

## Dispersion activity coefficient models. Part 3

**Citation for published version (APA):**

Krooshof, G. J. P., Tuinier, R., & de With, G. (2021). Dispersion activity coefficient models. Part 3: A topology preserving group contribution model. *Fluid Phase Equilibria*, 544-545, Article 113097. <https://doi.org/10.1016/j.fluid.2021.113097>

**Document license:**

TAVERNE

**DOI:**

[10.1016/j.fluid.2021.113097](https://doi.org/10.1016/j.fluid.2021.113097)

**Document status and date:**

Published: 15/09/2021

**Document Version:**

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

**Please check the document version of this publication:**

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

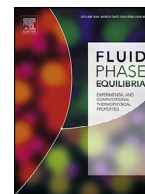
[www.tue.nl/taverne](http://www.tue.nl/taverne)

**Take down policy**

If you believe that this document breaches copyright please contact us at:

[openaccess@tue.nl](mailto:openaccess@tue.nl)

providing details and we will investigate your claim.



# Dispersion activity coefficient models. Part 3: A topology preserving group contribution model



Gerard J.P. Krooshof<sup>a,\*</sup>, Remco Tuinier<sup>b</sup>, Gijsbertus de With<sup>b</sup>

<sup>a</sup> DSM Applied Science Center, P.O.Box 1066, Geleen 6160 BB, the Netherlands

<sup>b</sup> Laboratory of Physical Chemistry, Department of Chemical Engineering and Chemistry, Eindhoven University of Technology, P.O. Box 513, Eindhoven 5600 MB, the Netherlands

## ARTICLE INFO

### Article history:

Received 28 March 2021

Revised 12 May 2021

Accepted 27 May 2021

Available online 9 June 2021

### Keywords:

Group contribution activity coefficient model

Dispersion

Topology theory

Occupation number

## ABSTRACT

The previously developed dispersion activity coefficient model (ACM) is extended from a molecular to a group contribution (GC) model providing results of equal accuracy as the molecular model. Contrary to other GC-ACMs the presented GC-ACM preserves the topology information of each molecule in the mixture. Therefore, in contrast to DISQUAC, UNIFAC or COSMOSAC, where the topology is eliminated by the assignment of groups or the construction of sigma-profiles, this new methodology makes it possible to quantify differences in phase behavior between isomers in a multicomponent mixture.

© 2021 Elsevier B.V. All rights reserved.

## 1. Introduction

In previous work [1] we derived a molecular dispersion activity model from the PC-SAFT equation of state, developed by Gross and Sadowski [2]. In combination with an off-lattice combinatorial contribution [3] this model yields the total activity coefficient of hydrocarbons in their mixtures. This model, which we called the improved perturbed chain (IPC) activity model and accounts explicitly for dispersion, can be extended to systems containing polar and hydrogen bond forming molecules by using the energetic contributions derived from the original UNIFAC [4], COSMO-RS [5] or COSMOSAC [6] model, where dispersion interaction has been neglected. It is not applicable to the modified UNIFAC (Dortmund) model [7], since the combinatorial term in this model has been modified on an ad-hoc basis in order to account for the neglected dispersive interaction between saturated alkyl groups. The benefit of the IPC activity model is that for the dispersive activity coefficient only the topology of the molecule is required.

In this paper we extend the IPC activity model to a group contribution (GC) method, in which the topology information of the molecule is preserved. This implies that this GC model will discriminate between isomers. A feature that is lost in the UNIFAC, DISQUAC [8,9] and COSMO-RS type of models. A further improvement in comparison to the molecular model is the use of the ex-

act solution of the generalized quasi-chemical theory of Guggenheim as is applied in the COSMO-RS family of models. This is an improvement over the DISQUAC, UNIQUAC and UNIFAC models, where the first order approximation of the quasi-chemical theory as derived by Barker [10] has been implemented.

This paper is arranged as follows. First, we summarize the molecular IPC activity coefficient model and point out its key features. Subsequently, we derive the proposed group contribution method. The obtained model (GC-IPC) is compared to experimental data, the molecular IPC model, the original UNIFAC and the modified UNIFAC (Dortmund) model. Finally, we discuss the obtained results.

## 2. Theory

### 2.1. The combinatorial activity coefficient model

The combinatorial contribution to the activity coefficient can be calculated with the generalized Guggenheim equation [3]

$$\ln(\gamma_k) = \ln\left(\frac{\phi_k}{x_k}\right) + \left(1 - \frac{\phi_k}{x_k}\right) \left[\frac{\ln\left(\frac{\phi_k}{\theta_k}\right)}{\frac{\phi_k}{\theta_k} - 1}\right], \quad (1)$$

where  $x_k$ ,  $\theta_k$  and  $\phi_k$  are the mole, surface area and volume fraction of compound  $k$  in the mixture, respectively. The volume fraction and surface area are calculated from the mole fraction, the number of spheres  $m_k$  and the diameter  $d_k$  of the spheres of molecule  $k$ .

\* Corresponding author.

E-mail address: [gerard.krooshof@dsm.com](mailto:gerard.krooshof@dsm.com) (G.J.P. Krooshof).

**Symbols**

$\gamma_j$	activity coefficient of a component in a mixture []
$\hat{\gamma}_j$	local activity coefficient of the methyl group of type $j$ in a pure and mixed state []
$\epsilon_0, \epsilon_1$	energy parameters of IPC activity model [K]
$\epsilon_{\text{CHx}}$	energy increment of CHx group [K]
$\epsilon_{ij}$	interaction energy between segments $i$ and $j$ [K]
$\Delta\epsilon_{ij}$	relative interaction energy between segments $i$ and $j$ [K]
$\tilde{\epsilon}_{ij}$	reduced energy []
$\phi_j$	volume fractions []
$\psi_j$	group fractions []
$\theta_j$	surface area fractions []
$\tau_{ij}$	interaction parameter []
$d_j$	diameter of sphere $j$ in a molecule [Å]
$g$	number of carbon atoms bonded to a methyl group of type $j$ []
$h_j^E$	excess enthalpy of compound $j$ [J mol <sup>-1</sup> ]
$h_{\text{mix}}^E$	enthalpy of mixing [J mol <sup>-1</sup> ]
$k_B$	Boltzmann constant $1.380649 \times 10^{-23}$ [J mol <sup>-1</sup> ]
$m_k$	number of repeating units in compound $k$ []
$q$	relative area in the UNIFAC models []
$r$	relative volume in the UNIFAC models []
$x_j$	mole fraction of compound $j$ in mixture []
$z$	coordination number in the UNIFAC models
$D_k$	occupation number of molecule $k$ (Eq. (4)) []
$D(g)$	occupation number of methyl group of type $g$ []
$J_{\text{QH}}$	quadratic hydrogen number []
$N_C$	carbon (alkane) number []
$T$	system temperature [K]
$T_0$	reference temperature 298.15 [K]
$R$	universal gas constant: 8.3143 [J mol <sup>-1</sup> K <sup>-1</sup> ]
$Z(g)$	number of local interacting spheres of methyl group $g$ []
$Z_k$	average number of interacting spheres of molecule $k$ []
$Z_{\text{max}}$	maximum number of interacting spheres around a sphere []
$Z_{\text{tot}}$	total number of interacting spheres around molecule $k$ []

The equations are

$$\phi_k = \frac{x_k m_k d_k^3}{\sum_j x_j m_j d_j^3}$$

$$\theta_k = \frac{x_k m_k d_k^2}{\sum_j x_j m_j d_j^2} \quad (2)$$

In case all molecules are composed of spheres of the same diameter,  $\phi_k = \theta_k$  and the term within the square brackets in Eq. 1 becomes unity. In that case the generalized Guggenheim combinatorial coefficient model reduces to the Flory-Huggins combinatorial activity coefficient model [11,12].

## 2.2. The molecular dispersion activity coefficient model

The dispersion contribution to the activity coefficient model can be calculated by the improved perturbed chain (IPC) activity coefficient model [1]. This model gives for component  $k$ , which consists of  $m_k$  spheres having diameter  $d_k$ , the activity coefficient:

$$\ln \gamma_k^{\text{disp}} = m_k \frac{Z_k}{2} \tilde{\epsilon}_k \left( \frac{T_0}{T} \right) \left[ 1 - \sum_j \left[ 1 + \frac{Z_j}{Z_k} \left( \frac{d_k}{d_j} \right)^3 \right] \phi_j \frac{\tilde{\epsilon}_{jk}}{\tilde{\epsilon}_k} \right]$$

$$+ \sum_{i,j} \frac{Z_i}{Z_k} \phi_i \phi_j \frac{\tilde{\epsilon}_{ij}}{\tilde{\epsilon}_k} \quad (3)$$

The quantity  $Z_k$  defines the average number of interacting spheres per molecule segment

$$Z_k = Z_{\text{max}} - \frac{D_k}{m_k} \quad (4)$$

where  $Z_{\text{max}} = 14.4$  is the maximum number of spheres around a single sphere at random close-packed (RCP) density, which is based on the work of Klumov et al. [13]. As detailed before [1], the number of spheres within the first and second shell of interaction, is 9.3 and 5.1, respectively. For a spherical molecule  $D_k = 0$ , and thus  $Z_k = Z_{\text{max}}$ . For non-spherical molecules, however,  $D_k > 0$ , resulting in  $0 \leq Z_k < 14.4$ . We mention that  $Z_{\text{max}}$  could be made temperature dependent to reflect the change in interaction energy as a result of the change in the packing density. We will not explore this particular feature here. The index  $D_k$ , which we will call the occupation number of molecule  $k$ , is determined by the topology of the molecule and is therefore temperature independent. The occupation number of a molecule is equal to the first Zagreb number [14] for linear and branched alkanes. The formula for the Zagreb index, and thus  $D_k$ , is given by

$$D_k = \sum_{j=1}^4 m_{v,j} j^2 \quad (5)$$

where  $m_{v,j}$  denotes the number of vertices, in our case non-hydrogen atoms, that have  $j$  covalent bonds with other non-hydrogen atoms. For instance, 3-methyl-pentane has 3 vertices ( $-\text{CH}_3$ ) with 1 C-C bond, 2 vertices ( $>\text{CH}_2$ ) with 2 C-C bonds, and 1 vertex ( $>\text{CH}-$ ) with 3 C-C bonds. In this case Eq. (5) reads:  $D_k = 3 \times 1^2 + 2 \times 2^2 + 1 \times 3^2 = 20$ . Thus, Eq. (4) gives the average number of segments that interact with a segment of 3-methyl-pentane at random close-packed condition as  $Z_k = 14.4 - 20/6 = 11.1$ .

The quantity  $\tilde{\epsilon}_k$  in Eq. (3) is the reduced interaction energy of pure compound  $k$  at reference temperature  $T_0$ , and is given by

$$\tilde{\epsilon}_k = \frac{\epsilon_k}{k_B T_0} \quad (6)$$

where  $k_B$  is the Boltzmann constant, and  $\epsilon_k$  the dispersion energy of molecule  $k$ , which is calculated using the empirical equation:

$$\epsilon_k = \frac{\epsilon_0 + \epsilon_1 J_{\text{QH}}}{m_k} \quad (7)$$

where  $\epsilon_0$  and  $\epsilon_1$  are treated as model constants, which were derived by fitting a large set of experimental data [1]. The calculation of the unlike interaction energy in Eq. (3) requires a mixing rule, which has been discussed in detail elsewhere [15]. For a model set-up with classical mixing rules the constants of Eq. (7) are  $\epsilon_0 = 125.24$  and  $\epsilon_1 = 12.69$ . The parameter  $J_{\text{QH}}$  is a constitutional index, called the quadratic hydrogen number, and is defined by the number of hydrogen atoms attached to carbons in molecule  $k$  according to

$$J_{\text{QH}} = 3^2 N_{\text{CH}_3} + 2^2 N_{\text{CH}_2} + 1^2 N_{\text{CH}} + 0^2 N_C \quad (8)$$

In this expression the parameter  $N_{\text{CH}_3}$ ,  $N_{\text{CH}_2}$ ,  $N_{\text{CH}}$ ,  $N_C$  is the number of carbons that have 3, 2, 1 and 0 hydrogen atoms, respectively. For instance, the compound 2,2,4-trimethylpentane, which consists of 5  $-\text{CH}_3$ , 1  $>\text{CH}_2$ , 1  $-\text{CH}<$  and 1  $>\text{C}<$  group. This results in  $D_k = 5 + 4 + 9 + 16 = 34$  and  $J_{\text{QH}} = 9 \times 5 + 4 \times 1 + 1 \times 1 = 50$ .

## 2.3. Extension to a group contribution method

The main feature of the molecular IPC model is the parameter  $Z_k$ , which contains topological information of the molecule.

**Table 1**

$D(g)$  and  $Z(g)$  of n-octane. Summation gives the molecular occupation number the total number of spheres that interact with the molecule.

Position	1	2	3	4	5	6	7	8	Compound
Group	CH <sub>3</sub>	CH <sub>2</sub>	CH <sub>2</sub>	CH <sub>2</sub>	CH <sub>2</sub>	CH <sub>2</sub>	CH <sub>2</sub>	CH <sub>3</sub>	Sum
$g$	1	2	2	2	2	2	2	1	
$D(g)$	2	3	4	4	4	4	3	2	$D_k = 26$
$Z(g) = Z_{\max} - D(g)$	12.4	11.4	10.4	10.4	10.4	10.4	11.4	12.4	$Z_{\text{tot}} = 89.2$

**Table 2**

$D(g)$  and  $Z(g)$  of 2,4,4-trimethylpentane. Summation gives the molecular occupation number and the total number of spheres that interact with the molecule.

Position	1	2	3	4	5	6	7	8	Compound
Group	CH <sub>3</sub>	CH <sub>3</sub>	CH	CH <sub>2</sub>	C	CH <sub>3</sub>	CH <sub>3</sub>	CH <sub>3</sub>	Sum
$g$	1	1	3	2	4	1	1	1	
$D(g)$	3	3	4	7	5	4	4	4	$D_k = 34$
$Z(g) = Z_{\max} - D(g)$	11.4	11.4	10.4	7.4	9.4	10.4	10.4	10.4	$Z_{\text{tot}} = 81.2$

One of the weak points in the PC-SAFT and thereby the molecular IPC activity coefficient model is the calculation of the average interaction energy using molar fractions. This implies that the liquid is regarded as a mixture of randomly distributed molecules. As has been pointed out by Guggenheim [16], molecular interaction creates non-randomness. His quasi-chemical approach, in which molecules form a quasi-bond at contact, leads to explicit equations for the chemical potential of compounds, of which molecules have one or two types of contacts. In COSMOSPACE [17] equivalent equations to those of the Guggenheim quasi-chemical model are obtained, but also formulae that can handle molecules having more than two different contact types. In this work there are four contact types (CH<sub>3</sub>-, -CH<sub>2</sub>-, >CH-, >C<). In the next section, where we discuss the set-up of a group contribution method for the GC-IPC model, we start with the definition of molecular groups. Subsequently, we explain the calculation of the activity coefficients, where we introduce a method to calculate the interaction energy between alkyl groups.

### 2.3.1. Definition of group increments

From the molecular IPC model we learned that a key parameter is the occupation number  $D_k$ , as defined by Eq. (5). From this the average number of interacting segments around a segment of molecule can be calculated with Eq. (4). To transpose this into a group contribution model, we define main groups, which are characterized by the same dispersion energy, and subgroups, which are defined by the connectivity of the groups they are bonded to. The molecular topology index  $D_k$  of molecule  $k$  is now subdivided into groups, i.e. carbon atoms  $D(g)$ , where  $g$  denotes the number of C-C bonds. The general expression for the occupation number of a group is:

$$D(g) = \sum_{i=1}^g \delta(i), \quad (9)$$

where  $\delta(i)$  denotes the connectivity index of the methyl group that is connected to the segment of interest, which is 1 for CH<sub>3</sub>, 2 for -CH<sub>2</sub>, 3 for -CH< and 4 for >C<. The number of interacting spheres for a group is given by

$$Z(g) = Z_{\max} - D(g). \quad (10)$$

The group contribution  $D(g)$  requires the sum of connectivity indices of a single, a pair, a triplet and a quadruplet of carbon atoms for a CH<sub>3</sub>, a -CH<sub>2</sub>, a -CH< or a >C< group, respectively. For instance, the  $D(g)$  of -CH<sub>3</sub> group is denoted as  $D(1)$  and depends only on one connection. If the -CH<sub>3</sub> group is connected to -CH<sub>3</sub>, >CH<sub>2</sub>, -CH, or >C<,  $D(1) = \delta(1) = 1, 2, 3$  or 4, respectively. This yields  $Z(1) = 13.4, 12.4, 11.4$  and 10.4, respectively. Likewise,

**Table 3**

Dispersion profile of n-octane and its isomer 2,4,4-trimethylpentane. Between parentheses the molecular fraction,  $\psi_j$ , of main group  $j$ .

Main group	n-Octane	2,4,4-trimethylpentane
$Z_{k3} = \sum Z(1)$	24.8 (0.278)	54.0 (0.665)
$Z_{k2} = \sum Z(2)$	64.4 (0.722)	7.4 (0.091)
$Z_{k1} = \sum Z(3)$	0 (0.000)	10.4 (0.128)
$Z_{k0} = \sum Z(4)$	0 (0.000)	9.4 (0.116)
$Z_{\text{tot}}$	89.2	81.2
$Z_k$	11.15	10.15

a -CH< group, which is connected to the triplet (-CH<sub>3</sub>, -CH<sub>3</sub>, -CH<sub>2</sub>-), has  $D(3) = \delta(1) + \delta(2) + \delta(3) = 1 + 1 + 2 = 4$  and, therefore,  $Z(3) = 10.4$

In the Supplementary Information a full table of 68 local occupation numbers, the  $\delta(i)$  increments, for branched alkanes are given. Tables 1 and 2 show the local occupation number  $D(g)$  and the local number of interacting spheres  $Z(g)$  of the groups in n-octane and 2,4,4-trimethylpentane, respectively.

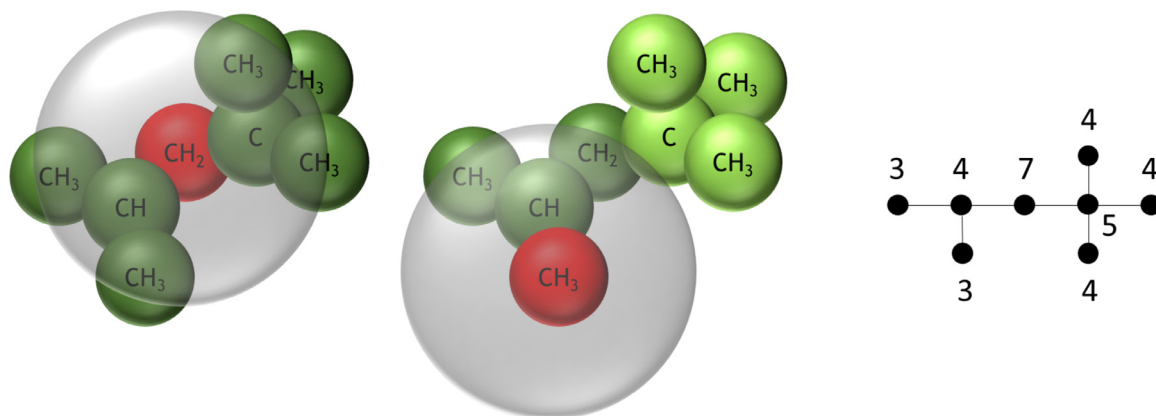
The last column of Tables 1 and 2 shows that the sum of group increments  $D(g)$  gives the same value as obtained from Eq. (5). The quantity  $Z_{\text{tot}}$  is the sum of  $Z(g)$ . The computation of  $D(g)$  values for 2,4,4-trimethylpentane is also illustrated by the two 3D representations and a graph in Fig. 1.

### 2.3.2. Definition of interaction energy

The subgroups, which belong to the same main group, i.e. having the same  $g$  value, are assumed to have the same contribution to the dispersion energy. This implies that we neglect the intramolecular interaction between neighboring groups in the same molecule, because we assume that these proximity effects cancel out in the calculation of the chemical potential difference of the compound in the pure and mixed state. The sum of  $Z(g)$  values for one particular  $g$  value quantifies the total amount of interactions of one type of methyl groups. Table 3 shows that n-octane has two CH<sub>3</sub> groups and the six CH<sub>2</sub> groups, interact with 24.8 and 64.4 groups from the surrounding, respectively. The isomer 2,4,4-trimethylpentane has five CH<sub>3</sub> groups, one CH<sub>2</sub> group, one CH and two C groups, which interact with 54.0, 7.4, 10.4 and 9.4 groups from the surrounding, respectively.

The total number of interactions is used to calculate the fractions of types of dispersion interaction using

$$\psi_j = \frac{\sum_k x_k Z_{k,j}}{\sum_{k,j} x_k Z_{k,j}}, \quad (11)$$



**Fig. 1.** Computation of groups in 2,4,4-trimethylpentane. Left and middle picture emphasize a methyl group (red) with its close connections (dark green): -CH<sub>2</sub>- with  $D(2) = 7$  and CH<sub>3</sub> with  $D(1) = 3$ , respectively. Right: Graph with all group increments  $D(g)$ . (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

where  $Z_{k,j}$  is the total number of interactions with the alkyl groups of type  $j$  in molecule  $k$ . Here the subscripts  $j = 3, 2, 1$  and  $0$  are short-cut notations for the interaction with the central CH<sub>3</sub>, CH<sub>2</sub>, CH, and C group, respectively. In Table 3 the fractions  $\psi_j$  of the main groups in a molecule are given between the parentheses. We emphasize that the definition of the parameter  $\psi_j$  contains topology information, which means that it takes into account the proximity effect of intramolecular interactions, whereas in the COSMOSPACE and COSMO-RS models interactions are computed on the basis of pairwise interacting surface areas. In the latter model a molecule is divided into small surface parts considered to have a constant charge density and the local electrostatic interaction energy is calculated from the charge density between two surface patches. This idea is thus translated into the  $Z_j$  of our GC-ACM, by which the fraction of induced dipole interactions  $\psi_j$  can be quantified. Different from the COSMOSPACE and COSMO-RS model we are not dealing with a pairwise interaction, but with the interaction around a central group, for which the number of groups in the first and second interaction shell are taken into account. This difference implies that the GC-IPC model takes intramolecular interactions (steric hindering) into account.

The local activity coefficient of a methylene group in a mixture,  $\hat{\gamma}_j$ , is found by using the methodology for the residual part of the COSMOSPACE model [17], where the surface fractions have been replaced with the fraction definition of Eq. (11). Worthwhile to mention is that the methodology of the COSMOSPACE model is the exact solution of the generalized Guggenheim quasi-chemical model [18]. This is one of the key differences with the group contribution methods DISQUAC [8] and UNIFAC [4], where the first order approximation of the quasi-chemical theory, which was derived by Barker [10] first, is applied to calculate the residual energy. For more details about the difference between the exact and first order solution of the quasi-chemical theory of Guggenheim, we refer to the COSMOSPACE paper [17]. Consequently, the local activity coefficients of the 4 types of methyl groups are obtained by simultaneously solving the set of 4 equation, defined by

$$\frac{1}{\hat{\gamma}_j} = \sum_{l=0}^3 \psi_l \hat{\gamma}_l \tau_{lj}, \quad (12)$$

where  $j$  and  $l$  run from 0 to 3 to denote the different alkyl groups. The parameter  $\tau_{ij}$  is the coefficient of interaction between two alkyl groups and is defined by the interaction energy difference. We have

$$\tau_{ij} = \exp \left[ -\frac{\epsilon_{ij} - (\epsilon_{ii} + \epsilon_{jj})/2}{k_B T} \right] = \exp \left[ -\frac{\Delta \epsilon_{ij}}{k_B T} \right], \quad (13)$$

where  $\epsilon_{ij}$  is the dispersion energy between an alkyl pair, as is computed from the dispersion energy of the various groups. In the appendix we show that for square-well potentials Eq. (13) reduces to

$$\tau_{ij} = \exp \left[ -\frac{(i-j)^2 \epsilon_{HH}}{2k_B T} \right], \quad (14)$$

where  $i$  and  $j$  denote the number of hydrogen atoms in the segments that interact, and  $\epsilon_{HH}$  is the induced dipole interaction energy between two hydrogen atoms of different alkyl groups. This implies that the energy differences of unlike pairs are correlated by the ratios:

$$\Delta \epsilon_{1,2} : \Delta \epsilon_{2,3} : \Delta \epsilon_{3,4} : \Delta \epsilon_{1,3} : \Delta \epsilon_{2,4} : \Delta \epsilon_{1,4} = 1 : 1 : 1 : 4 : 4 : 9 \quad (15)$$

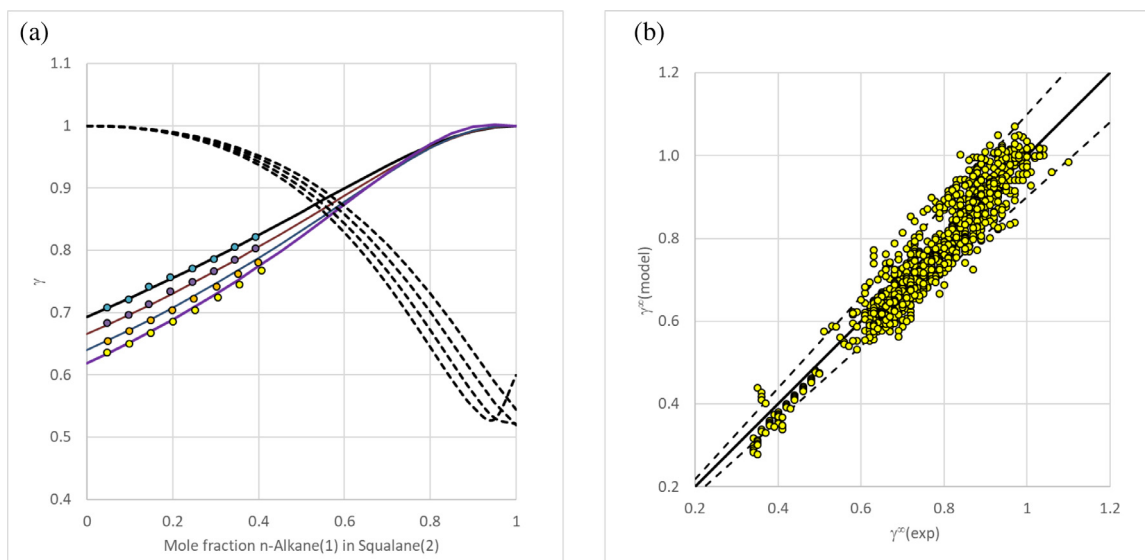
However, these ratios are not valid for a distance dependent interaction potential, as illustrated in the appendix for the London dispersion model [19]. Taking further into account that steric hindering influences the available space for methyl groups in the first and second shell of interaction, we regard the values in Eq. (15) as indicative. For this work it implies that we consider the  $\Delta \epsilon_{ij}$  of Eq. (13) as model parameters, which will be determined by minimization of the sum of squared differences between experimental and calculated infinite dilution activity coefficient data.

In previous work [1], we demonstrated that topological information improves the description of non-ideal behavior of alkane mixtures. Using nearest-neighbor theory [3] we showed that the molecular activity coefficient of compound  $k$  is calculated from ratio in surface activity coefficients of that compound in the mixed and pure state, and that the logarithm of this ratio requires a factor that is given by half the number of nearest neighbors  $Z_k$ . This led to the following expression for the dispersive activity coefficient of compound  $k$  in a molecular mixture:

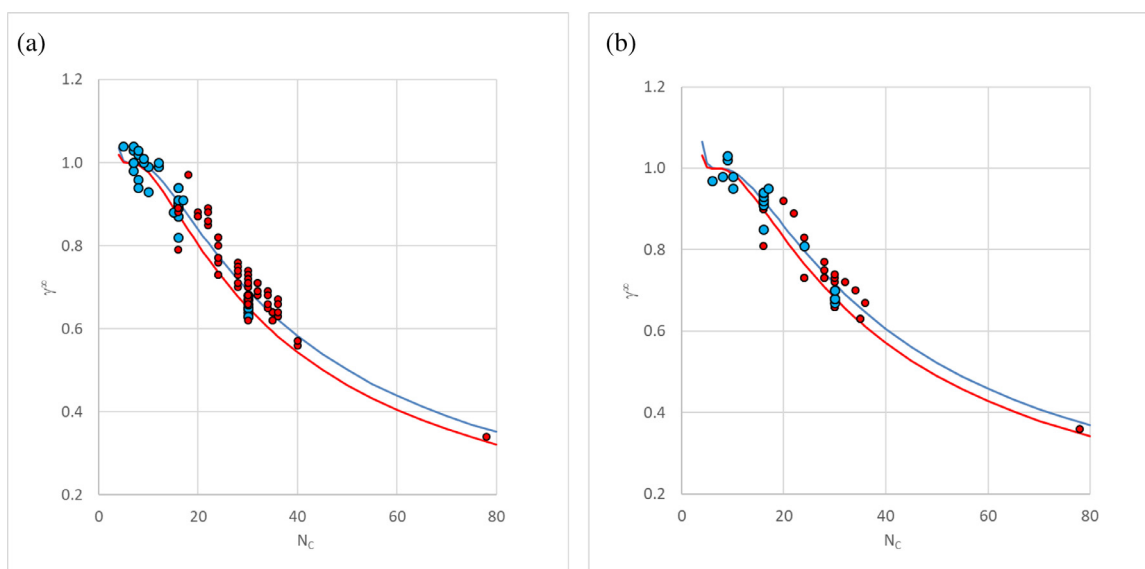
$$\ln(\gamma_k^{\text{disp}}) = \sum_j \frac{Z_{kj}}{2} \ln \left( \frac{\hat{\gamma}_j^x}{\hat{\gamma}_j^p} \right), \quad (16)$$

where  $\hat{\gamma}_j^x$  and  $\hat{\gamma}_j^p$  denote the local activity coefficient of group  $j$  in the mixture and in the pure compound, respectively, as calculated using Eq. (12).

We mention that in the molecular IPC model we have the term  $m_k Z_k$ , which in Eq. (16) corresponds with the summation of the group values  $Z_{kj}$ . The reduced energy term  $\tilde{\epsilon}_k$  of the molecular IPC model is now enclosed in the local activity coefficients of Eq. (12) by means of parameter  $\tau_{ij}$ . In the next section the dispersion energy between two hydrogen atoms of different segments  $\Delta \epsilon_{ij}$ , will be obtained by fitting the limiting activity coefficients of linear and branched alkanes.



**Fig. 2.** Left panel: Parity plot of IDAC of binary alkane mixtures between experimental data and the proposed GC model, where the dashed lines indicates the 10% experimental error in IDAC values and the solid line  $y = x$ . Yellow dots are experimental IDAC data from DDBST [21] used for fitting of the binary energy parameters of Table 5. Right panel: Experimental activity coefficients of n-pentane (yellow), n-hexane (orange), heptane (purple), n-octane (blue) in squalane at 303 K from Ashworth [22]. Solid and dashed curves represent model result for alkanes in squalane and squalane dissolved in octane (top), heptane, hexane and pentane (bottom), respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



**Fig. 3.** Limiting activity coefficients of hexane (left panel) and heptane (right panel) in alkanes. Blue and red symbols are experimental data at low and high temperature collected from DDBST [21]. Red and blue curves are the results of GC-IPC model at 313 and 373 K, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

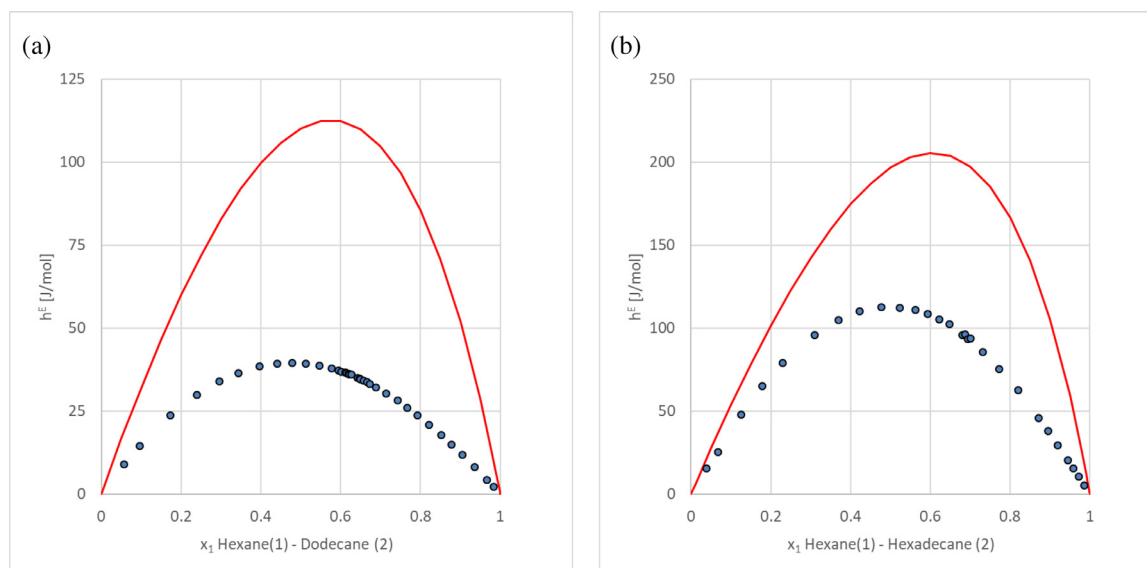
In this work, we have not included cyclic alkanes, because this needs more research. Prior to implementation we have to verify whether the constraint of random closed packing which lead to  $Z_{\max}$  is still valid. Extensive computational research, as was done by Klumov [13], has to be available to verify this. Since cyclo-alkanes have a significant higher normal boiling point than n-alkanes [20], an additional dispersive interaction on top of the interaction defined by the alkyl groups is probably required. The introduction of cyclic groups in the modified UNIFAC (Dortmund), which have different values than the linear groups, support this hypothesis. We speculate that this extra dispersive interaction is caused by a synchronized movement of the hydrogen electrons around the perimeter of the cyclic entity. This hypothesis needs a quantum mechanical proof before either the existing alkyl groups

are modified for cyclic entities and is added as a new set of groups as was done in the modified UNIFAC (Dortmund) model or a single new group, that reflect this collective interaction, is introduced.

### 3. Results

#### 3.1. Activity coefficients

Fig. 2 and Fig. 3 depict the results for the infinite dilution activity coefficient (IDAC) obtained with the GC-IPC model, which is the combination of Eqs. (16) and (1). We observe a good correlation between experiment and model for the limiting activity coefficients of binary hydrocarbon mixtures and for the activity coefficients of pentane to octane in squalane.



**Fig. 4.** Enthalpy of mixing of the binary systems n-hexane - n-dodecane (left panel) and of n-hexane - n-hexadecane (right panel). Experimental data measured at  $T = 298$  K from [26] and [27]. Curves are the GC-IPC model at  $T = 298$  K.

**Table 4**

Comparison of the molecular IPC models with different mixing rules and the GC-IPC model. The statistical results of the original and the modified UNIFAC (Dortmund) models were taken from [7]. (Np = No. of parameters, Nd = No. of datapoints).

	Mixing rule	Molecular IPC [1]			GC IPC	UNIFAC	UNIFAC(Do)
		IC	WH	Kong	None	None	None
Statistical info	Np → Nd ↓	2	2	2	6	Np = 2 Nd = 250	Np = 6 Nd = 250
AAD% IDAC [21]	1188	4.7%	4.9%	3.7%	4.4%	25%	3.3%
AAD% C5/8 in squalane [22]	32	0.004%	0.004%	0.007%	0.03%	–	–
$\chi_{red}^2$ with $\sigma = 5\%$	1210	1.55	1.61	1.21	1.46	–	–

**Table 5**

Absolute (top) and relative (bottom) dispersion interaction energies.

$\Delta\epsilon_{ij}$ [K]	-CH <sub>3</sub>	-CH <sub>2</sub> -	-CH<	>C<
-CH <sub>3</sub>	0	42	177	294
-CH <sub>2</sub> -	42	0	36	71
-CH<	177	36	0	0
>C<	294	71	0	0
$\Delta\epsilon_{ij}/\Delta\epsilon_{12}$	-CH <sub>3</sub>	-CH <sub>2</sub> -	-CH<	>C<
-CH <sub>3</sub>	0	1	4.24	7.07
-CH <sub>2</sub> -	1	0	0.87	1.69
-CH<	4.24	0.87	0	0
>C<	7.07	1.69	0	0

The reduced Chi-parameter in Table 4, which is based on an experimental error of 5%, shows that the GC-IPC model is of equal accuracy as the molecular IPC model, where we tested three different mixing rules: the Wong-Hagler (WH) [23], the Kong [24] and the ideal classical (IC) mixing [25], where  $k_{ij} = 0$ .

For comparison the reported statistical results of the original UNIFAC [4] and the modified UNIFAC (Dortmund) [7] models are included. Note that the number of parameters in the original UNIFAC model to describe alkanes is 2, which refers to the volume and area of the reference molecule needed to calculate the relative area and volume of molecule groups. The modified UNIFAC (Dortmund) model has 6 parameters: the exponent 3/4, the choice to set all the relative volumes of the alkyl groups to  $r = 0.6325$ , and the 3

optimized relative areas of -CH<sub>3</sub>, -CH<sub>2</sub>, and -CH<, while keeping that of >C< to zero (Fig. 3).

Table 5 shows at the left side the absolute interaction energies,  $\Delta\epsilon_{ij}$ , and on the right the values relative to the interaction energy between a CH<sub>3</sub> group and a CH<sub>2</sub> group.

We observe that the regression gives for the relative interaction of CH<sub>3</sub> with the other alkyl groups the ratios of 1:4.24:7.07. It follows roughly the ratio 1:4:9 of Eq. (15), as expected for the square-well potential. In the appendix we demonstrate that the relative ratio's are nearly the same for the London dispersion potential.

### 3.2. Excess enthalpy

The prediction of the excess enthalpy is shown in Fig. 4. It is qualitatively correct, and comparable to the results of our earlier work [1], where the classical combining rules of Lorentz and Berthelot were used.

This systematic deviation can be improved by applying a temperature dependence for the interaction energy or scaling the parameter  $Z_{max}$  with the liquid density of the mixture. We mention that in the modified UNIFAC (Dortmund) model [7] the systematic error in the excess enthalpy has been improved by using for the interactions between different classes of molecules a temperature dependent interaction energy. This implies that the constant  $\epsilon_{HH}$  in Eq. (14) is replaced by  $a_{ij} + b_{ij}T + c_{ij}T^2$ . However, in the modified UNIFAC (Dortmund) model the class of linear and branched alkanes fall within the same class of molecules and, therefore, the interaction energy is zero and similar results as in Fig. 4 will be obtained.

#### 4. Discussion and conclusions

In this work we have shown that the molecular dispersion activity coefficient model of our previous work for the calculation of the dispersion activity coefficient contribution can be extended into a group-contribution method including Boltzmann statistics to account for interaction probability. The proposed GC-ACM gives a very good description for the IDAC and has the same accuracy as the molecular ACM (IPC). Both models are a clear improvement for the alkane interaction description in comparison to the original UNIFAC model. The original UNIFAC gives systematically lower for the IDAC, due to the absence of an explicit dispersion term. The modified UNIFAC (Dortmund) model [7] shows by its AAD% of 3% a better result for the alkanes than our proposed GC model. However, the physical basis of the combinatorial part of the UNIFAC (Dortmund) model is weak due to the applied modifications, which were introduced to capture the dispersive contribution. First, the entropic solvency effect was reduced by using an empirical obtained exponent for the calculation of volume fractions. This exponent was set somewhat arbitrary to the value 3/4. Second, in contrast to the original UNIFAC model the relative volumes,  $r$ , of all the methyl sub groups were set to one value, while the relative areas,  $q$ , of the 4 alkyl groups were optimized. Using this process, the relative area of the CH<sub>3</sub>, CH<sub>2</sub> and CH groups become relatively larger than those of the original UNIFAC model [4]. This adjustment to incorporate the dispersion contribution into the combinatorial term created a fundamental inconsistency in terms of the coordination number  $z$ . If this number would be calculated by the original connectivity equation for linear and branched structures,  $z = 2(r - 1)/(r - q)$ , one would discover that the modified UNIFAC (Dortmund) model yields negative coordination numbers. For instance, the relative volume and area of n-hexane would yield  $z = -4.8$ . In the original UNIFAC model  $z = 10.9$ . In fact, the connectivity relation would yield for an infinite linear chain the correct  $z = 10$  in the original UNIFAC model, but  $z = -16.4$  in the modified UNIFAC (Dortmund) model. Because the coordination number is fixed to 10 and because the residual part compensates the systematic errors arising from the exponent 3/4, the modified UNIFAC (Dortmund) gives reliable and accurate results. On the other hand, the GC-IPC model has nearly the same accuracy but has a profound physical basis.

In the proposed GC-ACM the key parameter is the local occupation number  $D(g)$ , from which the fraction of interacting spheres of a group  $j$  in molecule  $k$ , i.e.  $\psi_{k,j}$ , can be calculated. The energy parameters for the interaction between two methyl groups were obtained by regressing experimental infinite dilution activity coefficient data, and follow roughly a quadratic relation with the difference in hydrogen atoms between two groups. The deviation from a quadratic relation, i.e. Eq. (15) vs. Table 5, can be attributed to the steric hindering of groups, which changes the ratio of molecules in the first and second shell of interaction. This change becomes effective when instead of a square-well potential, a London dipole-dipole interaction potential [19] is applied (see appendix). The reason for the larger deviation of the interaction energy between the >CH- and >C< groups is not fully understood, but might be explained by the observation made by Schreiner et al. [28]. They showed that the crowding of carbons in tertiary butyl side groups, stretches the adjacent C-C bond, but that this does not lead to a lower stability of this covalent bond, because the crowded side groups give a compensating attractive dispersion interaction between the intramolecular H..H surfaces that stabilizes the stretched bond. This might imply that induced dipole interaction between the >CH- and >C< methyl groups are different than expected, due to the occurrence of these stretched bonds. We also note that we neglected the fact that carbons have an ellipsoidal shape [29]. The ellipsoidal shape changes the ratio in area and surface fraction of

compounds, and thereby the value for the combinatorial activity contribution. However, this gives at most a 5% change in the activity coefficient, as was discussed in [3].

We mention further that the proposed GC-ACM can be extended to other atom groups that interact fully or partly by dispersion. In the later case additional activity coefficient contributions, such as a permanent polar or an associative interaction contribution, need to be included. For instance, the dispersive part of a hydroxyl group in a primary, secondary and tertiary alcohol will have  $D(1) = 1$  for methanol, while it is 2 and 3 for the other alcohols, respectively. This yields  $Z_{OH} = 13.4$ , 12.4, and 11.4 respectively, and this can be used to quantify a dispersive interaction.

Finally, we like to remark that the model can be used for vapor-liquid equilibria description at elevated pressures by introducing a fugacity correction for the gas phase and that the excess enthalpy description can be improved quantitatively by making  $Z_{max}$  dependent on the liquid density. This is, however, out of the present scope.

#### 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.

#### Appendix A. Dispersion interaction model

Here we demonstrate why the ratios of the interaction energies in matrix (Table 5) differ from those in Eq. (15). First we start with the derivation of Eq. (15), subsequently we use the London dispersion model to explain the observed differences.

The energy difference between two different atoms in reference to two equal atoms can be defined by

$$\Delta\epsilon_{ij} = \epsilon_{ij} - (\epsilon_{ii} + \epsilon_{jj})/2 \quad (\text{A.1})$$

where  $\epsilon_{ij}$  is the dispersion energy between atom  $i$  and  $j$ . Consider a CH<sub>3</sub> end group interacting by dispersion with a CH<sub>2</sub> chain group. We assume that each atom can form an induced dipole, so that the two atom groups result in 12 induced dipole pairs, as indicated in Table A.6.

The interaction between a CH <sub>$m$</sub>  and a CH <sub>$n$</sub>  group involves one C..C induced dipole pair,  $m + n$  C..H induced dipole pairs and  $m.n$  H..H induced dipole pairs. The corresponding dipole energies are  $\epsilon_{CC}$ ,  $\epsilon_{CH}$ , and  $\epsilon_{HH}$ , respectively. The energy difference is given by:

$$\begin{aligned} \Delta\epsilon_{CH_m, CH_n} &= \epsilon_{CH_m, CH_n} - \frac{1}{2}(\epsilon_{CH_m, CH_m} + \epsilon_{CH_n, CH_n}) \\ &= [\epsilon_{CC} + (m+n)\epsilon_{CH} + mn\epsilon_{HH}] \\ &\quad - \frac{1}{2}[(\epsilon_{CC} + 2m\epsilon_{CH} + m^2\epsilon_{HH}) + (\epsilon_{CC} + 2n\epsilon_{CH} + n^2\epsilon_{HH})] \\ &= -\frac{1}{2}(m-n)^2\epsilon_{HH}, \end{aligned} \quad (\text{A.2})$$

where  $\epsilon_{HH}$ ,  $\epsilon_{CC}$  is the dispersion energy between two hydrogen and two carbon atoms, respectively and  $\epsilon_{CH}$  is the dispersion energy between a hydrogen and a carbon atom.

Eq. (A.2) shows that only the hydrogen-hydrogen induced dipoles matter. The relative energies in Table A.7 illustrate this.

**Table A.6**

Example of CH<sub>3</sub> CH<sub>2</sub> induced dipole interaction energies.

	C	H	H
C	$\epsilon_{CC}$	$\epsilon_{CH}$	$\epsilon_{CH}$
H	$\epsilon_{CH}$	$\epsilon_{HH}$	$\epsilon_{HH}$
H	$\epsilon_{CH}$	$\epsilon_{HH}$	$\epsilon_{HH}$
H	$\epsilon_{CH}$	$\epsilon_{HH}$	$\epsilon_{HH}$



**Table A.7**  
Relative interaction energies between two methyl groups.

	$-2\Delta\epsilon_{ij}/\epsilon_{HH}$	$j=\text{CH}_3$	$j=\text{CH}_2$	$j=\text{CH}$	$j=\text{C}$
$i=\text{CH}_3$	0	1	4	9	
$i=\text{CH}_2$	1	0	1	4	
$i=\text{CH}$	4	1	0	1	
$i=\text{C}$	9	4	1	0	

**Table A.8**  
Correction factors for the induced energy due to steric hindering.

Methyl group	Spheres in 1st shell	Spheres in 2nd shell	Fraction 1st shell	Fraction 2nd shell	Correction factor	Relative correction
$\text{CH}_3$	8.3	4.1	0.67	0.33	0.711	100%
$\text{CH}_2$	7.2	3.1	0.70	0.30	0.737	104%
$\text{CH}$	6.3	2.1	0.75	0.25	0.781	110%
$\text{C}$	3.3	3.1	0.52	0.48	0.576	81%

**Table A.9**  
Relative induced energy corrected for steric hindering.

Methyl group	$\text{CH}_3$	$\text{CH}_2$	$\text{CH}$	$\text{C}$
$\text{CH}_3$	0	$\frac{100\% \times 104\%}{100\% \times 104\%}$	$4 \times \frac{100\% \times 111\%}{100\% \times 104\%}$	$9 \times \frac{100\% \times 80\%}{100\% \times 104\%}$
$\text{CH}_2$	1	0	$\frac{104\% \times 110\%}{100\% \times 104\%}$	$4 \times \frac{104\% \times 81\%}{100\% \times 104\%}$
$\text{CH}$	4.24	1.10	0	$\frac{110\% \times 31\%}{100\% \times 104\%}$
$\text{C}$	7.04	3.24	0.86	0

We observe that Table A.7 is symmetric around both matrix diagonals, and that the interaction energies increases quadratically with increasing difference in hydrogen number. These ratios hold for interactions which are distance independent, i.e. interactions that can be described with a square-well potential.

This approach is not valid in case the interaction is described by the London dispersion model [19]. The London dispersion energy  $U$  is given by

$$U = -\frac{3}{4} \frac{h\nu_0\alpha^2}{R_0^6}, \quad (\text{A.3})$$

where  $h$  is the Plank constant, while  $\nu_0$ ,  $R_0$  and  $\alpha$  are the characteristic frequency, the distance, and the polarizability of the two identical systems, respectively. Different from the square-well potential we have an interaction energy that drops with the sixth power of the distance. This implies that the total interaction of a central methyl group depends on the number of methyl groups in the first and second shell of interaction as well as the type of group in the center. For instance, the  $\text{CH}_3$ - group, due to the presence of bonded  $-\text{CH}_X-$  groups, has a reduced number of interacting methyl groups. In the first shell the number of interaction spheres reduces from 9.3 to 8.3 [1], and in the second shell from 5.1 to 4.1 or even to 3.1 or 2.1 when the first shell contains a  $-\text{CH}<$  or  $>\text{C}<$  group, respectively. This implies that the fraction over the two shells changes from 0.65 and 0.35 to 0.67 and 0.33. As a result the interaction of a  $\text{CH}_3$  with a methyl group  $\text{CH}_X$  from the surrounding is

$$\epsilon_{\text{CH}_3.\text{CH}_X} = -\frac{3}{4} h\nu_0\alpha^2 \left[ \frac{0.67}{R_0^6} + \frac{0.33}{(\sqrt{2}R_0)^6} \right] = -\frac{3}{4} \frac{h\nu_0\alpha^2}{R_0^6} \times 0.711. \quad (\text{A.4})$$

Here 0.711 is a factor to scale from a square-well to a London potential. The results for the other groups are summarized in Table A.8.

For the  $\text{CH}_3$ , the  $-\text{CH}_2$  and the  $-\text{CH}<$  methyl group, the number of hindering groups is set to 2, 4 and 6, respectively. These are

equally distributed over the first and second shell of interaction. The 8 groups of the  $>\text{C}<$  group were not distributed equally, but optimized to 6 in the first and 2 in the second shell of interaction in order to bring the results of Table A.9 closer in line with the experimental ratio, i.e. Table 5.

The main feature of this matrix is that it is no longer symmetric along the  $\text{C}-\text{CH}_3$   $\text{CH}_3-\text{C}$  line, which indicates that the steric hindering of groups also effects the effective interaction energy. This asymmetry cannot be explained by a distance independent potential, such as the square-well potential, as the above calculations with the London dispersion model illustrates.

## Supplementary material

Supplementary material associated with this article can be found, in the online version, at [10.1016/j.fluid.2021.113097](https://doi.org/10.1016/j.fluid.2021.113097).

## CRedit authorship contribution statement

**Gerard J.P. Krooshof:** Conceptualization, Methodology, Writing - original draft. **Remco Tuinier:** Supervision, Writing - review & editing. **Gijsbertus de With:** Supervision.

## References

- [1] G. Krooshof, R. Tuinier, G. de With, Dispersion activity coefficient models. Part 2: perturbed chain equations of state, Fluid Phase Equilib. 502 (2019) 112286.
- [2] J. Gross, G. Sadowski, Perturbed-chain SAFT: an equation of state based on a perturbation theory for chain molecules., Ind. Eng. Chem. Res. 40 (4) (2001) 1244.
- [3] G. Krooshof, R. Tuinier, G. de With, Generalization of Guggenheim's combinatorial activity coefficient equation, J. Mol. Liq. 266 (2018) 467–471.
- [4] A. Fredenslund, R. Jones, J. Prausnitz, Group-contribution estimation of activity coefficients in nonideal liquid mixtures., AIChE J. 21 (6) (1975) 1086.
- [5] A. Klamt, V. Jonas, T. Burger, J. Lohrenz, Refinement and parametrization of COSMO-RS, J. Chem. Phys. A 102 (1998) 5074–5085.
- [6] S.-T. Lin, S. Sandler, A priori phase equilibrium prediction from a segment contribution solvation model., Ind. Eng. Chem. Res. 41 (5) (2002) 899–913.
- [7] U. Weidlich, J. Gmehling, A modified UNIFAC model. 1. Prediction of VLE,  $h^E$ , and  $\gamma^\infty$ , Ind. Eng. Chem. Res. 26 (7) (1987) 1372–1381.
- [8] H. Kehiaian, J. Grollier, G. Benson, Thermodynamics of organic mixtures. a generalized quasichemical theory in terms of group surface interactions., J. Chim. Phys. 75 (1) (1978) 1031–1048.
- [9] H. Kehiaian, Thermodynamics of binary liquid organic mixtures., Pure Appl. Chem. 57 (1) (1985) 15–30.
- [10] J.-A. Barker, Cooperative orientation effects in solutions., J. Chem. Phys. (1952) 1526–1532.
- [11] P. Flory, Thermodynamics of high polymer solutions, J. Chem. Phys. 10 (1) (1942) 51–61.
- [12] M. Huggins, Solutions of long chain compounds, J. Chem. Phys. 9 (5) (1941) 440.
- [13] B. Klumov, Y. Jin, H. Makse, Structural properties of dense hard sphere packings, J. Phys. Chem. B 118 (36) (2014) 10761–10766, doi:10.1021/jp504537n.
- [14] I. Gutman, B. Ruscic, N. Trinajstić, C. Wilcox, Graph theory and molecular orbitals. XII. Acyclic polyenes, J. Chem. Phys. 62 (9) (1975) 3399–3405, doi:10.1063/1.430994.
- [15] G. Krooshof, R. Tuinier, G. de With, Dispersion activity coefficient models. Part 1: cubic equations of state, Fluid Phase Equilib. 501 (2019) 112275.
- [16] E. Guggenheim, Mixtures, Clarendon Press Oxford, 1952.
- [17] A. Klamt, G. Krooshof, R. Taylor, COSMOSPACE: Alternative to conventional activity-coefficient models, AIChE J. 48 (10) (2002) 2332–2349.
- [18] E. Guggenheim, Statistical thermodynamics of mixtures with zero energies of mixing, Proc. R. Soc. A 183 (1944) 203–212.
- [19] F. London, The general theory of molecular forces, Trans. Faraday Soc. 33 (1937) 8b–26.
- [20] G. Krooshof, Combinatorial and dispersion activity coefficient models for molecular solutions., 2019, ([https://pure.tue.nl/ws/files/140457858/20191217\\_Krooshof.pdf](https://pure.tue.nl/ws/files/140457858/20191217_Krooshof.pdf)). See appendix N of the thesis.
- [21] Dortmund Data Bank GmbH software package, 2018, (<http://www.ddb.st.de>).
- [22] A. Ashworth, Activity coefficients of C5 to C8 hydrocarbons in squalane and dinonyl phthalate at 30 C determined by a vacuum microbalance technique, J. Chem. Soc. Faraday Trans. 1 69 (1973) 459–466.
- [23] D. Wong, S. Sandler, A theoretically correct mixing rule for cubic equations of state, AIChE J. 38 (1992) 671.
- [24] C. Kong, Combining rules for intermolecular potential parameters II. Rules for the Lennard-Jones (12-6) potential and the morse potential., J. Chem. Phys. 59 (1973) 2464–2467.
- [25] H. Lorentz, Über die Anwendung des Satzes vom Virial in der kinetischen Theorie der Gase, Ann. Phys. 248 (1881) 127–136.

- [26] J. Ott, K. Marsh, R. Stokes, Excess enthalpies, excess volumes, and excess Gibbs free energies for (n-hexane + n-dodecane) at 298.15 and 308.15 K., *J. Chem. Thermodyn.* 13 (1981) 371–376.
- [27] R. Miller, A. Williamson, Excess molar enthalpies for (n-hexane + n-hexadecane) and for three- and four-component alkane mixtures simulating this binary mixture, *J. Chem. Thermodyn.* 16 (1984) 793–799.
- [28] P. Schreiner, Overcoming lability of extremely long alkane carbon-carbon bonds through dispersion forces, *Nature* 477 (2011) 308–311.
- [29] S. Batsanov, Van der Waals radii of elements., *Inorg. Mater.* 37 (9) (2001) 871.