4.

Mixture	1	2	3	4	5	6
$x_{\text{dimethoxymethane}}$ / g/g	0.078	0.114	0.092	0.088	0.088	0.200
x_{water} / g/g	0.907	0.858	0.865	0.855	0.849	0.568
$x_{\text{formaldehyde}}$ / g/g	0.015	0.029	0.044	0.057	0.063	0.232

Table S.2: Overall compositions of the mixtures from Figure 5.

Mixture	1	2	3	4	5
$x_{\text{dimethoxymethane}}$ / g/g	0.099	0.100	0.104	0.100	0.101
x_{water} / g/g	0.849	0.800	0.744	0.699	0.649
$x_{\text{formaldehyde}}$ / g/g	0.025	0.050	0.074	0.100	0.125
x_{methanol} / g/g	0.027	0.050	0.078	0.101	0.125

Table S.3: True compositions of the mixtures from Figure 4 calculated with the the activity-based chemical equilibrium model from Kuhnert et al.[1]

Mixture	1	2	3	4	5	6
$x_{\text{dimethoxymethane}}$ / g/g	0.078	0.114	0.092	0.088	0.088	0.200
x_{water} / g/g	0.898	0.842	0.841	0.826	0.817	0.490
$x_{\text{formaldehyde}}$ / g/g	0.000	0.000	0.000	0.000	0.000	0.000
x_{MG_1} / g/g	0.023	0.039	0.055	0.066	0.071	0.103
x_{MG_2} / g/g	0.001	0.005	0.010	0.016	0.020	0.105
x_{MG_3} / g/g	0.000	0.000	0.001	0.003	0.004	0.056
x_{MG_4} / g/g	0.000	0.000	0.000	0.000	0.001	0.026
x_{MG_5} / g/g	0.000	0.000	0.000	0.000	0.000	0.011
x_{MG_6} / g/g	0.000	0.000	0.000	0.000	0.000	0.005
x_{MG_7} / g/g	0.000	0.000	0.000	0.000	0.000	0.002
x_{MG_8} / g/g	0.000	0.000	0.000	0.000	0.000	0.001

Table S.4: True compositions of the mixtures from Figure 5 calculated with the the activity-based chemical equilibrium model from Kuhnert et al.[1]

Mixture	1	2	3	4	5
$x_{\text{dimethoxymethane}}$ / g/g	0.100	0.100	0.104	0.100	0.101
x_{water} / g/g	0.838	0.783	0.723	0.673	0.621
$x_{\text{formaldehyde}}$ / g/g	0.000	0.000	0.000	0.000	0.000
x_{methanol} / g/g	0.021	0.032	0.045	0.054	0.062
x_{MG_1} / g/g	0.027	0.042	0.049	0.056	0.059
x_{MG_2} / g/g	0.002	0.007	0.012	0.017	0.022
x_{MG_3} / g/g	0.000	0.001	0.002	0.003	0.004
x_{MG_4} / g/g	0.000	0.000	0.000	0.000	0.001
x_{HF_1} / g/g	0.012	0.032	0.057	0.082	0.108
x_{HF_2} / g/g	0.001	0.003	0.007	0.013	0.019
x_{HF_3} / g/g	0.000	0.000	0.001	0.001	0.002

Assignment of groups to components

In Table S.5, the groups that are assigned to the components of the present work for the fully specified mixtures according to modified UNIFAC (Dortmund)[2,3] are summarized. Included are all components that are part of the mixtures for which modified UNIFAC (Dortmund) was used to calculate the activity coefficients for the fully specified mixtures.

In Table S.6, the group assignment for all components that are part of the mixtures for which the chemical equilibrium model[1,4] based on original UNIFAC[5] was used to calculate the activity coefficients for the fully specified mixtures, is summarized.

In Tables S.7 - S.17, the estimated compositions of the poorly specified mixtures and the estimated stoichiometry of the mean unknown component $\tilde{u}$ calculated with NEAT for all experimentally studied mixtures of this work and all studied variants are summarized. The mass fraction of the target component is always the same as for the fully specified mixture.

Table S.5: Assignment of groups from modified UNIFAC (Dortmund) to the components used in this work for the fully specified mixtures. The labels and the numbers in parentheses are the identifiers for the subgroup and the corresponding main group from the original papers.[2,3]

Component	Modified UNIFAC (Dortmund) groups
acetic acid	1 x CH3 (1, 1), 1 x COOH (42, 20)
acetone	1 x CH3 (1, 1), 1 x CH3CO (18, 9)
acetonitrile	1 x CH3CN (40, 19)
L-ascorbic acid	1 x CH2 (2, 1), 1 x CH (3, 1), 1 x OH(P) (14, 5), 1 x C=C (70, 2), 1 x COO (77, 41), 1 CY-CH (79, 42), 1 x OH(S) (81, 5), 2 x OH(T) (82, 5)
1,4-butanediol	4 x CH2 (2, 1), 2 x OH(P) (14, 5)
2-butanone	1 x CH3 (1, 1), 1 x CH2 (2, 1), 1 x CH3CO (18, 9)
tert-butylhydroquinone	3 x CH3 (1, 1), 1 x C (4, 1), 3 x ACH (9, 3), 1 x AC (10, 3), 2 x ACOH (17, 8)
citric acid	2 x CH2 (2, 1), 1 x C (4, 1), 3 x COOH (42, 20,) 1 x OH(T) (82, 5)
cyclohexanone	1 x CH2CO (19, 9), 4 x CY-CH2 (78, 42)
dimethoxymethane	1 x CH2 (2, 1), 2 x CH3O (24, 13)
ethanol	1 x CH3 (1, 1), 1 x CH2 (2, 1), 1 x OH(P) (14, 5)
D-glucose	1 x CH2 (2, 1), 1 x OH(P) (14, 5), 1 x THF (27, 43), 3 x CY-CH (79, 42), 4 x OH(S) (81, 5)
methanol	1 x CH3OH (15, 6)
methyl acetate	1 x CH3 (1, 1), 1 x CH3COO (21, 11)
methyl oleate	2 x CH3 (1, 1), 13 x CH2 (2, 1), 1 x CH=CH (6, 2), 1 x CH2COO (22, 11)
methyl palmitate	2 x CH3 (1, 1), 13 x CH2 (2, 1), 1 x CH2COO (22, 11)
methyl stearate	2 x CH3 (1, 1), 15 x CH2 (2, 1), 1 x CH2COO (22, 11)
oleic acid	1 x CH3 (1, 1), 14 x CH2 (2, 1), 1 x CH=CH (6, 2), 1 x COOH (42, 20)
1-propanol	1 x CH3 (1, 1), 2 x CH2 (2, 1), 1 x OH(P) (14, 5)
pyridine	3 x ACH (9, 3), 1 x AC2H2N (37, 18)
water	1 x H2O (16, 7)
D-xylose	1 x THF (27, 43), 4 x OH(S) (81, 5), 3 x CY-CH (79, 42)

Table S.6: Assignment of groups from the chemical equilibrium model[1,4] based on original UNIFAC for the components for the fully specified mixtures.

Component	original UNIFAC groups
dimethoxymethane	1 x CH3OCH2OCH3
formaldehyde	1 x CH2O
methanol	1 x CH3OH

Table S.7: Estimated compositions in the component space (target component t + water w + mean unknown component $\tilde{u}$) for the poorly specified mixtures of the system (ascorbic acid t + water w + cyclohexanone u) and stoichiometry of $\tilde{u}$ for $M_{\tilde{u}} = 150$ g/mol obtained from ^{13}C NEAT, cf. Figure 1 in the main text. The labels are the identifiers for the subgroup from modified UNIFAC (Dortmund) according to the original papers.[2,3]

	x_i / g/g		Stoichiometry of $\tilde{u}$					
No.	x_t	$x_{\tilde{u}}$	'CH2'	'OH(P)'	'HC=CH'	'COOH'	'CH2CO'	Variant
1	0.099	0.019	5.586	0.044	0.111	0.310	1.286	^{13}C NEAT
2	0.098	0.031	6.536	0.000	0.000	0.000	1.387	
3	0.096	0.050	6.302	0.000	0.032	0.000	1.445	
4	0.094	0.066	6.348	0.097	0.012	0.012	1.391	
5	0.092	0.086	6.245	0.000	0.000	0.046	1.436	

Table S.8: Estimated compositions in the component space (target component t + water w + mean unknown component $\tilde{u}$) for the poorly specified mixtures of the system (1,4-butanediol t + water w + cyclohexanone u_1 + pyridine u_2 + citric acid u_3) and stoichiometry of $\tilde{u}$ for $M_{\tilde{u}} = 150$ g/mol obtained from ^{13}C NEAT, cf. Figure 1 in the main text. The labels are the identifiers for the subgroup from modified UNIFAC (Dortmund) according to the original papers.[2,3]

	x_i / g/g		Stoichiometry of $\tilde{u}$					
No.	x_t	$x_{\tilde{u}}$	'CH2'	'OH(P)'	'HC=CH'	'COOH'	'CH2CO'	Variant
1	0.099	0.017	1.875	0.375	1.781	0.000	1.688	^{13}C NEAT
2	0.094	0.058	1.890	0.378	1.620	0.000	1.782	
3	0.091	0.077	2.336	0.356	1.841	0.000	1.504	
4	0.089	0.096	2.307	0.308	1.984	0.000	1.445	
5	0.085	0.130	2.123	0.372	1.795	0.000	1.598	

Table S.9: Estimated compositions in the component space (target component t + water w + mean unknown component $\tilde{u}$) for the poorly specified mixtures of the system (acetic acid t + water w + xylose u) and stoichiometry of $\tilde{u}$ for $M_{\tilde{u}} = 150$ g/mol obtained from ^{13}C NEAT, cf. Figure 3 in the main text. The labels are the identifiers for the subgroup from modified UNIFAC (Dortmund) according to the original papers.[2,3]

	x_i / g/g		Stoichiometry of $\tilde{u}$					
No.	x_t	$x_{\tilde{u}}$	'CH2'	'OH(P)'	'HC=CH'	'COOH'	'CH2CO'	Variant
1	0.071	0.090	3.983	4.768	0.500	0.000	0.000	^{13}C NEAT
2	0.145	0.099	3.847	4.889	0.495	0.000	0.000	
3	0.239	0.099	3.746	4.989	0.483	0.000	0.000	
4	0.324	0.099	3.765	4.962	0.491	0.000	0.000	
5	0.402	0.098	3.700	5.059	0.462	0.000	0.000	
6	0.464	0.099	3.635	5.103	0.469	0.000	0.000	
7	0.525	0.096	3.689	5.045	0.478	0.000	0.000	

Table S.10: Estimated compositions in the component space (target component t + water w + mean unknown component $\tilde{u}$) for the poorly specified mixtures of the reactive system (dimethoxymethane t + water w + formaldehyde u) and stoichiometry of $\tilde{u}$ for $M_{\tilde{u}} = 150$ g/mol obtained from ^{13}C NEAT, cf. Figure 4 in the main text. The labels are the identifiers for the subgroup from original UNIFAC according to the original paper.[5]

	x_i / g/g		Stoichiometry of $\tilde{u}$					
No.	x_t	$x_{\tilde{u}}$	'CH2'	'OH(P)'	'HC=CH'	'COOH'	'CH2CO'	Variant
1	0.078	0.016	4.869	4.803	0.000	0.000	0.000	^{13}C NEAT
2	0.114	0.029	4.833	4.833	0.000	0.000	0.000	
3	0.092	0.047	4.827	4.827	0.000	0.000	0.000	
4	0.088	0.061	4.828	4.828	0.006	0.000	0.000	
5	0.088	0.068	4.825	4.818	0.014	0.000	0.000	
6	0.200	0.258	4.833	4.833	0.000	0.000	0.000	

Table S.11: Estimated compositions in the component space (target component t + water w + mean unknown component $\tilde{u}$) for the poorly specified mixtures of the reactive system (dimethoxymethane t + water w + formaldehyde u_1 + methanol u_2) and stoichiometry of $\tilde{u}$ for $M_{\tilde{u}} = 150$ g/mol obtained from ^{13}C NEAT, cf. Figure 5 in the main text. The labels are the identifiers for the subgroup from original UNIFAC according to the original paper.[5]

	x_i / g/g		Stoichiometry of $\tilde{u}$					
No.	x_t	$x_{\tilde{u}}$	'CH2'	'OH(P)'	'HC=CH'	'COOH'	'CH2CO'	Variant
1	0.100	0.033	6.570	2.445	0.624	0.000	0.000	^{13}C NEAT
2	0.100	0.067	6.202	2.393	0.857	0.000	0.000	
3	0.104	0.090	6.187	2.053	1.087	0.000	0.000	
4	0.100	0.116	6.109	1.936	1.205	0.000	0.000	
5	0.101	0.142	6.061	1.833	1.298	0.000	0.000	

Table S.12: Estimated compositions in the component space (target component t + solvent s + mean unknown component $\tilde{u}$) for the poorly specified mixtures of the system (ethanol t + oleic acid u_1 + *tert*-butylhydroquinone (TBHQ) u_2 + acetone u_3 + 1-propanol u_4) and stoichiometry of $\tilde{u}$ for $M_{\tilde{u}} = 150$ g/mol obtained from ^{13}C NEAT and ^{13}C NEAT, known solvent $s = u_1$, cf. Figure 6 in the main text. The labels are the identifiers for the subgroup from modified UNIFAC (Dortmund) according to the original papers.[2,3]

	x_i / g/g		Stoichiometry of $\tilde{u}$					
No.	x_t	$x_{\tilde{u}}$	'CH2'	'OH(P)'	'HC=CH'	'COOH'	'CH2CO'	Variant
1	0.048	0.952	7.765	0.038	0.524	0.555	0.043	^{13}C NEAT
2	0.046	0.954	7.551	0.082	0.519	0.569	0.046	
3	0.040	0.960	7.130	0.170	0.530	0.572	0.040	
4	0.035	0.965	6.719	0.259	0.529	0.581	0.035	
5	0.030	0.970	6.366	0.349	0.530	0.575	0.030	
1	0.048	0.163	6.848	0.191	0.489	0.643	0.214	^{13}C NEAT
2	0.046	0.212	6.154	0.341	0.473	0.682	0.353	known s
3	0.040	0.294	5.377	0.515	0.525	0.650	0.544	
4	0.035	0.383	4.959	0.616	0.524	0.648	0.645	
5	0.030	0.468	4.913	0.661	0.527	0.612	0.679	

Table S.13: Estimated compositions in the component space (target component t + solvent s + mean unknown component $\tilde{u}$) for the poorly specified mixtures of the system (ethanol t + oleic acid u_1 + *tert*-butylhydroquinone (TBHQ) u_2 + acetone u_3 + 1-propanol u_4) and stoichiometry of $\tilde{u}$ for $M_{\tilde{u}} = 150$ g/mol obtained from ^{1}H NEAT and ^{1}H NEAT, known solvent $s = u_1$, cf. Figure 6 in the main text. The labels are the identifiers for the subgroup from modified UNIFAC (Dortmund) according to the original papers.[2,3]

	x_i / g/g		Stoichiometry of $\tilde{u}$						
No.	x_t	$x_{\tilde{u}}$	'CH2'	'OH(P)'	'HC=CH'	'CH2COO'	'ACH'	'C'	Variant
1	0.048	0.952	8.000	0.522	0.535	0.049	0.049	0.954	^{1}H NEAT
2	0.046	0.954	7.846	0.544	0.505	0.093	0.099	0.905	
3	0.040	0.960	7.719	0.578	0.448	0.181	0.203	0.589	
4	0.035	0.965	7.321	0.688	0.378	0.258	0.296	0.575	
5	0.030	0.970	7.284	0.753	0.326	0.363	0.421	0.000	
1	0.048	0.039	4.007	0.348	0.488	1.045	1.115	0.000	^{1}H NEAT
2	0.046	0.100	3.159	0.396	0.172	0.825	0.941	2.860	known s
3	0.040	0.200	4.645	0.681	0.092	0.835	0.957	0.833	
4	0.035	0.303	4.600	0.987	0.038	0.807	0.923	0.740	
5	0.030	0.419	5.122	1.034	0.026	0.807	0.963	0.044	

Table S.14: Estimated compositions in the component space (target component t + solvent s + mean unknown component $\tilde{u}$) for the poorly specified mixtures of the system (ethanol t + methyl oleate u_1 + methyl stearate u_2 + methyl palmitate u_3 + *tert*-butylhydroquinone (TBHQ) u_4) and stoichiometry of $\tilde{u}$ for $M_{\tilde{u}} = 150$ g/mol obtained from ^{13}C NEAT and ^{13}C NEAT, known solvent $s = u_1$, cf. Figure 8 in the main text. The labels are the identifiers for the subgroup from modified UNIFAC (Dortmund) according to the original papers.[2,3]

	x_i / g/g		Stoichiometry of $\tilde{u}$					
No.	x_t	$x_{\tilde{u}}$	'CH2'	'OH(P)'	'HC=CH'	'COOH'	'CH2CO'	Variant
1	0.010	0.990	8.094	0.005	0.382	0.587	0.001	^{13}C NEAT
2	0.020	0.980	8.089	0.000	0.391	0.585	0.001	
3	0.030	0.970	8.095	0.003	0.388	0.580	0.005	
4	0.040	0.960	8.128	0.000	0.383	0.577	0.001	
5	0.048	0.952	8.116	0.003	0.390	0.574	0.003	
1	0.010	0.341	8.057	0.010	0.220	0.689	0.001	^{13}C NEAT
2	0.020	0.350	8.047	0.000	0.243	0.683	0.001	known s
3	0.030	0.378	8.060	0.006	0.232	0.674	0.011	
4	0.040	0.412	8.136	0.000	0.220	0.667	0.003	
5	0.048	0.397	8.108	0.007	0.240	0.657	0.007	

Table S.15: Estimated compositions in the component space (target component t + solvent s + mean unknown component $\tilde{u}$) for the poorly specified mixtures of the system (ethanol t + methyl oleate u_1 + methyl stearate u_2 + methyl palmitate u_3 + *tert*-butylhydroquinone (TBHQ) u_4) and stoichiometry of $\tilde{u}$ for $M_{\tilde{u}} = 150$ g/mol obtained from ^{1}H NEAT and ^{1}H NEAT, known solvent $s = u_1$, cf. Figure 8 in the main text. The labels are the identifiers for the subgroup from modified UNIFAC (Dortmund) according to the original papers.[2,3]

	x_i / g/g		Stoichiometry of $\tilde{u}$						
No.	x_t	$x_{\tilde{u}}$	'CH2'	'OH(P)'	'HC=CH'	'CH2COO'	'ACH'	'C'	Variant
1	0.010	0.990	7.127	0.004	0.322	0.687	0.128	0.000	^{1}H NEAT
2	0.020	0.980	6.560	0.042	0.301	0.628	0.082	0.994	
3	0.030	0.970	7.110	0.048	0.325	0.688	0.084	0.000	
4	0.040	0.960	7.051	0.049	0.325	0.678	0.084	0.117	
5	0.048	0.952	7.096	0.051	0.330	0.687	0.089	0.000	
1	0.010	0.300	7.250	0.015	0.001	0.753	0.332	0.000	^{1}H NEAT
2	0.020	0.358	5.950	0.125	0.021	0.583	0.245	2.237	known s
3	0.030	0.324	6.518	0.149	0.006	0.637	0.263	1.291	
4	0.040	0.349	5.600	0.134	0.007	0.535	0.231	2.907	
5	0.048	0.329	6.503	0.153	0.038	0.637	0.268	1.224	

Table S.16: Estimated compositions in the component space (target component t + solvent s + mean unknown component $\tilde{u}$) for the poorly specified mixtures of the system (dimethoxymethane t + methyl oleate u_1 + *tert*-butylhydroquinone (TBHQ) u_2 + methanol u_3 + acetone u_4) and stoichiometry of $\tilde{u}$ for $M_{\tilde{u}} = 150$ g/mol obtained from ^{13}C NEAT and ^{13}C NEAT, known solvent $s = u_1$, cf. Figure 9 in the main text. The labels are the identifiers for the subgroup from modified UNIFAC (Dortmund) according to the original papers.[2,3]

	x_i / g/g		Stoichiometry of $\tilde{u}$					
No.	x_t	$x_{\tilde{u}}$	'CH2'	'OH(P)'	'HC=CH'	'COOH'	'CH2CO'	Variant
1	0.200	0.980	7.645	0.049	0.522	0.581	0.052	^{13}C NEAT
2	0.035	0.965	7.641	0.001	0.513	0.567	0.092	
3	0.074	0.926	7.374	0.138	0.470	0.521	0.203	
4	0.106	0.894	7.087	0.345	0.425	0.466	0.302	
5	0.135	0.865	6.922	0.458	0.386	0.425	0.379	
1	0.200	0.246	6.309	0.104	0.570	0.805	0.206	^{13}C NEAT
2	0.035	0.227	6.250	0.005	0.529	0.740	0.360	known s
3	0.074	0.292	5.813	0.373	0.391	0.554	0.642	
4	0.106	0.340	5.445	0.858	0.293	0.400	0.795	
5	0.135	0.375	5.393	1.017	0.231	0.319	0.873	

Table S.17: Estimated compositions in the component space (target component t + solvent s + mean unknown component $\tilde{u}$) for the poorly specified mixtures of the system (dimethoxymethane t + methyl oleate u_1 + *tert*-butylhydroquinone (TBHQ) u_2 + methanol u_3 + acetone u_4) and stoichiometry of $\tilde{u}$ for $M_{\tilde{u}} = 150$ g/mol obtained from ^{1}H NEAT and ^{1}H NEAT, known solvent $s = u_1$, cf. Figure 9 in the main text. The labels are the identifiers for the subgroup from modified UNIFAC (Dortmund) according to the original papers.[2,3]

	x_i / g/g		Stoichiometry of $\tilde{u}$						
No.	x_t	$x_{\tilde{u}}$	'CH2'	'OH(P)'	'HC=CH'	'CH2COO'	'ACH'	'C'	Variant
1	0.020	0.980	6.414	0.067	0.440	0.788	0.131	0.000	^{1}H NEAT
2	0.035	0.965	6.231	0.066	0.381	0.865	0.105	0.000	
3	0.074	0.926	5.532	0.059	0.329	1.060	0.099	0.000	
4	0.106	0.894	4.998	0.053	0.285	1.215	0.084	0.000	
5	0.135	0.865	5.297	0.054	0.284	1.142	0.085	0.000	
1	0.020	0.059	0.678	0.678	0.327	1.779	1.321	0.000	^{1}H NEAT
2	0.035	0.101	1.873	0.420	0.000	1.860	0.665	0.000	known s
3	0.074	0.158	1.576	0.214	0.008	2.057	0.357	0.000	
4	0.106	0.224	1.550	0.142	0.004	2.116	0.225	0.000	
5	0.135	0.268	2.371	0.145	0.004	1.917	0.226	0.000	

Additional results for aqueous poorly specified mixtures

In Figures S.2 - S.4, results for the prediction of the activity coefficients of target components t in aqueous poorly specified mixtures with NEAT are shown. The mixtures have already been studied in an earlier work of our group.[6] The predictions in the earlier work were obtained based on ^{13}C NMR spectroscopy and information on t as well as on the water mass fraction in the respective mixture. In the present work, predictions for the same mixtures were made based on ^{13}C NMR spectroscopy and only information on t. The accuracy of the predictions of this work is very similar to that of the predictions from the earlier work. Hence, the results of this work show that also for the poorly specified mixtures studied in the earlier work, excellent predictions for the activity coefficients of target components can be obtained with NEAT based on ^{13}C NMR spectroscopy and only information on t.

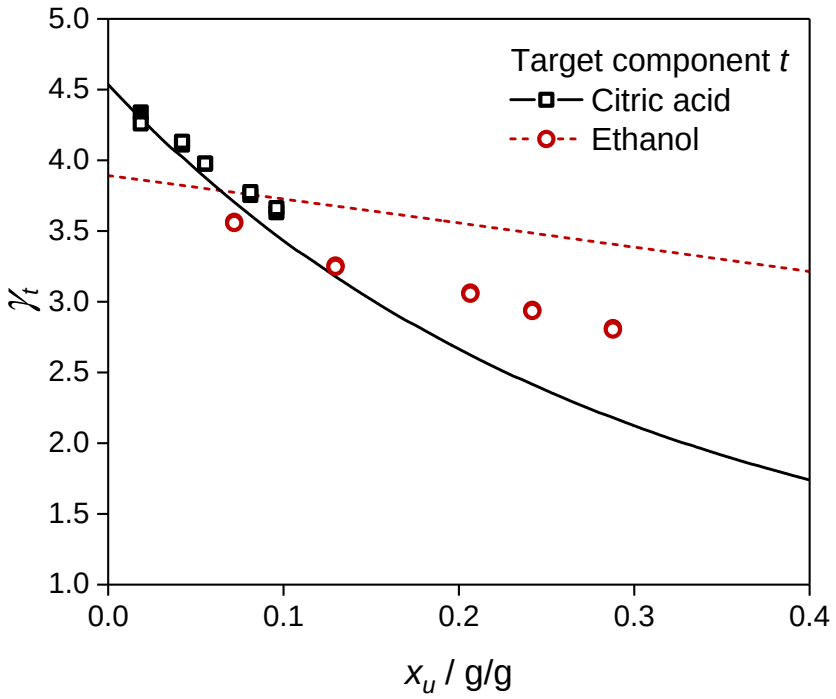

Figure S.2: Activity coefficients γ_t of target components (t = citric acid or ethanol) in ternary mixtures of t with water (w) and one unknown component (u = glucose) at 298 K and 1 bar. The mass ratio of (t : w) is (0.1 : 0.9) for all mixtures. Lines: results from modified UNIFAC (Dortmund) for the fully specified mixtures. Open symbols: predictions with NEAT based on ^{13}C NMR and only information on t, this work. Filled symbols: predictions with NEAT based on ^{13}C NMR and information on t and the water mass fraction, previous work.[6]

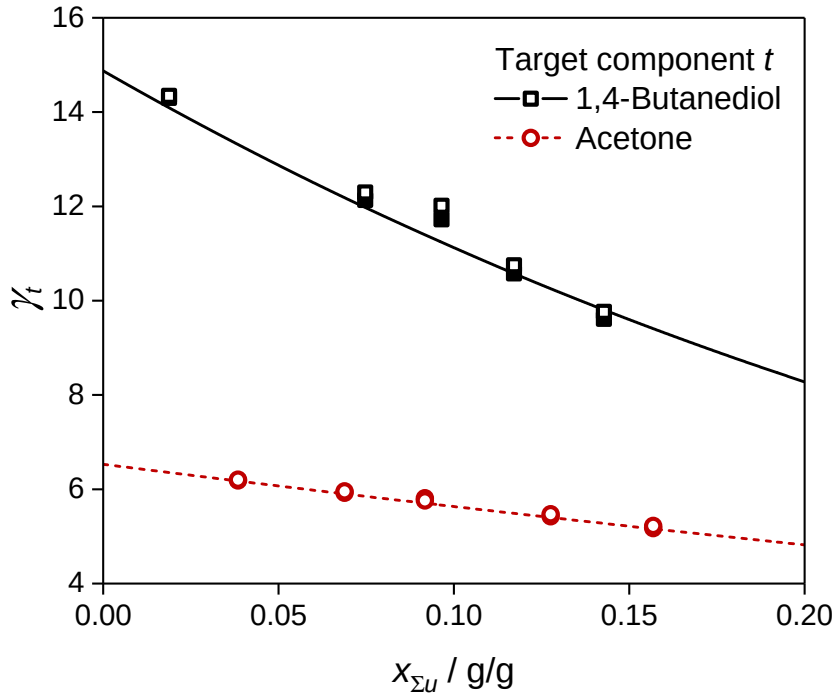

Figure S.3: Activity coefficients γ_t of target component (t = 1,4-butanediol or acetone) in five-component mixtures of t with water (w) and three unknown components (u = cyclohexanone, acetonitrile, methyl acetate or xylose, acetic acid, methyl acetate, each in the same mass ratio) at 298 K and 1 bar. The mass ratio of (t : w) is (0.1 : 0.9) for all mixtures. Lines: results from modified UNIFAC (Dortmund) for the fully specified mixtures. Open symbols: predictions with NEAT based on ^{13}C NMR and only information on t, this work. Filled symbols (hidden under the open symbols): predictions with NEAT based on ^{13}C NMR and information on t and the water mass fraction, previous work.[6]

In the following Tables, the estimated composition of the poorly specified mixtures and the estimated stoichiometry of the mean unknown component $\tilde{u}$ for the mixtures from Figures S.2 - S.4 calculated with the variants of NEAT from the previous work[6] and from the present work are summarized. The mass fraction of the target component is always the same as for the fully specified mixture.

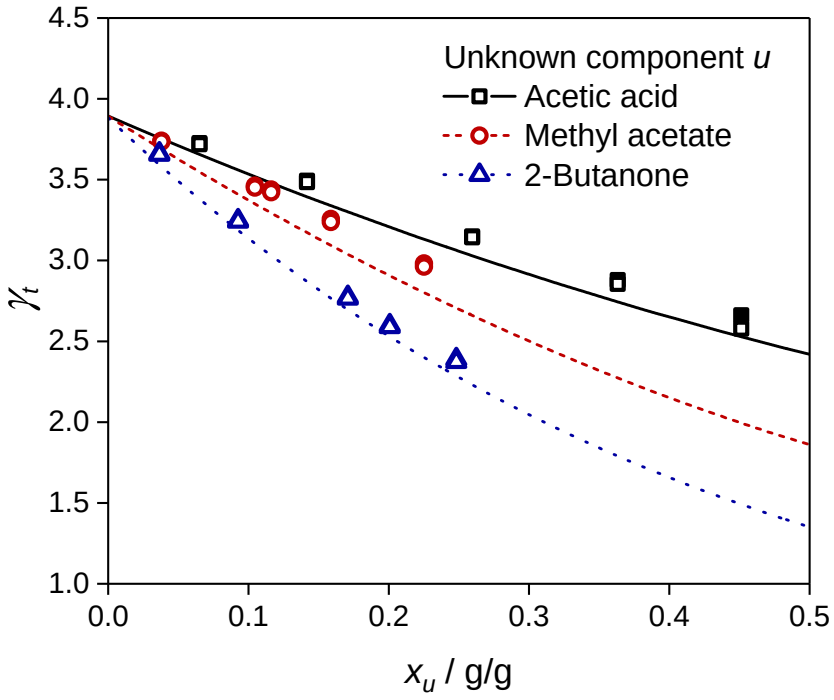

Figure S.4: Activity coefficient γ_t of target component (t = ethanol) in ternary mixtures of t with water (w) and one unknown components (u = acetic acid or methyl acetate or 2-butanone) at 298 K and 1 bar. The mass ratio of (t : w) is (0.1 : 0.9) for all mixtures. Lines: results from modified UNIFAC (Dortmund) for the fully specified mixtures. Open symbols: predictions with NEAT based on ^{13}C NMR and only information on t, this work. Filled symbols (mostly hidden under the open symbols): predictions with NEAT based on ^{13}C NMR and information on t and the water mass fraction, previous work.[6]

Table S.18: Estimated compositions in the component space (target component t + water w + mean unknown component $\tilde{u}$) for the poorly specified mixtures of the system (citric acid t + water + glucose u) and stoichiometry of $\tilde{u}$ for $M_{\tilde{u}} = 150$ g/mol obtained from ^{13}C NEAT using information on the water mass fraction (results from an earlier work[6]) and without using information on the water mass fraction (results from this work), cf. Figure S.2. The labels are the identifiers for the subgroup from modified UNIFAC (Dortmund) according to the original papers.[2,3]

	x_i / g/g		Stoichiometry of $\tilde{u}$					
No.	x_t	$x_{\tilde{u}}$	'CH2'	'OH(P)'	'HC=CH'	'COOH'	'CH2CO'	Variant
1	0.101	0.025	3.238	3.238	0.254	0.952	0.000	Ref.[6]
2	0.098	0.042	4.253	4.643	0.436	0.000	0.000	
3	0.097	0.055	4.045	4.090	0.437	0.273	0.000	
4	0.094	0.081	4.167	4.457	0.417	0.109	0.000	
5	0.093	0.096	4.113	4.485	0.407	0.121	0.000	
1	0.101	0.025	3.238	3.238	0.254	0.952	0.000	this work
2	0.098	0.039	4.568	4.334	0.469	0.000	0.000	
3	0.097	0.055	4.091	4.036	0.442	0.276	0.000	
4	0.094	0.078	4.327	4.289	0.433	0.113	0.000	
5	0.093	0.092	4.302	4.286	0.425	0.126	0.000	

Table S.19: Estimated compositions in the component space (target component t + water w + mean unknown component $\tilde{u}$) for the poorly specified mixtures of the system (ethanol t + water + glucose u) and stoichiometry of $\tilde{u}$ for $M_{\tilde{u}} = 150$ g/mol obtained from ^{13}C NEAT using information on the water mass fraction (results from an earlier work[6]) and without using information on the water mass fraction (results from this work), cf. Figure S.2. The labels are the identifiers for the subgroup from modified UNIFAC (Dortmund) according to the original papers.[2,3]

	x_i / g/g		Stoichiometry of $\tilde{u}$					
No.	x_t	$x_{\tilde{u}}$	'CH2'	'OH(P)'	'HC=CH'	'COOH'	'CH2CO'	Variant
1	0.102	0.072	3.789	5.093	0.393	0.000	0.000	Ref.[6]
2	0.111	0.130	3.856	5.039	0.391	0.000	0.000	
3	0.087	0.207	4.106	4.802	0.411	0.000	0.000	
4	0.085	0.242	4.101	4.809	0.410	0.000	0.000	
5	0.080	0.288	4.046	4.864	0.404	0.000	0.000	
1	0.102	0.061	4.446	4.446	0.461	0.000	0.000	this work
2	0.111	0.112	4.454	4.454	0.452	0.000	0.000	
3	0.087	0.190	4.474	4.444	0.447	0.000	0.000	
4	0.085	0.222	4.478	4.440	0.448	0.000	0.000	
5	0.080	0.260	4.475	4.445	0.446	0.000	0.000	

Table S.20: Estimated compositions in the component space (target component t + water w + mean unknown component $\tilde{u}$) for the poorly specified mixtures of the system (1,4-butanediol t + water + cyclohexanone u_1 + acetonitrile u_2 + methyl acetate u_3) and stoichiometry of $\tilde{u}$ for $M_{\tilde{u}} = 150$ g/mol obtained from ^{13}C NEAT using information on the water mass fraction (results from an earlier work[6]) and without using information on the water mass fraction (results from this work), cf. Figure S.3. The labels are the identifiers for the subgroup from modified UNIFAC (Dortmund) according to the original papers.[2,3]

	x_i / g/g		Stoichiometry of $\tilde{u}$					
No.	x_t	$x_{\tilde{u}}$	'CH2'	'OH(P)'	'HC=CH'	'COOH'	'CH2CO'	Variant
1	0.096	0.019	3.974	2.769	0.448	0.457	0.355	Ref.[6]
2	0.090	0.075	4.231	2.142	0.534	0.562	0.357	
3	0.088	0.096	3.733	2.904	0.470	0.516	0.304	
4	0.086	0.117	4.460	1.695	0.542	0.635	0.379	
5	0.083	0.143	4.595	1.333	0.573	0.658	0.436	
1	0.096	0.013	5.793	0.000	0.653	0.666	0.518	this work
2	0.090	0.057	5.588	0.000	0.705	0.742	0.472	
3	0.088	0.065	5.566	0.000	0.701	0.769	0.453	
4	0.086	0.95	5.521	0.000	0.671	0.786	0.469	
5	0.083	0.121	5.414	0.000	0.675	0.775	0.514	

Table S.21: Estimated compositions in the component space (target component t + water w + mean unknown component $\tilde{u}$) for the poorly specified mixtures of the system (acetone t + water + xylose u_1 + acetic acid u_2 + methyl acetate u_3) and stoichiometry of $\tilde{u}$ for $M_{\tilde{u}} = 150$ g/mol obtained from ^{13}C NEAT using information on the water mass fraction (results from an earlier work[6]) and without using information on the water mass fraction (results from this work), cf. Figure S.3. The labels are the identifiers for the subgroup from modified UNIFAC (Dortmund) according to the original papers.[2,3]

	x_i / g/g		Stoichiometry of $\tilde{u}$					
No.	x_t	$x_{\tilde{u}}$	'CH2'	'OH(P)'	'HC=CH'	'COOH'	'CH2CO'	Variant
1	0.096	0.038	2.975	2.645	0.194	1.294	0.000	Ref.[6]
2	0.093	0.069	3.080	2.916	0.140	1.190	0.000	
3	0.091	0.092	3.126	2.750	0.154	1.230	0.000	
4	0.087	0.128	3.078	2.506	0.142	1.344	0.000	
5	0.084	0.157	3.199	2.507	0.139	1.307	0.000	
1	0.096	0.031	3.640	1.266	0.237	1.582	0.000	this work
2	0.093	0.056	3.792	1.551	0.172	1.465	0.000	
3	0.091	0.075	3.804	1.434	0.187	1.497	0.000	
4	0.087	0.108	3.646	1.341	0.168	1.592	0.000	
5	0.084	0.133	3.772	1.378	0.164	1.541	0.000	

Table S.22: Estimated compositions in the component space (target component t + water w + mean unknown component $\tilde{u}$) for the poorly specified mixtures of the system (ethanol t + water + acetic acid u) and stoichiometry of $\tilde{u}$ for $M_{\tilde{u}} = 150$ g/mol obtained from ^{13}C NEAT using information on the water mass fraction (results from an earlier work[6]) and without using information on the water mass fraction (results from this work), cf. Figure S.4. The labels are the identifiers for the subgroup from modified UNIFAC (Dortmund) according to the original papers.[2,3]

	x_i / g/g		Stoichiometry of $\tilde{u}$					
No.	x_t	$x_{\tilde{u}}$	'CH2'	'OH(P)'	'HC=CH'	'COOH'	'CH2CO'	Variant
1	0.094	0.065	2.198	0.693	0.000	2.385	0.000	Ref.[6]
2	0.086	0.142	2.425	0.349	0.000	2.445	0.000	
3	0.074	0.259	2.519	0.099	0.000	2.510	0.000	
4	0.064	0.363	2.515	0.000	0.000	2.548	0.000	
5	0.055	0.451	2.529	0.000	0.000	2.544	0.000	
1	0.094	0.060	2.386	0.000	0.000	2.589	0.000	this work
2	0.086	0.136	2.525	0.000	0.000	2.545	0.000	
3	0.074	0.257	2.548	0.000	0.000	2.538	0.000	
4	0.064	0.370	2.515	0.000	0.000	2.548	0.000	
5	0.055	0.479	2.529	0.000	0.000	2.544	0.000	

Table S.23: Estimated compositions in the component space (target component t + water w + mean unknown component $\tilde{u}$) for the poorly specified mixtures of the system (ethanol t + water + methyl acetate u) and stoichiometry of $\tilde{u}$ for $M_{\tilde{u}} = 150$ g/mol obtained from ^{13}C NEAT using information on the water mass fraction (results from an earlier work[6]) and without using information on the water mass fraction (results from this work), cf. Figure S.4. The labels are the identifiers for the subgroup from modified UNIFAC (Dortmund) according to the original papers.[2,3]

	x_i / g/g		Stoichiometry of $\tilde{u}$					
No.	x_t	$x_{\tilde{u}}$	'CH2'	'OH(P)'	'HC=CH'	'COOH'	'CH2CO'	Variant
1	0.097	0.038	3.640	1.000	0.000	1.820	0.000	Ref.[6]
2	0.090	0.105	3.857	1.051	0.000	1.733	0.000	
3	0.088	0.116	3.758	1.335	0.000	1.657	0.000	
4	0.084	0.159	3.841	1.072	0.000	1.730	0.000	
5	0.078	0.225	3.924	0.761	0.000	1.822	0.000	
1	0.097	0.034	4.106	0.000	0.000	2.053	0.000	this work
2	0.090	0.093	4.348	0.063	0.000	1.953	0.000	
3	0.088	0.099	4.429	0.000	0.000	1.952	0.000	
4	0.084	0.140	4.363	0.020	0.000	1.965	0.000	
5	0.078	0.206	4.295	0.000	0.000	1.994	0.000	

Table S.24: Estimated compositions in the component space (target component t + water w + mean unknown component $\tilde{u}$) for the poorly specified mixtures of the system (ethanol t + water + 2-butanone u) and stoichiometry of $\tilde{u}$ for $M_{\tilde{u}} = 150$ g/mol obtained from ^{13}C NEAT using information on the water mass fraction (results from an earlier work[6]) and without using information on the water mass fraction (results from this work), cf. Figure S.4. The labels are the identifiers for the subgroup from modified UNIFAC (Dortmund) according to the original papers.[2,3]

	x_i / g/g		Stoichiometry of $\tilde{u}$					
No.	x_t	$x_{\tilde{u}}$	'CH2'	'OH(P)'	'HC=CH'	'COOH'	'CH2CO'	Variant
1	0.098	0.037	3.744	1.858	0.000	0.000	1.567	Ref.[6]
2	0.092	0.093	4.199	0.726	0.000	0.000	1.873	
3	0.084	0.171	4.163	0.597	0.000	0.000	1.938	
4	0.081	0.201	4.282	0.549	0.000	0.000	1.918	
5	0.076	0.248	4.108	0.663	0.000	0.000	1.929	
1	0.098	0.029	4.743	0.000	0.000	0.000	1.986	this work
2	0.092	0.085	4.576	0.000	0.000	0.000	2.041	
3	0.084	0.159	4.466	0.000	0.000	0.000	2.078	
4	0.081	0.188	4.566	0.000	0.000	0.000	2.045	
5	0.076	0.229	4.442	0.000	0.000	0.000	2.086	

References

(1) Kuhnert, C.; Albert, M.; Breyer, S.; Hahnenstein, I.; Hasse, H.; Maurer, G. Phase Equilibrium in Formaldehyde Containing Multicomponent Mixtures: Experimental Results for Fluid Phase Equilibria of (Formaldehyde + (Water or Methanol) + Methylal)) and (Formaldehyde + Water + Methanol + Methylal) and Comparison with Predictions. *Industrial & Engineering Chemistry Research* **2006**, *45*, 5155–5164.

(2) Weidlich, U.; Gmehling, J. A Modified UNIFAC Model. 1. Prediction of VLE, h^E, and γ^∞. *Industrial & Engineering Chemistry Research* **1987**, *26*, 1372–1381.

(3) Constantinescu, D.; Gmehling, J. Further Development of Modified UNIFAC (Dortmund): Revision and Extension 6. *Journal of Chemical & Engineering Data* **2016**, *61*, 2738–2748.

(4) Schmitz, N.; Homberg, F.; Berje, J.; Burger, J.; Hasse, H. Chemical Equilibrium of the Synthesis of Poly(oxymethylene) Dimethyl Ethers from Formaldehyde and Methanol in Aqueous Solutions. *Industrial & Engineering Chemistry Research* **2015**, *54*, 6409–6417.

(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.

(6) Jirasek, F.; Burger, J.; Hasse, H. Method for Estimating Activity Coefficients of Target Components in Poorly Specified Mixtures. *Industrial & Engineering Chemistry Research* **2018**, *57*, 7310–7313.