PAPER

Check for updates

A modified group contribution method for estimating thermodynamic parameters of methanol-to-olefins over a SAPO-34 catalyst†

Junyi Yu,[ab] Hua Li,*[a] Mao Ye ✦[a] and Zhongmin Liu[ab]

Group contribution (GC) methods, a semi-empirical approach based on the additivity of guest molecular properties, are widely applied to obtain the thermodynamic properties of complex reaction networks. In molecular sieve catalyzed processes, however, the interaction between guest molecules and host active sites also affects thermodynamic properties. In this study, therefore, we propose a modified group contribution (mGC) method by considering the interaction between the groups of guest molecules and independent active site functional groups (IASFGs) in molecular sieves. The mGC method has been used to estimate the thermodynamic properties of guest molecules as well as elementary reactions for the initial stage of methanol to olefins (MTO) reaction over SAPO-34 molecular sieves. It shows that mGC is more accurate than the conventional GC (cGC) methods when compared with the reference data calculated by density functional theory (DFT), indicating that mGC provides an effective way for batch calculation of thermodynamic properties in molecular sieve catalyzed processes.

1. Introduction

Chemical processes are widely encountered in various industrial sectors, and have produced most of, if not all, the synthesized materials necessary for human life. Usually, chemical reactions can be very sophisticated and involve numerous intermediate species and elementary reaction steps, forming quite complex reaction networks with multiple potential reaction pathways. One of the central tasks in developing a chemical process is the on-target production of certain chemical species demanded in the market *via* optimal reaction pathways to save cost and energy. Effective and concise reaction pathways are of great significance for the development of new chemical processes. In this regard, thermodynamic properties are highly desired in the understanding of reaction mechanisms and identification of possible reaction pathways.

Quantitatively deriving the thermodynamic properties of molecular species in a chemical process, usually including enthalpy, entropy, Gibbs free energy, and some others, is quite challenging because a great number of chemical species (*i.e.*, reactants, intermediates, and products) need to be considered.

In principle, this can be done through well-established experiments or fundamental quantum mechanics and molecular modeling. Normally experiments provide the most direct way to obtain accurate and reliable thermodynamic data, which, however, are not always feasible as the number of chemical species and relevant elementary reactions might be huge. Moreover theoretical calculations based on quantum mechanics and molecular modeling are also quite time consuming and laborious, constrained by computational power and time within reach.

Group contribution (GC) methods, which are based on the additivity of molecular properties, represent a general and effective approach to calculate thermodynamic properties.[1] Compared with DFT calculation, GC methods can be considered as a semi-empirical approach, which involves no quantum chemical calculations and can quickly obtain the thermodynamic parameters of certain species. In GC methods, the properties of a given molecule are presumably treated as the sum of elemental contributions of all individual structural groups composing the molecule, and for the same individual structural group in different molecules, the elemental contribution keeps the same.[2] In addition to the advantages of simplicity, universality and quickness,[1] GC methods can essentially reduce the computational cost, even for a reaction network with thousands of chemical species and elementary reactions, because the number of functional groups constituting these molecular species is limited.[3]

[a] *National Engineering Laboratory for Methanol-to-Olefins, Dalian National Laboratory for Clean Energy, Dalian Institute of Chemical Physics, Chinese Academy of Sciences, 116023 Dalian, China. E-mail: lihua@dicp.ac.cn, maoye@dicp.ac.cn*

[b] *University of Chinese Academy of Sciences, 10049 Beijing, China*

† Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d3cp01719b

This journal is © the Owner Societies 2023

Phys. Chem. Chem. Phys., 2023, **25**, 21631–21639 | **21631**

The first application of GC methods can date back to 1944.[4] The most widely used group additivity method for estimating thermodynamic properties was proposed by Benson and Buss in 1958. Since then, many other GC methods have been developed and used in different applications.[1–3,5–8] For example, GC methods have been extended to estimate the thermodynamic properties such as free energy change,[9] adsorption isotherms,[10] enthalpy of formation and Gibbs free energy,[11] enthalpies of formation and entropies,[10,12] to name a few. It has been shown that GC methods can achieve satisfactory accuracy in estimating thermodynamic properties for a wide variety of chemical species.

Methanol-to-olefins (MTO) is an important chemical process that has received considerable attention in the past decades, owing to that MTO can convert methanol, a chemical readily produced from many non-oil resources such as coal, syngas, biomass and CO$_2$, to ethylene and propylene, the basic materials for manufacturing polymers. In the industrial MTO process SAPO-34 molecular sieves are used as the catalyst due to their low activation energy barrier and unique shape-selectivity effect.[13,14] It has been shown that MTO follows the dual-cycle hydrocarbon pool mechanism,[15,16] which makes the reaction network of MTO quite complicated and involves hundreds of elementary reactions and various carbocation active intermediates. The thermodynamic properties of elementary reactions in MTO are not only dependent on the guest molecular species engaged in the reaction network, but also relevant to the interplay between guest molecules and active sites of the catalyst. The guest molecules can be absorbed at the acidic sites. Compared with gas phase reactions, the reaction energy barriers in molecular sieves are greatly reduced and thus the reaction rate and product selectivity have been enhanced.[17] Martinis utilized a GC method to calculate the thermodynamic data of hydrocarbon species over molecular sieves,[18] in which the carbocation, alkene, and their ortho carbon atoms were considered as groups with second-order contributions. However, he neglected the group interactions between the guest molecules and active sites of the catalyst. Solely summing up the contributions of individual groups of guest molecules might result in inaccurate calculation of thermodynamic data in molecular sieve catalyzed processes.[3,12,19]

In this work, we propose a modified group contribution (mGC) method to calculate the thermodynamic data of species encountered in the MTO process over the SAPO-34 catalyst, in which the interaction between guest molecules and active sites in the molecular sieves will be considered. This is partly inspired by the recent work by He et al.,[3] in which they extended the GC methods to account for the intramolecular interactions. The intramolecular interactions are incorporated as an exponential decay function of group distance, which is defined as the number of bonds in between. However, He et al. only considered interactions between different guest molecular groups. In our proposal, we treat the active sites, which are essentially the acidic centers in molecular sieves, as independent functional groups in addition to the

groups of guest molecules in the reaction network. In the thermodynamic data calculations, these independent active site functional groups (IASFGs) can interact with the groups of guest molecules. By doing so, all the contributions from individual groups of guest molecules, interactions between groups of guest molecules, and interactions between IASFGs and groups of guest molecules are taken into account. The purpose of this work is to develop a semi-empirical model for rapidly calculating the thermodynamic properties for all reactants, products, and intermediate species in catalytic processes over molecular sieves. Our ultimate goal is to establish an intelligent micro-kinetics generation approach, which requires the establishment of a detailed reaction network, as well as a model for quickly obtaining thermodynamic and kinetic parameters. To this end, in this paper we focus on the model for obtaining the thermodynamic parameters, including the reaction entropy and enthalpy, for all species and the corresponding reactions.

2. Method

2.1 Group contribution methods

The GC method represents an approximate calculation approach based on the additive nature of molecular properties. It divides the molecule into different groups, and then adds up the properties of each group. GC methods can be applied to predict various physical and thermodynamic properties of molecules. The commonly used group classification principle is Benson's method, in which the contribution of atoms or atomic groups is determined by taking the atom or atomic group as the starting point and considering the property differences of the connected atoms or atomic groups in turn. GC methods can be used to estimate the standard formation enthalpy, standard absolute entropy and standard molar heat of ideal gaseous organic compounds by

$$Y_f = \sum_{i=1}^{n_c} ng_i \cdot gc_i \qquad (1)$$

where Y_f represents the thermodynamic parameters (such as enthalpy H_f, entropy S_f, etc.), gc_i is the energy of each group i, n_c is the number of group categories, and ng_i is the number of segments of each group i. For example, in Fig. 1, 1-butene has four carbon atoms, so it can be divided into four groups, i.e., "CH$_2$=", "CH=", "CH$_2$" and "CH$_3$". Note that there is only one segment for each group in the 2-butene; eqn (1) can be written as $Y_f = gc_{"CH_2="} + gc_{"CH="} + gc_{"CH_2"} + gc_{"CH_3"}$.

2.2 Modified group contribution method

For molecular sieves, the interactions between IASFGs and groups of guest molecules should be taken into account in thermodynamic data calculations. In this article, IASFGs are related to the locations of active sites or the corresponding acid sites of the molecular sieves. In the real calculations, it represents the first group which is closest to the molecular sieves, such as [CH$_2$+] in Fig. 1(b). Assuming that y_{MS}

21632 | Phys. Chem. Chem. Phys., 2023, 25, 21631–21639

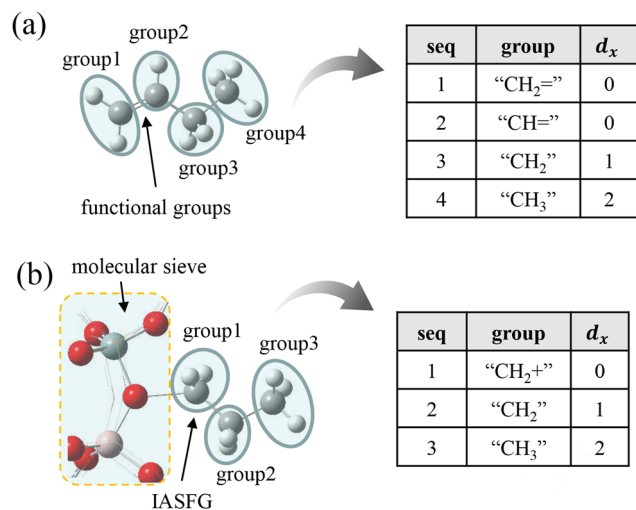This journal is © the Owner Societies 2023

(a)

(b)

Fig. 1 Schematic diagram of the mGC method for different functional groups: (a) carbon–carbon double bond and (b) independent active site functional groups (IASFGs).

represents the influence brought by molecular sieves and y_{MOL} represents the influence of guest molecular functional groups, we obtain

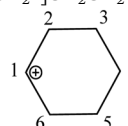$$Y_f = \sum_{i=1}^{n_c} ng_i \cdot gc_i + y_{MS} + y_{MOL} \qquad (2)$$

$$y_{MS} = P_{[C]} \sum_x \left(f_{[C]} \right)^{d_x} \qquad (3)$$

$$y_{MOL} = P_{[O]} \sum_x \left(f_{[O]} \right)^{d_x} + P_{[N]} \sum_x \left(f_{[N]} \right)^{d_x} \qquad (4)$$

Here P is the correction coefficient describing the interaction between groups, f is the distance factor which can be obtained as an empirical parameter, and d_x is the topological distance that describes the number of atoms between atom x and the corresponding functional group. For functional groups with subscripts of P and f, "[C]" represents the carbenium of IASFGs, "[O]" represents the carbon–carbon double bond, and "[N]" represents the ring. The number of atoms in between two different atoms is defined as the topological distance of these two atoms. For a given guest molecule, the interaction between each group constituting this molecule and the active sites or functional groups is closely related to the distance between them. The closer the distance, the stronger the interaction. Therefore, a distance correction should be considered in GC methods. In this work, a power distance function, $\sum_x f^{d_x}$, is used to describe the influence of the active sites or functional groups of molecular sieves on guest molecules. For example, in Fig. 1(a), 1-butene has a functional group of carbon–carbon double bond, so d_x values of groups 1 and 2 are both 0 and d_x values of groups 3 and 4 are 1 and 2, respectively. In Fig. 1(b), the IASTG corresponding to the propylene carbocation in molecular sieves is "CH$_2$+", so d_x values of groups 1, 2 and 3

Table 1 Calculation of the groups and corresponding topological distances of guest molecules

Name	ng				d_x
	C–CH$_3$	C–C$_2$H$_2$	Cdb–CH	C + –CH$_2$	Topological distance
CH$_3$CH$_2$CH$_2$CH$_3$	2	2	0	0	0
CH$_3$CH=CHCH$_3$	2	0	2	0	1, 0, 0, 1
[CH$_2$+]CH$_2$CH$_2$CH$_3$	1	2	0	1	0, 1, 2, 3
[hexagon structure, positions 1–6]	Cc–C$_2$H$_2$		Cc + –C$_2$H		0, 1, 1, 2, 2, 3
	5		1		

are 0, 1, and 2, respectively. Among them, ng and d_x need to be obtained from the topological structure and properties of the guest molecule of interest, while the remaining parameters are obtained by fitting.

The number of segments of the same group within guest molecules, ng, can be obtained from the simplified structural formula. Each carbon atom and its connected hydrogen atom can be considered as a group. For example, as shown in Table 1, n-butane contains two methylene groups and two methyl groups. Specifically, 2-butene contains two methylene groups and two methyl groups, and the corresponding d_x values of these four groups are 1, 0, 0, and 1, respectively. Meanwhile 1-butene carbocation contains one 'C + –CH$_{2+}$' group, two methylene groups, and one methyl group, and in 1-butene carbocation the corresponding d_x values of the groups are 0, 1, 2, and 3, respectively.

3. Thermodynamic data

3.1 Methanol to olefins reaction network

Methanol to olefins (MTO) opens up a new way for the utilization of non-oil resources to produce light olefins for alleviating the shortage of oil resources.[20,21] As discussed above, the reaction network of the MTO process over SAPO-34 molecular sieves is extremely complex, containing numerous intermediate species and elementary reaction steps. To illustrate the effectiveness of the mGC method in calculating the thermodynamic parameters, we consider the initial stage of MTO reaction as an example in the following.

The reaction mechanism of MTO has been well investigated, and the initial stage of MTO over SAPO-34 molecular sieves mainly involves alkane and alkene conversion as well as ring formation.[22,23] The main species appeared at this stage are alkanes (R, gas phases), olefins (O, gas phases), carbocations (C, chemisorption state), olefin carbocations (OC, chemisorption state), five-membered ring carbocations (N5C, chemisorption state), and six-membered ring carbocations (NC, chemisorption state), as shown in Table S1 (in the ESI†). The reactions occurring at the initial stage of MTO include protonation (PRO), deprotonation (DEP), methyl shift (MSI), PCP branching (PCP), methylation (MET), alkylation (ALK), β-scission (BSC), hydrogen transfer (HTR), cyclization-five (CYC1) and

This journal is © the Owner Societies 2023

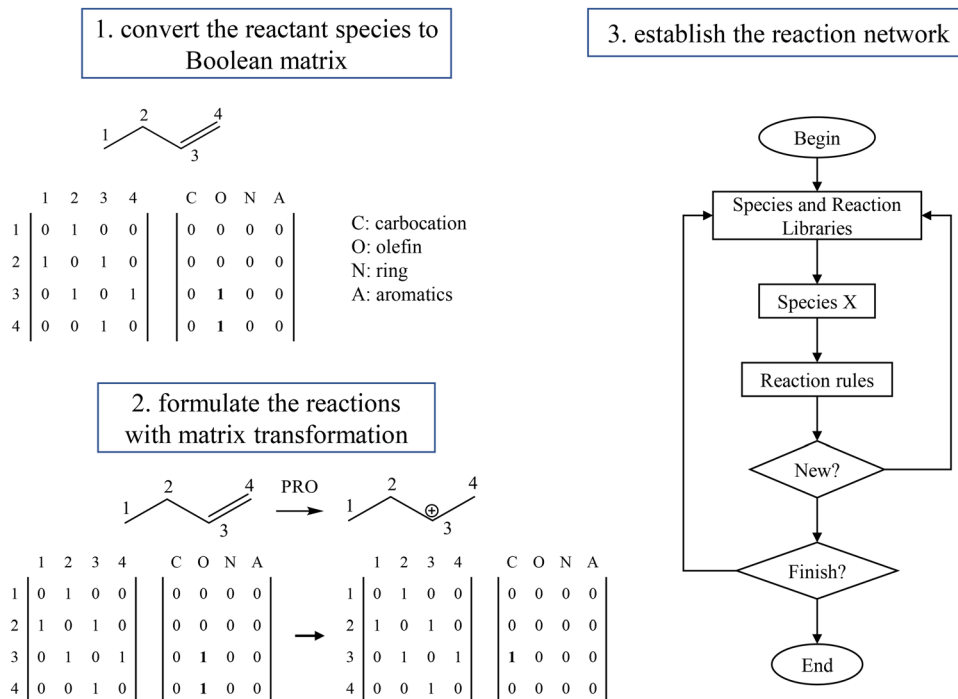Phys. Chem. Chem. Phys., 2023, **25**, 21631–21639 | **21633**

Fig. 2 A method for automatic generation of reaction networks.

cyclization-six (CYC2), as listed in Table S2 (in the ESI†). An approach of automatic generation of reaction networks based on the established reaction mechanism has been developed in a parallel work (Fig. 2, see ESI† for detailed procedures). In the current work, we use this approach to generate the reaction network. The approach consists of three steps: converting the reactant species to a Boolean matrix, formulating the reactions with matrix transformation, and establishing the reaction network.

For simplicity, we define a maximum number of carbon atoms in the guest molecules to limit the indefinite expansion of species number in the MTO process. We argue that this is reasonable as our main purpose here is to illustrate the effectiveness of the mGC method. The corresponding thermodynamic properties (including enthalpy and entropy) of all species considered are obtained at four temperatures (698, 723, 748, and 763 K) based on the first principles density functional theory (DFT) calculation, which are defined as the reference data in this work. We first assumed that the maximum number of carbon atoms is 7, which leads to the 186 species and 753 reactions enclosed in the initial stage of the MTO process which were used as the training set. Based on the molecular structures of 186 guest molecules in the training set,

we identified all possible groups constituting guest molecules, and derived the thermodynamic properties of each group by data fitting. Then, we calculated the thermodynamic data of the training set using both the cGC (eqn (1)) and mGC (eqn (2)–(4)) methods, and made a detailed comparison. To further examine the reliability of the mGC method, we also extended the maximum carbon atom number to 8, and selected 65 additional species and 100 additional reactions from the reaction network, which were used as the testing set. The thermodynamic properties of these extra species and reactions were calculated using the mGC method based on the group parameters of the training set. The detailed analysis is in Sections 4.1 and 4.2.

3.2 Thermodynamic database by density functional theory

As mentioned above, the thermodynamic properties (including enthalpy and entropy) of all 186 species in the training set and 65 species in the testing set were derived based on the DFT calculation, and for each species four temperatures (698 K, 723 K, 748 K, and 763 K) were considered. We used well-developed software to obtain the thermodynamic parameters of species at different temperatures. Specifically (Fig. 3), we first used Gauss
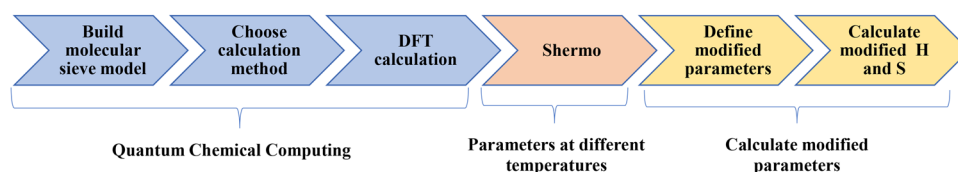

Fig. 3 Flowchart for calculating thermodynamic parameters.

21634 | *Phys. Chem. Chem. Phys.*, 2023, **25**, 21631–21639

This journal is © the Owner Societies 2023

View 5.0 software to edit the structure of species, and then used Gaussian 09W and Shermo 2.0.5 software to calculate the thermodynamic parameters of species at different temperatures.[24–26] Finally, in order to simplify calculations and improve model fitting, we propose the utilization of modified thermodynamic parameters. This approach is particularly beneficial when dealing with thermodynamic values obtained from DFT calculations, which tend to be excessively large and pose challenges for effective fitting.

The detailed procedure is as follows:

(1) Calculating the absolute enthalpy and entropy of species at a specified temperature. For gas phase species, we directly edited the structure through Gauss View 5.0 and calculated it with Gaussian 09W. For molecular sieve adsorption species, the H-SAPO-34 molecular sieve structure was derived by selecting the extended 74T($SiP_{36}Al_{37}O_{119}H_{59}$)[27] structure extracted from CHA crystals, in which the acidic sites were located in the 8-MR window[28] (Fig. 4).

After the structure was defined, we then obtained the thermodynamic parameters of adsorbed species of H-SAPO-34. In doing so, we first conducted geometry optimization and frequency calculations via ONIOM (ωB97XD/6-31G(d,p):am1), and then derived energy via ωB97XD/6-311G (d,p).[29–31] In Gaussian 09W calculations, the results for a certain temperature T can be stored in "*.log file", which includes single point energy ε_{ele}, enthalpy correction H_{corr}, and Gibbs free energy correction G_{corr}. Finally, we obtained the absolute enthalpy H and entropy S of the species at this temperature via

$$H(T) = \varepsilon_{ele} + H_{corr}(T) \tag{5}$$

$$G(T) = \varepsilon_{ele} + G_{corr}(T) \tag{6}$$

$$S(T) = \frac{H(T) - G(T)}{T} \tag{7}$$

(2) Calculating the absolute enthalpy and entropy of species at other temperatures. By use of Shermo 2.0.5 software, we can quickly calculate thermodynamic parameters at other

temperatures based on the "*.log file" for one specified temperature through the partition function methods.

(3) Generating the modified thermodynamic parameters of species based on the absolute data. Usually, the absolute enthalpy or entropy of a given species calculated from different methods may vary because the standard conditions could be different.

For convenience, we defined the modified parameters of formation Y_f^0 (including the modified enthalpy of formation H_f^0 and modified entropy of formation S_f^0), and calculated the modified parameters of each species "C_xH_y" as well as the modified parameters of carbon atoms $Y_f^0(C,g)$ and hydrogen molecules $Y_f^0(H_2,g)$.

$$Y_f^0(C_xH_y) = Y_f^0(C_xH_y, g) - x \times Y_f^0(C, g)$$
$$- \frac{y}{2} \times Y_f^0(H_2, g) \tag{8}$$

Taking the reaction " $ZH + CH_2{=}CH_2 \rightarrow ZH{-}CH_2CH_3$ (723 K)" as an example, we have

$$Y_f^0(ZHC_xH_y) = Y_f^0(ZHC_xH_y) - Y_f^0(ZH)$$
$$- x \times Y_f^0(C, g) - \frac{y}{2} \times Y_f^0(H_2, g) \tag{9}$$

Note that the single point energy ε_{ele}, enthalpy correction H_{corr}, and Gibbs free energy correction G_{corr} are all calculated by GAUSS 09W, and the absolute enthalpy of the species (unit conversion: 1 H/P $\approx$ 2625.5 kJ mol^{-1}) can be obtained through formulas (5) and (6). Then we can calculate the modified enthalpy of formation H_f^0 by formulas (8) and (9). We can calculate the modified enthalpy of formation of ethylene by $H_f^0(CH_2{=}CH_2) = H(CH_2{=}CH_2, g) - 2 \times H(C,g) - 2 \times H(H_2,g)$. The modified formation enthalpy calculation of ethylene carbocation is thus $H_f^0(ZHCH_2CH_3) = H(ZHCH_2CH_3) - H(ZH) - 2 \times H(C,g) - 2 \times H(H_2,g)$.

4. Results and discussion

4.1 Training process

In the training process, we first identified the groups using the 186 species in the training set and then calculated the group parameters, including distance factor f, group energy parameter gc and distance coefficient P. In this work, 15 groups which can be classified as four types (i.e., the alkane groups, alkene groups, carbocation groups, and ring groups) were identified. The alkane groups were further divided into methyl group (C–CH$_3$) and methylene group (C–C$_2$H$_2$, C–C$_3$H and C–C$_4$). The alkene groups, according to the number of hydrogens, were divided into groups Cdb–H$_2$, Cdb–CH and Cdb–C$_2$. Similarly, the carbocation groups were divided into groups C + –CH$_2$, C + –C$_2$H and C + –C$_3$. The ring groups include groups Cc–C$_2$H$_2$, Cc–C$_3$H, Cc–C4, Cc + –C$_2$H, and Cc + –C$_3$.

As shown in Table 2, in order to describe the effects of active sites and functional groups of molecular sieves on the thermodynamic properties of guest species, for different guest molecules, three distance factors ($f_{[O]}, f_{[C]}$ and $f_{[NC]}$) and five distance
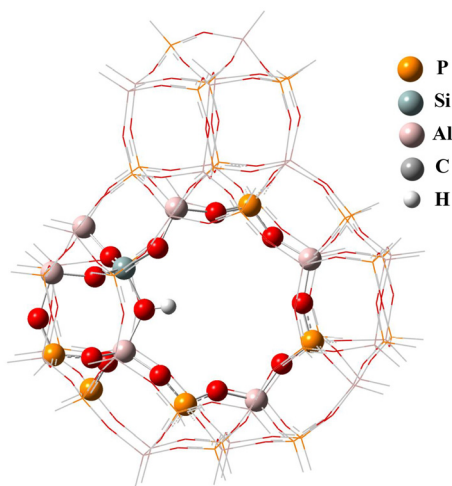

Fig. 4 H-SAPO-34(74T) model.

This journal is © the Owner Societies 2023

Phys. Chem. Chem. Phys., 2023, **25**, 21631–21639 | **21635**

Table 2 Calculation models for different guest species

Species	Calculation model
R	$Y_{\mathrm{f}}^{0} = \sum_{i=1}^{n_c} \mathrm{ng}_i \cdot \mathrm{gc}_i$
O	$Y_{\mathrm{f}}^{0} = \sum_{i=1}^{n_c} \mathrm{ng}_i \cdot \mathrm{gc}_i + P_{[\mathrm{O}]} \sum_x \left(f_{[\mathrm{O}]} \right)^{d_x}$
C	$Y_{\mathrm{f}}^{0} = \sum_{i=1}^{n_c} \mathrm{ng}_i \cdot \mathrm{gc}_i + P_{[\mathrm{C}]} \sum_x \left(f_{[\mathrm{C}]} \right)^{d_x}$
OC	$Y_{\mathrm{f}}^{0} = \sum_{i=1}^{n_c} \mathrm{ng}_i \cdot \mathrm{gc}_i + P_{[\mathrm{C\_OC}]} \sum_x \left(f_{[\mathrm{C}]} \right)^{d_x} + P_{[\mathrm{O\_OC}]} \sum_x \left(f_{[\mathrm{O}]} \right)^{d_x}$
N5C	$Y_{\mathrm{f}}^{0} = \sum_{i=1}^{n_c} \mathrm{ng}_i \cdot \mathrm{gc}_i + P_{[\mathrm{NC}]} \sum_x \left(f_{[\mathrm{NC}]} \right)^{d_x}$
NC	$Y_{\mathrm{f}}^{0} = \sum_{i=1}^{n_c} \mathrm{ng}_i \cdot \mathrm{gc}_i + P_{[\mathrm{NC}]} \sum_x \left(f_{[\mathrm{NC}]} \right)^{d_x}$

coefficients ($P_{[\mathrm{O}]}$, $P_{[\mathrm{C}]}$, $P_{[\mathrm{O\_OC}]}$, $P_{[\mathrm{C\_OC}]}$ and $P_{[\mathrm{NC}]}$) were further considered.

$$ \mathrm{RMSE} = \sqrt{\frac{1}{m} \sum_{i=1}^{m} (y_i' - y_i)^2} \qquad (10) $$

$$ \mathrm{MAE} = \frac{1}{m} \sum_{i=1}^{m} |y_i' - y_i| \qquad (11) $$

To illustrate the procedure, we take the training set as an example (in which the temperature is set as 723 K), and the detailed results are included in the ESI.† The fitting process is conducted by two steps, and root mean square error (RMSE) and mean absolute error (MAE) are used to describe the fitting effect (as shown in eqn (10) and (11)), where m is the sample size, y_i' is the model predicted value of sample i, and y_i is the reference value. In the first step, the initial group parameters are obtained by fitting according to eqn (1) with DFT data as reference data. RMSE as a function of distance correction factor is shown in Fig. 5. The distance factor is determined by that corresponding to the minimum RMSE for either enthalpy or entropy. In the following step, according to the reference enthalpy and entropy of the guest species by DFT calculation, the thermodynamic parameters of different groups as well as the corresponding distance factors were obtained by least square fitting, as shown in Table S3 (in the ESI†).

We used both the cGC and mGC methods to estimate the thermodynamic parameters for 186 species using the reference data and group parameters obtained above. We examined the results carefully, and analyzed the distribution of deviations of thermodynamic parameters calculated by mGC and cGC methods for all the species (see Fig. 6). Compared with the traditional group contribution method, the important improvement of mGC lies in the number distribution of average error and the root mean square error has been essentially narrowed. This means that the estimation of thermodynamic parameters has been essentially improved for most of the species by use of mGC. Meanwhile mGC can realize thermodynamic parameter calculation for molecular sieve adsorbed species. Furthermore, mGC is more physics-sound as it takes into account all contributions from intramolecular functional groups as well as interactions between the zeolite and the guest molecule.

We have calculated the enthalpy change and entropy change of each reaction at the initial stage of MTO, and analyzed the deviation between the estimated thermodynamic parameters and corresponding reference data for the training set. In Fig. 7, in total 753 reactions classified into 10 types are shown, with the ordinate representing the DFT data and the abscissa
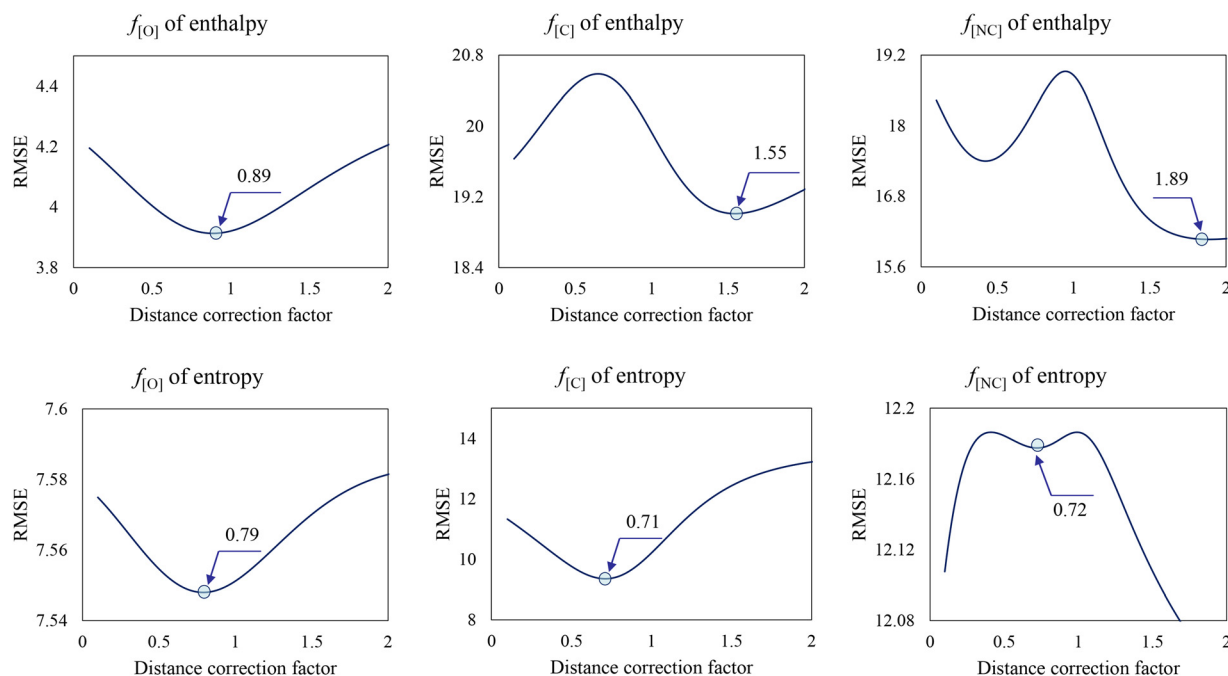

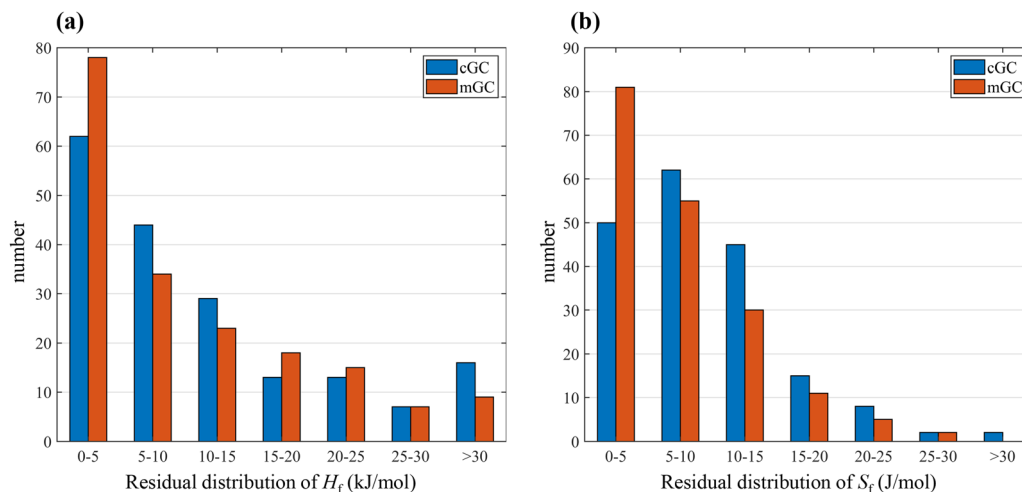Fig. 5 Distance factor f of enthalpy and entropy.

21636 | *Phys. Chem. Chem. Phys.*, 2023, **25**, 21631–21639

This journal is © the Owner Societies 2023

Fig. 6 Comparison of calculation error between cGC and mGC: (a) H_f and (b) S_f.
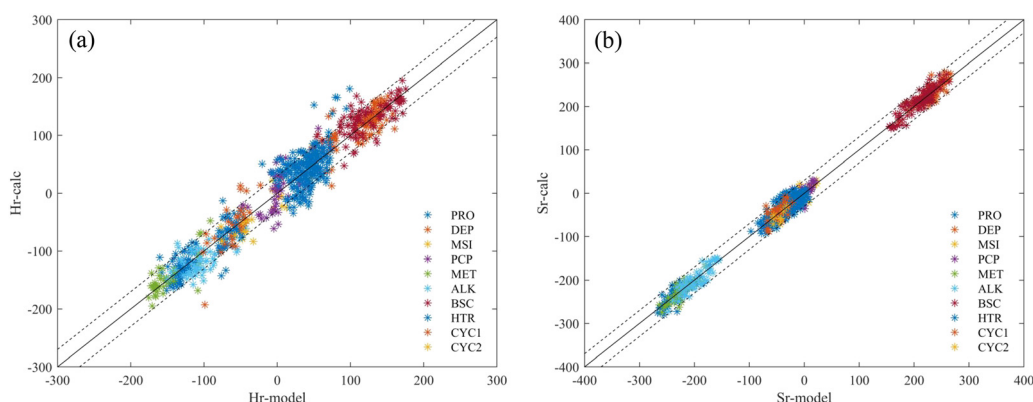

Fig. 7 Comparison of enthalpy change (a) and entropy change (b) of reactions estimated by the mGC method and reference data by DFT calculation for the training set. The dashed lines represent the deviation of ± 30 kJ for enthalpy change (a) and ± 30 J for entropy change (b).

representing the estimated parameters using the mGC method. Note that the dotted lines represent the deviation of ± 30 kJ in Fig. 7(a) and ± 30 J in Fig. 7(b); it can be concluded that both the enthalpy change and entropy change of reactions estimated by the mGC method agree well with the reference data by DFT calculations. As shown in Table 3, the MAE of the enthalpy change of reactions in the training set, which was obtained by comparing the data estimated by the mGC method with the reference data by DFT, is about 18.80 kJ mol^{-1}, while the MAE

of the entropy change is about 10.84 J mol^{-1}. Apparently, the mGC method is superior over the cGC methods in terms of enthalpy and entropy change calculations for the initial stage of MTO over molecular sieves.

4.2 Testing process

To further check the reliability and applicability of the proposed mGC method, we assumed the maximum number of carbon atoms to be 8 in the initial stage of MTO over SAPO-34 molecular sieves. Thus, we obtained 100 extra reactions with 65 extra species. Again, to ensure the accuracy of results, we used DFT described in Section 3.2 to calculate the thermodynamic properties of these extra species and reactions which were used for evaluating the performance of the mGC method. In estimating the thermodynamic properties of these extra species and reactions, we used the information of groups and distance parameters obtained in the training process. As listed in Table 3, MAEs of both enthalpy and entropy of these extra species relative to the reference data are 17.31 kJ mol^{-1} and 14.19 J mol^{-1}, respectively. Meanwhile, we can observe that

Table 3 Error analysis of thermodynamic parameters by the mGC method and DFT calculations

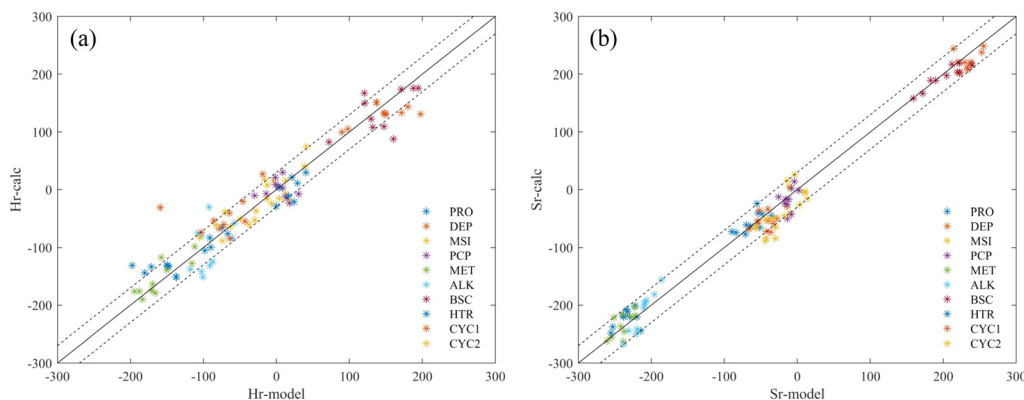| Data		Number	MAE		RMSE	
			H_f (kJ mol^{-1})	S_f (J mol^{-1})	H_f (kJ mol^{-1})	S_f (J mol^{-1})
Training set	Species	186	10.62	7.12	15.49	9.08
	Reaction	753	18.80	10.84	24.23	13.68
Testing set	Species	65	17.31	14.19	25.45	16.74
	Reaction	100	22.43	16.25	29.18	19.76

Fig. 8 Comparison of enthalpy change (a) and entropy change (b) of reactions estimated by the mGC method and reference data calculated by DFT for the testing set. The dashed lines represent the deviation of ± 30 kJ for enthalpy change (a) and ± 30 J for entropy change (b).

RMSEs of enthalpy and entropy of the extra species are, respectively, 25.45 kJ mol^{-1} and 16.74 J mol^{-1}.

Similarly, we calculated the enthalpy change and entropy change of these 100 extra reactions. We found that MAEs, as shown in Table 3, are 22.43 kJ mol^{-1} and 16.25 J mol^{-1} for the enthalpy change and entropy change, respectively. The RMSE of enthalpy change of reactions relative to the reference data is 29.18 kJ mol^{-1}, and the RMSE of the entropy change is 19.76 J mol^{-1}. Fig. 8 compares the reaction enthalpy change and entropy change estimated by the mGC method and reference data calculated by DFT. Note that the dotted lines represent the deviation of ± 30 kJ in Fig. 8(a) and ± 30 J in Fig. 8(b), respectively; it can be concluded that both the enthalpy change and entropy change of reactions estimated by the mGC method agree well with that derived by DFT calculations. It can be argued that, within a certain error range, the rapid calculation of thermodynamic parameters of guest species in molecular sieves can be achieved by the proposed mGC method.

5. Conclusion

For the catalytic process over molecular sieves, we develop a semi-empirical mGC method for rapid estimation of thermodynamic parameters of guest species in either the gas phase or inside molecular sieves. In doing so, the effects of active sites and functional groups of molecular sieves, specifically the interaction between the guest molecular groups and the active centers or functional groups in the molecular sieves, were first considered. Then, the initial stage of MTO over SAPO-34 molecular sieves was studied as an example to illustrate the effectiveness of the proposed mGC method. A total of 186 species with a carbon atom number less than 7 were selected as the training set, and 65 species with a carbon atom number of 8 were used as the test set. The thermodynamic parameters (including enthalpy and entropy) of these guest species at different temperatures (698, 723, 748, and 763 K), which were used as the reference data in the current study, were derived by use of DFT calculation with the ωB97XD/6-31G (d,p) model. The comparison between the cGC and mGC methods shows that the

mGC method has higher prediction accuracy, even extending to other guest molecules with higher carbon atom number. The MAEs of enthalpy and entropy estimated by mGC for guest molecules with a carbon atom number of 8 are only 17.31 kJ mol^{-1} and 14.19 J mol^{-1}, respectively. In fact, the mGC method can be applied for rapid and batch calculation of thermodynamic properties of guest molecules in a complex catalytic reaction network, which is essential for automatic generation of reaction kinetics.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

References

1 Z. Kolska, M. Zabransky and A. Randov, *Thermodynamics – Fundamentals and Its Application in Science*, IntechOpen, London, UK., 2012, ch. 6, DOI: **10.5772/49998**.

2 A. Salmon and D. Dalmazzone, *J. Phys. Chem. Ref. Data*, 2007, **36**, 19–58.

3 T. He, S. Li, Y. Chi, H. B. Zhang, Z. Wang, B. Yang, X. He and X. You, *Phys. Chem. Chem. Phys.*, 2016, **18**, 23822–23830.

4 J. Anderson, G. Beyer and K. Watson, *News Tech. Sect.*, 1944, **36**, R476–484.

5 S. W. Benson and J. H. Buss, *J. Chem. Phys.*, 1958, **29**, 546–572.

6 Y. Yoneda, *Bull. Chem. Soc. Jpn.*, 1979, **52**, 1297–1314.

7 T. Thinh, J. L. Duran and R. S. Ramalho, *Ind. Eng. Chem. Process Des. Dev.*, 1971, **10**, 576–582.

8 L. Constantinou and R. Gani, *AIChE J.*, 1994, **40**, 1697–1710.

9 C. A. Mora Osorio and A. F. Gonzalez Barrios, *BMC Bioinf.*, 2016, **17**, 522.

21638 | *Phys. Chem. Chem. Phys.*, 2023, **25**, 21631–21639

This journal is © the Owner Societies 2023

10 R. Z. Li, J. M. Herreros, A. Tsolakis and W. Z. Yang, *Fuel*, 2021, **304**, 121437.

11 E. Potrich, F. A. P. Voll, V. F. Cabral and L. Cardozo, *Chem. Ind. Chem. Eng. Q.*, 2019, **25**, 153–162.

12 F. vom Lehn, L. Cai, B. Copa Cáceres and H. Pitsch, *Combust. Flame*, 2021, **232**, 111525.

13 A. Tanimu, G. Tanimu, H. Alasiri and A. Aitani, *Energy Fuels*, 2022, **36**, 5152–5166.

14 M. Yang, D. Fan, Y. Wei, P. Tian and Z. Liu, *Adv. Mater.*, 2019, **31**, e1902181.

15 M. Westgård Erichsen, S. Svelle and U. Olsbye, *Catal. Today*, 2013, **215**, 216–223.

16 K. De Wispelaere, K. Hemelsoet, M. Waroquier and V. Van Speybroeck, *J. Catal.*, 2013, **305**, 76–80.

17 A. Corma, *Chem. Rev.*, 1997, **97**, 2373–2420.

18 J. M. Martinis and G. F. Froment, *Ind. Eng. Chem. Res.*, 2006, **45**, 954–967.

19 E. R. Ritter and J. W. Bozzelli, *Int. J. Chem. Kinet.*, 1991, **23**, 767–778.

20 P. Tian, Y. X. Wei, M. Ye and Z. M. Liu, *ACS Catal.*, 2015, **5**, 1922–1938.

21 M. Ye, H. Li, Y. Zhao, T. Zhang and Z. Liu, *Mesoscale Modeling in Chemical Engineering Part II*, Elsevier, 2015, vol. 47, pp. 279–335.

22 C.-M. Wang, Y.-D. Wang and Z.-K. Xie, *J. Catal.*, 2013, **301**, 8–19.

23 S. M. Alwahabi and G. F. Froment, *Ind. Eng. Chem. Res.*, 2004, **43**, 5098–5111.

24 W. N. Zhang, Y. C. Zhi, J. D. Huang, X. Q. Wu, S. Zeng, S. T. Xu, A. M. Zheng, Y. X. Wei and Z. M. Liu, *ACS Catal.*, 2019, **9**, 7373–7379.

25 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, Gaussian, Inc., Wallingford CT, 2013.

26 T. Lu and Q. X. Chen, *Comput. Theor. Chem.*, 2021, **1200**, 113249.

27 P. J. O'Malley and J. Dwyer, *Zeolites*, 1988, **8**, 317–321.

28 D. Lesthaeghe, B. De Sterck, V. Van Speybroeck, G. B. Marin and M. Waroquier, *Angew. Chem., Int. Ed.*, 2007, **46**, 1311–1314.

29 Y. Y. Chu, B. Han, A. M. Zheng and F. Deng, *J. Phys. Chem. C*, 2012, **116**, 12687–12695.

30 D. K. Papayannis, K. D. Papavasileiou and V. S. Melissas, *Microporous Mesoporous Mater.*, 2016, **226**, 1–9.

31 J. D. Chai and M. Head-Gordon, *Phys. Chem. Chem. Phys.*, 2008, **10**, 6615–6620.

This journal is © the Owner Societies 2023

Phys. Chem. Chem. Phys., 2023, **25**, 21631–21639 | **21639**